HANDBOOK OF SOCIOLOGICAL SCIENCE: CONTRIBUTIONS TO RIGOROUS SOCIOLOGY

RESEARCH HANDBOOKS IN SOCIOLOGY

Series Editor: Hans-Peter Blossfeld, *Professor of Sociology, University of Bamberg, Germany*

The Research Handbooks in Sociology series provides an up-to-date overview on the frontier developments in current sociological research fields. The series takes a theoretical, methodological and comparative perspective to the study of social phenomena. This includes different analytical approaches, competing theoretical views and methodological innovations leading to new insights in relevant sociological research areas. Each *Research Handbook* in this series provides timely, influential works of lasting significance. These volumes will be edited by one or more outstanding academics with a high international reputation in the respective research field, under the overall guidance of series editor Hans-Peter Blossfeld, Professor of Sociology at the University of Bamberg. The *Research Handbooks* feature a wide range of original contributions by well-known authors, carefully selected to ensure a thorough coverage of current research. The *Research Handbooks* will serve as vital reference guides for undergraduate students, doctoral students, postdoctorate students and research practitioners in sociology, aiming to expand current debates, and to discern the likely research agendas of the future.

Titles in the series include:

Research Handbook on the Sociology of Education
Edited by Rolf Becker

Research Handbook on the Sociology of the Family
Edited by Norbert F. Schneider and Michaela Kreyenfeld

Research Handbook on Environmental Sociology
Edited by Axel Franzen and Sebastian Mader

Research Handbook on Analytical Sociology
Edited by Gianluca Manzo

Handbook of Sociological Science
Contributions to Rigorous Sociology
Edited by Klarita Gërxhani, Nan Dirk de Graaf and Werner Raub

Handbook of Sociological Science

Contributions to Rigorous Sociology

Edited by

Klarita Gërxhani

Professor of Sociology, Department of Political and Social Sciences, European University Institute (EUI), Italy

Nan Dirk de Graaf

Professor of Sociology and Official Fellow, Nuffield College, University of Oxford, UK

Werner Raub

Professor of Sociology, Department of Sociology/ICS, Utrecht University, the Netherlands

RESEARCH HANDBOOKS IN SOCIOLOGY

Edward Elgar PUBLISHING

Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2022934532

This book is available electronically in the **Elgar**online
Sociology, Social Policy and Education subject collection
http://dx.doi.org/10.4337/9781789909432

ISBN 978 1 78990 942 5 (cased)
ISBN 978 1 78990 943 2 (eBook)

Typeset by Cheshire Typesetting Ltd, Cuddington, Cheshire

Contents

Contributors

Katrin Auspurg, Department of Sociology, Ludwig-Maximilians-Universität, Munich (Chapter 26)

Rolf Becker, Institute of Educational Science, University of Bern (Chapter 19)

Richard Breen, Nuffield College and Department of Sociology, University of Oxford (Chapters 15 and 18)

Josef Brüderl, Department of Sociology, Ludwig-Maximilians-Universität, Munich (Chapter 26)

Ronald S. Burt, Booth School of Business, University of Chicago, and Department of Management and Technology, Bocconi University (Chapter 20)

Vincent Buskens, Department of Sociology/ICS, Utrecht University (Chapter 9)

Rense Corten, Department of Sociology/ICS, Utrecht University (Chapter 9)

Nan Dirk de Graaf, Nuffield College, University of Oxford (Chapters 1 and 24)

Andreas Diekmann, Department of Humanities, Social and Political Science, ETH Zürich, and Institute of Sociology, University of Leipzig (Chapter 6)

Joseph Dippong, Department of Sociology, University of North Carolina at Charlotte (Chapter 13)

Peter Sheridan Dodds, Department of Computer Science, University of Vermont (Chapter 22)

Ivan Ermakoff, Department of Sociology, University of Wisconsin-Madison (Chapter 11)

Marie Evertsson, Swedish Institute for Social Research, Stockholm University (Chapter 10)

Andreas Flache, Department of Sociology/ICS, University of Groningen (Chapter 4)

Markus Gangl, Institute of Sociology, Goethe University Frankfurt am Main, and Department of Sociology, University of Wisconsin-Madison (Chapter 16)

Klarita Gërxhani, Department of Political and Social Sciences, European University Institute (Chapters 1 and 17)

John H. Goldthorpe, Nuffield College, University of Oxford (Chapter 18)

Rosemary L. Hopcroft, Department of Sociology, University of North Carolina at Charlotte (Chapter 13)

Michelle Jackson, Department of Sociology, Stanford University (Chapter 2)

Mads Meier Jæger, Department of Sociology, University of Copenhagen (Chapter 7)

Eva Jaspers, Department of Sociology/ICS, Utrecht University (Chapter 10)

Rachel Kail, School of Criminal Justice, University of Cincinnati (Chapter 13)

Frank Kalter, School of Social Sciences, University of Mannheim University (Chapter 8)

Marijn A. Keijzer, Institute for Advanced Study in Toulouse, University of Toulouse 1 Capitole (Chapter 4)

Hexuan Liu, School of Criminal Justice, University of Cincinnati (Chapter 13)

Gianluca Manzo, Department of Sociology, Sorbonne University (Chapter 3)

Michael Mäs, Department of Sociology, Karlsruhe Institute of Technology (Chapter 4)

Luis Miller, Institute of Public Goods and Policies, Spanish National Research Council (Chapter 17)

Melinda C. Mills, Nuffield College, University of Oxford (Chapter 14)

Werner Raub, Department of Sociology/ICS, Utrecht University (Chapters 1 and 9)

Matthew J. Salganik, Department of Sociology, Princeton University (Chapter 22)

Tom A.B. Snijders, Department of Sociology, University of Groningen, and Nuffield College, University of Oxford (Chapter 5)

Christian E.G. Steglich, Department of Sociology/ICS, University of Groningen, and Institute for Analytical Sociology, Linköping University (Chapter 5)

Arnout van de Rijt, Department of Political and Social Sciences, European University Institute and Utrecht University (Chapter 23)

Tanja van der Lippe, Department of Sociology/ICS, Utrecht University (Chapter 10)

Federico Varese, Nuffield College, University of Oxford (Chapter 12)

Balazs Vedres, Oxford Internet Institute, University of Oxford (Chapter 21)

Thomas Voss, Institute of Sociology, University of Leipzig (Chapter 25)

Duncan J. Watts, Annenberg School for Communication, University of Pennsylvania (Chapter 22)

Dingeman Wiertz, UCL Social Research Institute, University College London (Chapter 24)

Acknowledgments

The *Handbook of Sociological Science: Contributions to Rigorous Sociology* would have been impossible without the active involvement of numerous people and institutions. First of all, we would like to thank the contributors. They were willing to share their expertise and write informative chapters. They were likewise patient and thoughtful throughout an, at times, meticulous process that we as editors asked them to go through.

As part of the process, we held a series of online meetings in December 2020, during the Covid-19 pandemic, at which all contributors participated actively as reviewers and discussants of each other's chapters, with written contributions on each chapter that had been circulated ahead of the meetings. In this way, the individual chapters and the *Handbook* as a whole benefited from detailed and constructive comments and suggestions as well as from lively discussions. Contributors and we editors found these meetings a great success. Claudia Fanti (European University Institute) provided administrative support for these meetings and saw to it that everything worked smoothly. Subsequently, contributors revised their chapters. As editors, we then provided another round of comments for each chapter and contributors afterwards prepared the final versions.

This *Handbook* would not have been possible without the support of Hans-Peter Blossfeld, Editor of the Edward Elgar *Research Handbooks in Sociology Series*. Peter was enthusiastic about the idea for our *Handbook* from the very beginning and supported the project in various ways. We likewise acknowledge the collaboration with various people at Edward Elgar, our publisher. Preparing the complete manuscript for submission, Esmee Bosma (Utrecht University) worked carefully and efficiently on unifying formatting and layout of the chapters. Jesper Sørensen, Editor-in-Chief of the journal *Sociological Science* and his editorial team gently agreed with sharing the title of their journal with parts of the *Handbook*'s title, even though the *Handbook* is not affiliated with the journal.

We are grateful to authors of the 'showcase chapters' in Part IV for their willingness to include reprints of their work and for active participation in the meetings on reviewing and improving draft-versions of chapters. We likewise acknowledge the respective journals and publishers of these journals for their permission to reprint. More specifically, our *Handbook* comprises the following reprints:

Richard Breen and John H. Goldthorpe (1997), 'Explaining educational differentials: Towards a formal rational action theory', *Rationality and Society,* **9**, 275–305 (reprinted as Chapter 18). Reprinted with permission from Sage Publications.

Ronald S. Burt (2004), 'Structural holes and good ideas', *American Journal of Sociology*, **110**, 349–399 (reprinted as Chapter 20). Reprinted with permission from the University of Chicago Press; © 2004 by The University of Chicago. All rights reserved. 0002-9602/2004/11002-0004$10.00.

Matthew J. Salganik, Peter Sheridan Dodds, and Duncan J. Watts (2006), 'Experimental study of inequality and unpredictability in an artificial cultural market', *Science*, **311**, 854–856 (reprinted as Chapter 22). Reprinted with permission from AAAS.

Arnout van de Rijt (2019), 'Self-correcting dynamics in social influence processes', *American Journal of Sociology*, **124**, 1468–1495 (reprinted as Chapter 23). Reprinted with permission from the University of Chicago Press; © 2019 by The University of Chicago. All rights reserved. 0002-9602/2019/12405-0004$10.00.

We are very happy that all the 22 chapters of this *Handbook* that are original contributions are also available Open Access. We are confident that Open Access for 22 of the 26 *Handbook*-chapters will be attractive for readers as well for the authors of the chapters and will contribute to the circulation of the *Handbook*. Open Access will, for example, also facilitate use for teaching purposes. The four showcase chapters that are reprints, for copyright reasons, are not Open Access via the *Handbook* but are easily available in digital format via libraries.

Last but not least, some words are in order for our own home institutions: the European University Institute with its Department of Political and Social Sciences, Nuffield College at the University of Oxford, and Utrecht University with its Department of Sociology as well as the Interuniversity Center of Social Science Theory and Methodology (ICS). Our home institutions have contributed first of all by being academic and intellectual strongholds of rigorous sociology. Our home institutions also provided funding when needed, including considerable funding (also from the University Library and the Faculty of Social and Behavioral Sciences of Utrecht University) for making Open Access possible for all but the reprint-chapters of the *Handbook*.

Working on this *Handbook* has sometimes been demanding, but our collaboration as editors was always pleasant and predominantly entertaining. As editors, we learned a lot from each other as well as from the *Handbook*-contributors throughout the project. We pseudo-randomized the order of editorship, avoiding indication of any type of priority

Klarita Gërxhani, Nan Dirk de Graaf, Werner Raub

11. Validation strategies in historical sociology (and beyond)*

Ivan Ermakoff†

INTRODUCTION

Validation is a linchpin of scientific rigor. Claims relying on arguments by assertion, embedding themselves in self-validating discursive setups, or dodging critical assessments undercut the prospect of sound and cumulative knowledge. A significant stake is therefore attached to clear-cut validation yardsticks.

For inquiries that embed their object in history—namely, they identify these objects by reference to chronological coordinates—and that aim for generalizable claims, the issue deserves close attention. Inquiries of that kind are exposed to divergent epistemic injunctions. While the historical dimension of their object compels them to attend to the specifics of their case, their social-scientific ambition, on the other hand, requires that they probe claims abstracted from empirical specificities. Furthermore, a quick perusal of how they proceed to validate their claims reveals a quite disparate, if not fragmented, landscape. Unless we sort out the ins and outs of these ways of proceeding, it is unclear how we might achieve a sense of shared standards as this *Handbook* advocates (see the chapter by Raub, De Graaf & Gërxhani on rigorous sociology).

Accordingly, this chapter seeks to systematize our knowledge of validation practices in historical sociology. It does so by identifying the operating principles of seven distinct validation designs or strategies, and by outlining the challenges each faces. (1) The *descriptive fit* design denotes the use of empirical evidence to document a matter-of-fact claim preceding causal interpretation. (2) A strategy in terms of *observable implications* probes empirical predictions derived from an explanatory argument. (3) *Counterfactuals* validate causal claims by assessing the effects of a counter-to-the-fact condition. (4) The *natural experiment* design takes advantage of exogenous sources of variation bequeathed by history in order to ground causal hypotheses. (5) *Inductive comparisons* certify causal inferences based on the comparison of attributes across cases in light of formal criteria of consistency and parsimony. (6) *Process tracing* validates claims about mechanisms by documenting the process whereby change is effectuated. (7) *Simulation* adds to the plausibility of such claims through the computational modeling of interactions.

* K. Gërxhani, N.D. de Graaf and W. Raub (eds) (2022), *Handbook of Sociological Science: Contributions to Rigorous Sociology*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing. ISBN: 978 1 78990 942 5; http://dx.doi.org/10.4337/9781789909432.

† I would like to thank Werner Raub and Federico Varese for their constructive comments.

DESCRIPTIVE FITS

In the first third of the twentieth century, 'Blacks, Mexicans and European immigrants inhabited three separate worlds of reliefs', three worlds characterized by particular sets of race, labor and political relations (Fox 2012, p. 11). Controlling for differences across cities in urbanization, need, and fiscal capacity, it appears for instance that 'cities with more blacks and Mexicans . . . spent less on relief and were more privately oriented, while cities with more foreign-born whites spent more on relief and were more publicly oriented' (pp. 62–64). Fox establishes this descriptive diagnosis through a detailed and systematic analysis of data collected by the US Census Bureau at the municipality level. The soundness of the diagnosis sets the ground for subsequent explanatory explorations.

Assessing the fit between evidence and descriptions can be more or less formalized. At one end of the spectrum stand studies recording factual observations without relying on formal indicators (e.g., Adams 1996; Camic 1983, Chapter 1; Go 2011; Loveman 2014; Roy 1997; Zaret 1996). At the opposite end stand descriptive syntheses that draw on techniques of data reduction and formalization for the purpose of identifying phenomenal patterns. Suffice it to mention (among many others): probability tests (McLean & Padgett 1997, pp. 226–231), multivariate decompositions of temporal trends (Brooks & Manza 2013, p. 733), vital event and life course indicators (Lai et al., 2019, pp. 31–47), measures of spatial clustering (Grigoryeva & Ruef 2015, pp. 821–833; Whitt 2010, p. 157), measures of fit with known distributions (e.g. Biggs 2005, pp. 1697, 1702), block models of tie patterns (Slez 2020, pp. 98–99), latent class analysis (Bonikowski & DiMaggio 2016, pp. 957–962), or longitudinal analyses (O'Hearn 1994, pp. 598–610).

Reliability

For this validation strategy, the first challenge is whether the sources on which it relies are reliable. When sources are 'primary'—that is, witnesses, record keepers or participants produce them—gauging reliability rests on the ability to identify, and check, selection and measurement biases (Gould 1999, pp. 363–366). These can be due to actors selecting themselves into producing a certain type of sources (Hug 2003, p. 258), actors' cognitive inclinations or positional interests (e.g., in the case of news reporters: Earl et al. 2004), or actors' strategic interest in disguising some pieces of information (Hung 2011, p. 51; Su 2011, pp. 47–53). We are in a better position to address these biases once we identify the conditions of production, or enactment, of primary sources (Barkey & Van Rossem 1997, p. 1359–1360; Ermakoff 2008, Appendix A).

Studies that lean on secondary references for their part (e.g. Abu-Lughod 1989; Hui 2005; Wallerstein 2011) expose themselves to the risk of giving precedence to references that are readily accessible (convenience bias), or that fit their claims (confirmation bias) (Goldthorpe 2007, pp. 33–36). These issues are particularly salient when the historiography provides contrasted accounts. Acknowledging these contrasts and making selection rules transparent (Møller & Skaaning 2021, pp. 110–115) are steps in the right direction. However, they can hardly be enough. The critical engagement with secondary sources is unlikely to be sound if analysts do not develop their own historical expertise (Ermakoff 2019, p. 8).

The issue of reliability extends to descriptive inferences based on data formalization. Techniques of data reduction require operationalization choices regarding how to delimit the empirical set to which the method will be applied (Abbott & Hrycak 1990, p. 159), which descriptive categories and items of information to select (Griffin 1993, p. 1105; Franzosi 2010, p. 52; Vicari 2010, pp. 510–513), how to operationalize them, and which values should be ascribed to the method parameters at the data reduction stage. Analysts' interventions at these different steps may crucially affect the findings that the technique yields. At issue is the possibility that these interventions might produce 'artifactual results' (Abbott & Tsay 2000, p. 16).

Empirical Demarcations

The ability to demarcate negative from positive cases conditions the empirical validation of an argument. If the demarcation is fluid or non-existent, cases can be tailored to fit in. Hence, claims relying on a descriptive fit design cannot be duly validated unless the concepts they invoke are clearly demarcated from an empirical standpoint. We should be able to draw the line between cases that belong in the concept's court, so to speak, and cases that do not. *The Dark Side of Democracy* by Mann (2005) for instance conceptualizes 'modern' democracy as welded to an ideology of sovereignty celebrating an ethnic understanding of the 'people' ('ethnos') (pp. 2–3). In parallel to this ideology of sovereignty, the book's introduction alludes to an understanding of democracy as a type of regime (p. 4). This second meaning is implicitly institutional. Given this plurality of possible semantic imputations, 'democracy' has no precise empirical boundaries. Revealingly, none of the case studies included in *The Dark Side of Democracy* (settlers' democracies, Nazi Germany, the Soviet Union, Cambodia under the Red Khmers, and the former Yugoslavia under Communist rule) fits a minimal and institutional understanding of democratic regimes.

OBSERVABLE IMPLICATIONS

Validation through observable implications is a two-step affair. The first step deduces empirical inferences from an explanatory argument. The second step confronts these inferences with the evidence. In this design, inferences have the status of empirical predictions (Stinchcombe 1968, p. 17). Within this methodological compass, two types of empirical inferences can be distinguished.

One type refers to phenomenal observations. Depending on the unit of analysis, these may have the format of aggregate indicators (Emigh 1997, pp. 429–430), morphological patterns (Braun 2019, pp. 74–75, or descriptive accounts of behaviors, beliefs and representations (Kalyvas 2006, chapter 9). *The Disciplinary Revolution* (2003) by Gorski about the impact of Protestantism on state power in the early modern period offers an example of validation relying on expected societal implications. Gorski argues that Protestantism in general and, Calvinism in particular, contributed to the disciplining of state agents and citizens, thereby significantly enhancing the administrative and regulative capacity of state organizations. The predictive inference derived from this argument is that in Protestant countries the levels of administrative efficiency and social order should be

higher in the rest of Europe (p. 36). *The Disciplinary Revolution* examines these empirical expectations in light of a comparative analysis of taxation and crime rates. The Dutch state's ability to extract fiscal resources from its subjects was indeed much higher than that of its counterparts (pp. 49–50). Equally significant, murder rates were lower in the Netherlands and in England in the seventeenth century than elsewhere in Europe despite a less developed repressive apparatus (pp. 52–53). Contemporary testimonies corroborate these societal indicators (ibid).

The second type of predictive inferences is correlational (e.g., Braun 2019, Chapter 5; Mizruchi et al. 2006, pp. 322–327; Weeks 2014, Chapter 2). Consider Haveman's (2015) analysis of the material and cultural foundations of magazine expansion in the US between 1741 and 1860 (Chapter 3). Haveman argues that by providing a distribution infrastructure the development of the postal system in the first half of the nineteenth century 'reduced spatial barriers to interaction between magazines and their readers' (p. 67) and, as a result, made the presence of a readership and potential contributors in a given state less of a relevant factor for the decision to found a magazine. This argument, if valid, predicts a decline in the correlation between the founding rates of magazines by state and the number of in-state magazines (p. 67). Controlling for state population, infrastructure development and literacy, a fixed-effect model capturing both unmeasured time-varying and state-specific effects confirm this predictive claim (pp. 68–69, pp. 281–287).

Validation through observable implications extends to multiple claims vying with one another for explanatory pre-eminence (e.g., Soule & Olzak 2004, pp. 479–483; Wimmer & Feinstein 2010, pp. 767–771). Kaufman's (1999) investigation of the effects of civic associational activity on local government and democratic participation is a case in point. Kaufman draws the correlational implications of (1) the Tocquevillian approach to associations as an alternative to government; (2) the pluralist conception portraying associations as self-seeking interest groups; and (3) the social capital perspective conceptualizing them as a stimulus to democratic involvement (p. 1137). For instance, if associations spur democratic life, as the social capital perspective argues, an increase in associational activities should be positively correlated with voter turnout. Controlling for the effects of urbanization, industrialization, regional differences and immigration (pp. 1297, 1317–1325), data pertaining to 53 of the largest cities of the United States in 1880 suggest that associations operate more as self-seeking factions than as an alternative to government intervention or as a stimulus to democratic participation.

Deriving Inferences

Regarding the elaboration of predictive inferences, different issues have to be considered. One is of course the possibility of faulty inferences. While formal modeling, thanks to the precision it imposes (e.g., Gould 2002, p. 1155; Diekmann's chapter on rational choice sociology), may help mitigate this possibility, it provides no warrant against it. Faulty inferences result from a non sequitur (Chwe 1999, p. 138), a logical lapse (Oliver 1980, pp. 1373–1374), or the reliance on an unjustified auxiliary assumption (Simon 1986, p. S213). Assuming that faulty inferences are not an issue, a second challenge relates to the adequate specification of the 'conditions under which the implications of a theory might be expected to operate' (Lieberson & Horwich 2008, p. 11).

Multiple derivations from the same theoretical argument raise a third type of challenge. The empirical breadth of a theory—the range of empirical phenomena it purports to explain—makes it a priori amenable to several predictions. The 'theory of market transition,' which argues that 'in reforming socialist economies, the transition from redistributive to market coordination shifts sources of power and privilege to favor direct producers' (Nee 1989, p. 663) is a theory of that kind. When, applied to specific cases, empirical predictions outline a consistent overall diagnosis, this embarrassment of riches is naturally a blessing (Nee 1989, pp. 670–678; Gerber & Hout 1998, pp. 33–38). When, on the other hand, predictive inferences yield non-congruent conclusions, analysts may want to assess whether these inferences have different degrees of empirical likelihood. Derivations empirically more difficult to obtain have greater confirmatory leverage on an argument than less demanding inferences. This explains why in terms of research priority they have precedence (Rogowski 1995, p. 467). Conversely, their empirical refutation is less disconfirming (for a similar rationale regarding case selection: Eckstein 1975, pp. 178–179).

Empirical Probes

If we turn our attention to the second step of this validation strategy—the empirical assessment of predictive inferences—the main concerns pertain to the reliability of the probe. Inferences cast in phenomenal terms have to deal with the reliability issues mentioned in the section on descriptive fits: the reliability of the sources that undergird the probe, and the robustness of the patterns identified through a technique of data reduction.

Tests informed by correlational predictions for their part have to deal with the inferential challenges due to the usual culprits: selection biases, measurement errors, spurious associations, omitted explanatory variables and unobserved heterogeneity. The gamut of strategies available to control for selection biases includes correcting for case selection, modeling data production (Przeworski et al. 2000, Appendix 1), randomizing cases and control groups (e.g., Gould 1993, pp. 726–727; Hechter et al. 2016, p. 175), controlling for group parameters susceptible to bias representativeness (Traugott 2002, pp. 191–195) and opening up the empirical assessment to all possible outcomes (Hagen et al., 2013). Fixed-effect models address the issue of unobserved heterogeneity. In addition, robustness checks and sensitivity tests help assess the extent to which a correlational diagnosis is sensitive to model specifications and alternative measures (e.g., Chen 2007, pp. 1751–1758; Kadivar 2018, pp. 404–405; Sharkey & Elwert 2011, pp. 1954–1958; see the chapter by Auspurg & Brüderl on reproducibility and credibility).

COUNTERFACTUALS

Validation rests on a counterfactual rationale when it investigates a counter-to-the-fact condition to gauge a difference in outcomes (see Breen's chapter on causal inference). The rationale has a distinctive discursive structure: had condition ζ been present, what would have been the outcome? 'Condition ζ' can be cast as a counter-to-the-fact presence ('had the presidential candidate been present at this meeting, what would have been decided?') or absence ('if this category of voters had not been college graduates, what would have been their propensity to vote for this presidential candidate?'). In

any case, the raison d'être of this setup is to certify the responses it yields as causal statements.

Counterfactual assessments are amenable to two modes of analysis. One follows the precepts of a 'genetic' approach to causal investigation in the sense that it pays attention to the genesis of processes and outcomes (Ermakoff 2019, p. 11). Its focus is on the temporal unfolding of a process understood as a sequence of actions and reactions (Deluermoz & Singaravélou 2021, pp. 213–238). The specification of this process is precise enough so that one can plausibly gauge the consequences of a counter-to-fact condition. A processual understanding along these lines is therefore micro-analytical and forward-looking (e.g., Ermakoff 2008, pp. 246, 329; Griffin 1993, p. 1102). The second mode of counterfactual validation is variable-centered. The main question of interest is: how much variation in the dependent variable y would have been obtained had the value of the covariate x been m (unobserved) (Morgan & Winship 2015, p. 13)? Imputation of causality amounts to an estimation of the difference in outcome resulting from exposure to a given condition (treatment) compared with the counterfactual world in which such exposure would not have taken place (e.g., Alderson 1999. pp. 716–717; Lin & Tomaskovic-Devey 2013, pp. 1311–1312).

Theoretical Explication

There is no counterfactual statement that is not theoretically laden. That is, the meaning of a counterfactual assessment ultimately lies in the theoretical proposition justifying the putative connection between cause and effect (Elster 1978, p. 182; Weber 1949, p. 173). This point may seem obvious. It has nonetheless two implications that are of relevance for the credibility of this validation design. (1) If the connection between effect and cause is analytically well grounded, then it should be possible to explicate it in theoretical terms without reference to the specifics of the case. The ability to explicate the theoretical claim undergirding a counterfactual demarcates rigorous analyses from rhetorical wrappings without substance. (2) Once the theoretical underpinnings of a counterfactual claim have been formulated, the question that comes to the fore is whether its formulation is specific enough so that it is amenable to critical scrutiny. Here, the reference to conditional factors is key. Several points are worth underscoring. (a) Specifying conditional factors increases the ability to gauge whether the claim applies to the empirical class under consideration. (b) In limiting the scope of a theoretical claim, the specification of conditional factors paradoxically makes this claim more refutable than if its scope was left indeterminate. (c) The more precise conditional factors are, the more specific a counterfactual elaboration can be (Ermakoff 2019, p. 14).

Well-tempered Realism

Counterfactuals can hardly deliver what they promise if they are not empirically plausible. This requirement raises the following question: to what extent can a counter-to-the-fact condition deviate from factual observations and nonetheless be considered empirically plausible? A processual approach and a variable-centered approach tackle this question in different terms. Yet, and quite tellingly, they address it in the same spirit, that of a well-tempered realism.

In the variable-grounded conception, plausibility can be gauged in light of an assessment of how distant a counterfactual is from the observed data. This assessment helps 'ascertain whether the data can support a counterfactual inference without excessive model dependence' (King & Zeng 2007, p. 190). In the processual conception, the key rule of practice is the minimal rewrite criterion (Hawthorne 1991, pp. 158–159; Lebow 2000, p. 597): a counter-to-the-fact statement is plausible if it only minimally rewrites the historical record. In concrete terms, this means that only one factor of the historical scene—one action, one resource, one constraint or one happening—becomes counterfactual. The configuration of relations in which this counterfactual is envisioned on the other hand remains as it is.

To the minimal rewrite criterion can be added a clause relative to actors' own understanding of the realm of the possible. Here the clause is that when counterfactuals presuppose beliefs, representations or expectations beyond the boundaries of what is conceivable for historical actors, they lack realism. By way of consequence, such counterfactual speculations miss the point. Sohrabi's (2011) detailed analysis of the collective dynamics that at the turn of the twentieth century shaped the constitutionalist revolutions in Russia (1905), the Ottoman Empire (1906) and Iran (1908), fleshes out this point. In all three instances, revolutionaries modeled their action on the template provided by the French revolution of 1789: the executive had to be subsumed to the legislative. Ottoman revolutionaries in 1908 'refused to capture the executive not because they did not want to, but because they did not know how to: they lacked a theory of . . . the takeover of the executive' (Sohrabi 2011, p. 24). At this stage in the process of revolutionary contention, a counterfactual scenario hypothesizing revolutionaries' taking over the executive would lack empirical relevance.

NATURAL EXPERIMENTS

A natural experiment design takes advantage of a historical event that can be considered exogenous to the cause–effect nexus under study. Its legacy is presumed to be equivalent to a randomization (or quasi-randomization) procedure allocating actors between a group that experiences a given condition and a group that does not ('control group') (Dunning 2012, Chapter 2). The purpose of the design is to test and, hopefully, validate a hypothetical claim about the effect of this condition ('treatment'). The imputation of randomness rules out the prospect of obtaining observed differences reflecting 'pre-treatment' characteristics. The ambition of this validation design is therefore to free causal inferences from the Damoclean sword of unobserved heterogeneity and confounders.

Historical events that operate as external shocks, such as natural catastrophes, a priori fit the bill of this validation design. For instance, the impact of the Boll Weevil infestation on tenant farming in Southern US counties (1892–1930) helps test the claim that tenant farming was associated with early marriages among African Americans (Bloome et al. 2017, pp. 1040–1046).

'As-if Random'

The credibility of the design hinges on the claim that the contextual conditions in which actors find themselves approximate randomized or quasi-randomized trial conditions.

When this claim appears to be frail or dubious, the design is in trouble. Two procedures address the issue. One consists in running 'balance tests' to gauge the absence of real differences between the treatment and the control groups with regard to pre-treatment covariates (Dunning 2012, pp. 239–241). However, as Kocher & Monteiro (2016, p. 953) have pointed out, these tests cannot be viewed as definitive since they cannot rule out the possibility of unobserved heterogeneity and confounders. The second procedure culls direct evidence about the process whereby units were assigned to 'treatment' and 'control' (Hyde 2007, p. 46). At stake here is the ability of deep historical knowledge to debunk mistaken characterization of the assignment as quasi-random.

The tendency to interpret frontiers and administrative boundaries as 'arbitrary' and therefore prone to a natural experiment setting has not been exempt from this mischaracterization. Let us mention the study of differences in the rates of sabotage in districts (*départements*) adjacent to the line of demarcation in France during the Second World War (Ferwerda & Miller 2014). Underlying the validation strategy is the claim that the line of demarcation at the local level was 'as good as random' and that the territories thus demarcated exemplified two conditions of state power: under control of the occupier (northern zone) and devolved (southern zone) (p. 647). Neither claim is historically corroborated (Kocher & Monteiro 2016, pp. 964–965). The differential in rates of sabotage across the line reflects, not the impact of power devolution on resistance, but the more prosaic fact that double-track railways were located north of the demarcation line—a fact created by the German operatives who traced the line in June 1940 (ibid, pp. 966–969).

Observing 'Treatment'

For the promoters of the natural experiment design, the gold standard is the experimental control trial (e.g., Dunning 2012, p. 15). Unsurprisingly, the terminology that prevails in the literature regarding natural experiment adopts the language of 'treatment' and 'control.' For the sake of identifying a causal effect it becomes key to gauging whether a 'treatment' is not a bundle of different treatments combining their effects in various ways. At stake is the 'internal validity' of the inferential statements (see the chapter by Gërxhani & Miller on experimental sociology). Laboratory experiments are geared to avoid this possibility. Socio-historical processes, however, can interact in ways that make it difficult to pin down what the 'treatment' embodied in one specific process amounts to. The difficulty is compounded by the lack of clear-cut and bounded outcome specifications (Geddes 2003, pp. 43–45).

While a natural experiment design a priori approaches causal analysis from the standpoint of the 'effects of causes,' it is interesting to observe that debates about the validation of the causal inferences yielded by this design has shifted the attention to the process whereby an outcome is brought about. In methodological jargon, the focus is on 'mechanism causal process observations' (CPOs) (Collier et al. 2010, p. 184; Dunning 2012, p. 209). Yet, inasmuch as the methodological apparatus deployed for this purpose remains cast in terms of variables, it is unclear what 'process' is supposed to mean in this context. If the notion is equated with the identification of 'intervening variables,' the approach runs the risk of stumbling over unspecified or underspecified causal connections.

INDUCTIVE COMPARISONS

Inductive comparative analysis, first, identifies combinations of explanatory factors and outcomes—coded as either present or absent—across cases and, second, submits the combinations thus identified to Boole algebra's formal requirements of simplification and inference (Ragin 2014). When the independent variables defining explanatory factors are continuous, cases are identified by their degree of membership to the empirical sets defined by these variables (Ragin 2008). In this design, the formalism of Boole algebra and set theory certify the validity of the causal inferences derived from the systematic recording of differences and similarities across cases. Factorial associations denote 'causal configurations' that, taken together, form a 'causal recipe' (Ragin 2014, pp. xxiii, xxviii). Underlying this approach to causal inferences is a 'conjunctural' understanding of causation, which states that causation operates as conjunctions of factors and that these conjunctions vary across cases. This design has been used to tackle a wide range of objects of investigation from the occurrence and outcomes of guerrilla movements in Latin America (1956–1980; Wickham-Crowley 1992, Chapter 12) to 'Left-Libertarian party' successes (1980–1997) (Redding & Viterna 1999, pp. 497–502).

Conjunctures and Associations

Two sets of consideration call into question the grand inductive ambitions of this validation design: (1) the fragility of the diagnoses it yields, and (2) the discrepancy between its professed commitment to a conjunctural conception of causation and the operational understandings of causes built in the design.

(1) *Fragility of the diagnoses*. As for any variable-grounded approach to causal inference, Boolean comparative analysis has to contend with the possibility of biases resulting from the selection of cases, sources and variables (Ermakoff 2019, pp. 7–8). Case selection can significantly alter the identification and configuration of factorial associations when selection is endogenous, based on the dependent variable, or conducted without a clear-cut delimitation of the universe from which cases are drawn. Extensive reliance on secondary sources for the coding of cases takes the risk of convenience and confirmatory biases as defined in the first section of this chapter. As for the issues of reversed causation and variable omission related to variable selection, they remain unaddressed in this validation design.

This last point deserves attention. A Boolean approach identifies associations that ultimately are a function of the factors, or attributes, selected as explanatory variables. There is no reason to assume that the set adopted in any study is definitive. Selecting variables on the basis of their theoretical hypotheses, or the state of the literature, does not insure a Boolean comparative investigation against the drawback of missing variables. Furthermore, independent variables constructed at a macro level often have a compound character: they subsume various factors to the same category (Hall 1999, p. 162). For instance, the variable 'history of opposition' (Beck 2014, pp. 210–211) encompasses different types and modes of opposition (e.g., religious, ethnic, party-based, armed, violent). For the sake of explanation, an analyst may find it judicious to disaggregate a

compound category into two or more variables. When these are added to the set already taken into account, the associations constitutive of the 'causal recipe' are likely to be different. In particular, causes initially labeled 'necessary' or 'sufficient' may turn out to be neither, thus raising doubts about the significance of these causal imputations.

Concerns about the method's ability to validate causal inferences have further been raised in light of several observations. First, even when the data are random, the method is likely to identify patterns of association when there should be none (Lieberson 2004, p. 14; Lucas & Szatrowski 2014, pp. 59–60). Second, the causal configurations identified by the method are sensitive to coding decisions and the methods used to calibrate the degree of membership of cases to fuzzy sets (Goldthorpe 2007, pp. 46, 226–227; Hug 2013, pp. 260–261; Lucas & Szatrowski 2014, p. 47). Third, simulations assigning values to cases on the basis on a given Boolean factorial expression show that the method fails to identify the correct patterns of association (Lucas & Szatrowski 2014, pp. 19–20). Part of the problem lies in the algorithms and routines that assign values to contradictory combinations and deal with limited diversity, that is, combinations of attributes that although possible theoretically are not observed among the cases under investigation (Seawright 2014). These observations raise serious concerns about the reliability of the technique and, by way of consequence, its demonstrative value.

(2) *Static and conjunctural causation*. In addition, the static understanding of causes built in the design appears at odds with the conjunctural conception of causation it professes. Boolean comparative analyses 'freeze' the time of history—to use Burawoy's (1989, p. 782) formulation regarding macro comparative inductions. Attributes and factors are abstracted from their temporality, i.e., their unfolding in time. By way of consequence, the approach ignores the possibility that the same nominal factor pertaining to the same case—say 'economic pressure' (Beck 2014, p. 210)—may produce different effects depending on which moment in the temporal unfolding of the case is taken into account. In short, the method prevents itself from fully taking into account conjunctural effects understood in temporal terms. This inability is at odds with the causal philosophy—'causation is conjunctural'—that presided over the elaboration of the method.

The Language of Causes

If diagnoses regarding a given outcome are likely to vary from one study to another depending on the number of variables selected, if random attributes are likely to be picked up as associations, if due to its algorithms the method can misrepresent true associations of attributes, and if the static approach to causation prevents a truly conjunctural understanding of the temporality of causes, then a word of caution regarding causal imputations is in order. Yet, Boolean comparative analyses are rife with causal assertions often framed by reference to necessary or sufficient causes (e.g., Mahoney 2010, pp. 115–119, 183–188). At times, these are couched in a strong deterministic language, which interpret variables in a literal fashion and elevate them to the status of monadic causal agents. True enough, these agents operate for the most as elements of various combinations. But they are the building blocks of these combinations and, in some instances, the method identifies them as standing alone causal agents. This formal representation vests them with monad-like capacity.

A conjunctural understanding of causation leads us to expect that factors—assuming that these factors are clearly specified with regard to their empirical content—have different 'effects' depending on how they combine with other factors. Along these lines it is plausible to argue that effects are different because the processes undergirding their production are different as well. If so, the focus should be on processes of emergence or production—that is, 'genetic processes.' Ultimately, we can account for the conjunctural variation of effects if we have a clear understanding of the genetic processes at play. Boolean comparative analysis, however, stops short of processual specifications. By itself, it cannot highlight the genesis of effects.

These different points suggest that the design is best used to test hypotheses, not as an inductive method (Hug 2013, pp. 253–254). Putting at bay an inductive use may nonetheless prove quite challenging. Witness Brown & Boswell's (1995) Boolean comparative analysis of interracial solidarity in 16 northern US cities affected by the 1919 steel strike. This study first tests hypotheses derived from Heckathorn's (1990) formal model of collective action adapted to racially split labor market conditions. The focus is on two parameters: whether the union can be considered weak or strong, and whether the minority labor group is predominantly "local" or "migrant" (pp. 1485–1493). Analyzed in this light, it appears that the interracial solidarity observed in Cleveland and Wheeling contradicts the lack of solidarity in cities with similar combinations of independent variables (p. 1502). To inductively resolve this contradiction, Brown and Boswell include three additional independent variables: city size, steel company ownership and local government repression. Reexamined along these lines, a Boolean comparative inquiry concludes that interracial solidarity took place in cities with strong union organizations and nonrepressive governments (pp. 1502–1505).

PROCESS TRACING

Compendiums on 'process tracing' offer two characterizations of this validation strategy. One refers to 'the use of evidence from within a case' (Bennett & Checkel 2015, pp. 4, 8), and the other to the 'analysis of evidence' for the purpose of 'testing hypotheses about causal mechanisms' (ibid, p. 7). The reference to 'evidence from within a case' is peculiar: nothing in the fact of being 'within' a case a priori tells us something about the processual character of this evidence. The reference to the 'testing of hypotheses about mechanisms' appears more in tune with a focus on processes, provided that we have a clear understanding of what we mean by 'mechanism.'

Over the last three decades, 'mechanism' has become as faddish as 'process tracing.' By way of consequence, definitional statements have been profuse. Insofar as the language of mechanism is interpreted as reflecting a concern for precision and rigor, a 'mechanism' may be conceptualized as the 'analytical specification of change' (Ermakoff 2019, p. 12). On this score, two conceptions can be contrasted. A variable-based conception equates 'mechanisms' with variables intervening between an initial condition and an outcome (Imai et al. 2011, p. 765; Knight & Winship 2013, p. 283). An action-based conception by contrast views a 'mechanism' as the specification of how actions and interactions bring about a change in state (e.g., Gambetta 1998, p. 102–103; Hedström 2005, p. 25; Hernes 1998, p. 95).

'Processes' obviously are inscribed in time. We commonly say that a phenomenal reality is 'in process' in order to describe the fact that it is happening, which is also to say that it is bringing about change or reproduction. Hence, 'process' in a minimal sense can be conceptualized as the effectuation of change or reproduction. It follows from this simple observation that, if the point of tracing a process is to validate hypothetical claims about one or several mechanisms, a variable-based conception of the latter is somehow at odds with this empirical focus. Transposed to the realm of socio-historical phenomena, the drawback of a variable-based conception is that it loses sight of the temporal dimension of what it seeks to capture. In the action-based conception, on the other hand, this temporal dimension is front and center. Analysts can neither elude nor fend it off.

Documenting Processes

How then shall we validate claims cast in terms of mechanisms when the focus on processes is action-centered? We need to be able to observe processes in the making in order to assess which mechanism these observations corroborate or qualify. This means delving into the ins and outs of actions as they bring about change or reproduction in a collective setting. Three consequences follow. First, the focus is as much on actors as it is on their actions (e.g., Goldberg 2007, Chapters 4–5), their cognitive or psychological states (Paschel 2016, Chapter 4), or the relational configurations they have to deal with (e.g., Anderson 2021, Chapters 2–4). These actors can be individual or collective. Analyzing a collective actor's behavioral stance, however, requires investigating the dynamic of interactions that bring about this stance (Ermakoff 1997, pp. 413–417, 2008, Chapters 7–9; Padgett & McLean 2006, pp. 1468–1471). As a result, the individual actor may be considered the 'regulative' unit of analysis . Second, in order to clearly track a process in the making and its dynamic, it is key to clearly identify and thus delineate the group that experiences it (Bonnell 1983, Chapter 1). Third, since the point of tracing a process is to observe how change is effectuated in time, this mode of validation implies a forward-looking temporal standpoint making possible the reconstruction of change without presuming the outcome.

Jansen's (2017) account of the creation of a new form of political mobilization—populist mobilization—in Peru in 1931 (*Revolutionizing Repertoires*) is a case in point. This study probes the process whereby two political outsiders contending for the position of president—Haya de la Torre and Sánchez Cerro—drew on their experience as outsiders to experiment with new political practices geared to the mobilization of marginalized people, and infused with anti-elite rhetoric. Jansen carefully reconstructs sequences of actions and events with a close focus on issues of temporality and actors' subjective assessments, motives and initiatives. *Revolutionizing Repertoires* for instance shows how Sánchez Cerro initially favored a kind of political practice that can be dubbed 'military paternalism' before endorsing populist mobilization and rhetoric (p. 137). By documenting when the shift in practice took place, *Revolutionizing Repertoires* highlights the causal significance of outbidding in the emergence of a populist mode of mobilization.

Time Scales

The focus on actions and interactions is propitious to the adoption of a time frame confined to the short-term (days, weeks, months). Is this to say that process tracing, as a

validation design, cannot be deployed for the medium or the long run (years, decades, centuries)? Putative explanatory factors have causal relevance when their effects translate in terms of actors' dispositions, motivations, beliefs, or behaviors. Therefore, a factor theorized as a cause can be diagnosed as significant if we can somehow observe it at play in actors' behavioral make-up (Ermakoff 2008, p. xxiii). In light of this basic epistemic premise, three empirical protocols are conceivable to test claims about long-term processes.

One protocol selects one site of empirical investigation, explores the extent to which this site can be deemed representative with regard to the actors under consideration and, if the response to this query is positive, engages in a systematic account of relational patterns, beliefs, representations or behaviors. This focus is intended to provide a lens magnifying a process presumed to be at play in sites with similar characteristics (e.g., Bearman 1993, pp. 2, 175–181; Cerutti 1990, p. 23). A second protocol adopts a longitudinal perspective and examines whether within or across generations the focus group undergoes the transformative process specified by the mechanism being hypothesized (e.g., Elias 1994, regarding the civilizing process; Collins 2002, documenting the interaction rituals of intellectual and philosophical creativity). The third protocol, instead of examining subjective and behavioral patterns deployed *au long cours*, pays attention to conjunctures characterized by a high intensity of interactions in the course of which actors are likely to reveal how they think, feel or make choices. In numerous instances, these conjunctures are moments of confrontations (e.g., Steinmetz 1993, pp. 88–107).

Decisional Focus

Whether we are dealing with processes unfolding on a wide or a short time scale, investigating decisional moments can prove heuristically fruitful in three respects. First, this focus highlights which factors bear on actors' minds as they make their choice. In so doing, it offers us a point of entry to investigate whether and how supra-individual conditions translate, or not, into decisional parameters (e.g., Paschel 2016, p. 98; Simmons 2020, pp. 205–214; Vaughan 1996, Chapters 8–9). Second, the focus on decisional moments helps adjudicate alternative explanatory accounts theorized in terms of mechanisms (Erikson 2014, Chapter 4; Ermakoff 2008, Chapters 3–5). Third, it sheds light on the range of scenarios actors take into account and, when their relation to the immediate future betrays a sense of mutual uncertainty, on moments of collective contingency (Ermakoff 2015, pp. 92–97). This approach grounds empirically the notion of historical possibility.

SIMULATIONS

It is customary to distinguish between equation-based and agent-based simulation modeling (Macy & Willer 2002, p. 147). 'Whereas equation-based modeling attempts to express causal relations among variables, individual-based modeling represents interactions among the agents directly' (Cederman 2005, pp. 877–878). Since, from a realist standpoint, claims about socio-historical phenomena and processes are claims about actions and reactions, there is good ground to argue that for validation purposes, agent-based computational modeling fits the prospect of a realist approach better than the

variable-based format (see the chapters by Flache, Mäs & Keijzer on computational approaches and by Steglich & Snijders on stochastic network modeling).

Several steps inherent to the deployment of multiagent simulation undergird its potential as a validation technique. For one thing, it compels analysts to straighten out their claims about conditional factors and mechanisms. Not only does this approach request that hypotheses about actors' decision rules and about the dimensions of their relational environment that have an impact, be specified (Hedström et al. 2000, pp. 150–152), it also makes it possible to probe the empirical plausibility of different theoretical scenarios (Bruch 2014, pp. 1252–1255; Becker et al. 2020, pp. 878–883). Empirically grounded simulations of behavioral diffusion for instance have explored which characteristics of the relational configurations are likely to affect the dynamic of the process, and how individuals deal with decision problems (Manzo et al. 2018, pp. 1114–1116, 1134).

Furthermore, the ability to reconsider parameter values and analytical specifications, and to examine the extent to which such calibrations modify the degree of fit between predicted and observed outcomes, motivates a deep and dialectical engagement with the empirics of historically situated cases. For instance, Ziegler's (2008, p. 116–117) game-theoretical simulation of the circular exchange of gifts (*kula*), and of the network structures emerging from these exchanges is at odds with the patterns observed among tribal groups in the Massim region east of Papua New Guinea in the early twentieth century. However, close attention to the temporality of the phenomenon motivates the distinction between three historical phases in the development of the network that, once taken into account, make the simulated outcome approximate observed behaviors (ibid, pp. 118–121).

As these few remarks make clear, implementing this validation strategy requires devising a protocol for assessing the degree of fit between simulated and observed outcomes (e.g., Ziegler 2008, pp. 116–119; Manzo et al. 2018, pp. 1137–1141; Becker et al. 2020, pp. 883–888). Particularly significant are assessment protocols that seek to take into account all the relevant dimensions of variation (Manzo et al. 2018, pp. 1136–1142). A further challenge relates to the interpretation of the results when several processes get simulated: one cannot a priori rule out the possibility that as simulated processes interact they produce system properties that the theory has not anticipated and that fall under the radar of ex post diagnostics.

Coda

By way of taking stock, two points are worth underscoring. The first one concerns validation requirements for concepts and cases. The primary focus of this chapter has been on the validation of claims—explanatory or descriptive—pertaining to historical objects of investigation. Claims crown the edifice. They provide its ultimate raison d'être. But the edifice would have no consistency without concepts and cases. At issue is whether concepts and cases duly fulfill the tasks they are expected to perform and, as a result, can be validated as such. For concepts, this means unambiguous empirical demarcation (see the section on descriptive fits). We have a sound concept in hand when we can clearly delimit its empirical content in light of observable indicators. Transposed to the realm of socio-historical phenomena, the requirement of empirical demarcation takes the form of traceable temporal boundaries. Cases for their part can be viewed as empirically validated when no bias mars the sources on which they rest, and the criteria whereby they got selected.

Second, this chapter incidentally suggests that validation practices may be indebted to specific conceptions of causality. Undergirding the inductive comparisons design for instance is a linear conception of causation, which depicts causes as broad forces to which actors get exposed. Conversely, process tracing understood and practiced from a micro-analytical perspective gives credence to an interactionist understanding, which views causes as the generative processes shaping interactions or taking shape through them. Whereas, along the same lines, a processual approach to a counterfactual endorses a genetic conception of causality, a variable-based approach speaks to a conception in terms of reliable associations. So does the natural experiment design. In other words: we cannot fully engage validation practices without paying close attention to the conceptions of causality on which they draw.

REFERENCES

Abbott, A. and A. Hrycak (1990), 'Measuring resemblance in sequence data: An optimal matching analysis of musicians' careers', *American Journal of Sociology*, **96**, 144–185.

Abbott, A. and A. Tsay (2000), 'Sequence analysis and optimal matching methods in sociology: Review and prospect', *Sociological Methods and Research* **29**, 3–33.

Abu-Lughod, J.L. (1989), *Before European Hegemony: The Word System A.D. 1250–1350*, Oxford, UK: Oxford University Press.

Adams, J. (1996), 'Principals and agents, colonialists and company men: The decay of colonial control in the Dutch East Indies', *American Sociological Review*, **61**, 12–28.

Alderson, A.S. (1999), 'Explaining deindustrialization: Globalization, failure, or success?', *American Sociological Review*, **64**, 701–721.

Anderson, E. (2021), *Agents of Reform: Child Labor and the Origins of the Welfare State*, Princeton, NJ and Oxford, UK: Princeton University Press.

Barkey, K., and R. van Rossem (1997), 'Networks of contention: Villages and regional structure in the seventeenth-century Ottoman Empire', *American Journal of Sociology*, **102**, 1345–1382.

Bearman P. (1993), *Relations into Rhetorics: Local Elite Social Structure in Norfolk, England, 1540–1640*, New Brunswick, NJ: Rutgers University Press.

Beck, C. (2014), 'Reflections on the revolutionary wave in 2011', *Theory and Society*, **43**, 197–223.

Becker, S.O., Y. Hsiao, S. Pfaff, and J. Rubin (2020), 'Multiplex network ties and the spatial diffusion of radical innovations: Martin Luther's leadership in the Early Reformation', *American Sociological Review*, **85**, 857–894.

Bennett, A. and J.T. Checkel (2015), *Process Tracing: From Metaphor to Analytic Tool*, Cambridge, UK: Cambridge University Press.

Biggs, M. (2005), 'Strikes as forest fires: Chicago and Paris in the late nineteenth century', *American Journal of Sociology*, **110**, 1684–1714.

Bloome, D., J. Feigenbaum, and C. Muller (2017), 'Tenancy, marriage, and the boll weevil infestation, 1892–1930,' *Demography*, **54**, 1029–1049.

Bonikowski, B. and P. DiMaggio (2016), 'Varieties of American popular nationalism', *American Sociological Review*, **81**, 949–980.

Bonnell, V. (1983), *Roots of Rebellion; Workers' Politics and Organizations in St. Petersburg and Moscow, 1900–1914*, Berkeley, CA: University of California Press.

Braun, R. (2019), *Protectors of Pluralism: Religious Minorities and the Rescue of Jews in the Low Countries during the Holocaust*, Cambridge, UK: Cambridge University Press.

Brooks, C. and J. Manza (2013), 'A broken public? Americans' responses to the great recession', *American Sociological Review*, **78**, 727–748.

Brown, C. and T. Boswell (1995), 'Strikebreaking or solidarity in the Great Steel Strike of 1919: A split labor market, game-theoretic, and QCA analysis', *American Journal of Sociology*, **100**, 1479–1519.

Bruch, E. (2014), 'How population structure shapes neighborhood segregation', *American Journal of Sociology*, **119**, 1221–1278.

Burawoy, M. (1989), 'Two methods in search of science', *Theory and Society*, **18,** 759–785.

Camic, C. (1983), *Experience and Enlightenment: Socialization for Cultural Change in Eighteenth-Century Scotland*, Chicago, IL: The University of Chicago Press.

Cederman, L. (2005), 'Computational models of social forms: Advancing generative process theory', *American Journal of Sociology*, **110**, 864–893.
Cerutti, S. (1990), *La Ville et les Métiers: Naissance d'un Langage Corporatif (Turin, 17e-18e* siècles) [*The City and the Professions: Birth of a Corporate Language (Turin, 17th-18th centuries*)], Paris: EHESS.
Chen, A. (2007), 'The party of Lincoln and the politics of state fair employment practices legislation in the north, 1945–1964', *American Journal of Sociology*, **112**, 1713–1774.
Chwe, M. (1999), 'Structure and strategy in collective action,' *American Journal of Sociology*, **105**, 128–156.
Collier, D., H.E. Brady, and J. Seawright (2010), 'Sources of leverage in causal inference', in H. Brady and D. Collier (eds.), *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, Lanham, MD: Rowman & Littlefield.
Collins, R. (2002), *The Sociology of Philosophies: A Global Theory of Intellectual Change*, Cambridge, MA: Harvard University Press.
Deluermoz, Q. and P. Singaravélou (2021), *A Past of Possibilities: A History of Could Have Been*, New Haven, CN and London, UK: Yale University Press.
Dunning, T. (2012), *Natural Experiments in the Social Sciences: A Design-Based Approach*, Cambridge, UK: Cambridge University Press.
Earl, J., A. Martin, J.D. McCarthy, and S.A. Soule (2004), 'Newspapers and protest event analysis', *Annual Review of Sociology*, **30**, 65–80.
Eckstein, H. (1975). 'Case studies and theory in political science', in F. I. Greenstein and N. W. Polsby (eds.), *Handbook of Political Science. Political Science: Scope and Theory*, Reading, MA: Addison-Wesley, vol. 7, pp. 94–137.
Elias, N. (1994), *The Civilizing Process*, Oxford, UK: Blackwell 1939.
Elster, J. (1978), *Logic and Society*, Chichester, UK: John Wiley.
Emigh, R. (1997), 'The spread of sharecropping in Tuscany: The political economy of transaction costs', *American Sociological Review*, **62**, 423–442.
Erikson, E. (2014), *Between Monopoly and Free Trade: The English East India Company, 1600–1757*, Princeton, NJ and Oxford, UK: Princeton University Press.
Ermakoff, I. (1997), 'Prelates and princes: Aristocratic marriages, canon law prohibitions and shifts in norms and patterns of domination in the Central Middle Ages', *American Sociological Review*, **62**, 405–422.
Ermakoff, I. (2008), *Ruling Oneself Out. A Theory of Collective Abdications*, Durham, UK: Duke University Press.
Ermakoff, I. (2015), 'The structure of contingency', *American Journal of Sociology*, **121**, 64–125.
Ermakoff, I. (2019), 'Causality and history: Modes of causal investigation in historical social sciences', *The Annual Review of Sociology*, **45**, 581–606.
Ferwerda, J. and N.L. Miller (2014), 'Political devolution and resistance to foreign rule: A natural experiment', *American Political Science Review*, **108**, 642–660.
Fox, C. (2012), *Three Worlds of Relief: Race, Immigration, and the American Welfare State from the Progressive Era to the New Deal*, Princeton, NJ: Princeton University Press.
Franzosi, R. (2010), *Quantitative Narrative Analysis*, Los Angeles, CA: Sage Publications.
Gambetta, D. (1998), 'Concatenations of mechanisms', in P. Hedström and R. Swedberg, *Social Mechanisms: An Analytical Approach to Social Theory*, Cambridge, UK: Cambridge University Press, pp. 102–124.
Geddes, B. (2003), *Paradigms and Sand Castles: Theory Building and Research Design in Comparative Politics*, Ann Arbor, MI: University of Michigan Press.
Gerber, T. and M. Hout (1998), 'More shock than therapy: Market transition, employment, and income in Russia, 1991–1995', *American Journal of Sociology*, **104**, 1–50.
Go, J. (2011), *Patterns of Empire: The British and American Empire, 1688 to the Present*, Cambridge, UK: Cambridge University Press.
Goldberg, C. (2007), *Citizens and Paupers: Relief, Rights, and Race, from the Freedmen's Bureau to Workfare*, Chicago, IL: University of Chicago Press.
Goldthorpe, J. (2007), *On Sociology, Volume One: Critique and Program*, Stanford, CA: Stanford University Press.
Gorski, P.S. (2003), *The Disciplinary Revolution. Calvinism and the Rise of the State in Early Modern Europe*, Chicago, IL: The University of Chicago Press.
Gould, R.V. (1993), 'Trade cohesion, class unity, and urban insurrection: Artisanal activism in the Paris commune,' *American Journal of Sociology*, **98**, 721–754.
Gould, R.V. (1999), 'Collective violence and group solidarity: Evidence from a feuding society', *American Sociological Review*, **64**, 356–830.
Gould, R.V. (2002), 'The origins of status hierarchies: A formal theory and empirical test', *American Journal of Sociology*, **107**, 1143–1179.
Griffin, L.J. (1993), 'Narrative, event-structure analysis, and causal interpretation in historical sociology', *American Journal of Sociology*, **98**, 1094–1133.

Grigoryeva, A. and M. Ruef (2015), 'The historical demography of racial segregation', *American Sociological Review*, **80**, 814–842.
Hagen R., K. Makovi, and P. Bearman (2013), 'The influence of political dynamics on Southern lynch mob formation and lethality', *Social Forces*, **92**, 757–787.
Hall, J. R. (1999), *Cultures of Inquiry: From Epistemology to Discourse in Sociohistorical Research*. Cambridge, UK: Cambridge University Press.
Haveman, H.A. (2015), *Magazines and the Making of America: Modernization, Community, and Print Culture, 1741–1860*, Princeton, NJ: Princeton University Press.
Hawthorne, G. (1991), *Plausible Worlds. Possibility and Understanding in History and the Social Sciences*. Cambridge, UK: Cambridge University Press.
Hechter M., S. Pfaff and P. Underwood (2016), 'Grievances and the genesis of rebellion: Mutiny in the Royal Navy, 1740 to 1820', *American Sociological Review*, **81**, 165–189.
Heckathorn, D.D. (1990), 'Collective sanctions and compliance norms: A formal theory of group-mediated social control', *American Sociological Rev*iew, **55**, 366–384.
Hedström, P. (2005), *Dissecting the Social: On the Principles of Analytical Sociology*, Cambridge, UK: Cambridge University Press.
Hedström, P., R. Sandell, and C. Stern (2000), 'Mesolevel networks and the diffusion of social movements: The case of the Swedish Social Democratic Party', *American Journal of Sociology*, **106**, 145–172.
Hernes, G. (1998), 'Real virtuality', in P. Hedström and R. Swedberg (eds.), *Social Mechanisms: An Analytical Approach to Social Theory*, Cambridge, UK: Cambridge University Press. pp. 74–101
Hug S. (2003), 'Selection bias in comparative research: The case of incomplete data sets', *Political Analysis*, **11**, 255–274.
Hug S. (2013), 'Qualitative comparative analysis: How inductive use and measurement error lead to problematic inference', *Political Analysis*, **21**, 252–265.
Hui, V. (2005), *War and State Formation in Ancient China and Early Modern Europe*, Cambridge, UK: Cambridge University Press.
Hung H. (2011), *Protest with Chinese Characteristics: Demonstrations, Riots, and Petitions in the Mid-Qing Dynasty*, New York, NY: Columbia University Press.
Hyde, S. (2007), 'The observer effect in international politics: Evidence from a natural experiment', *World Politics*, **60**, 37–63.
Imai, K., L. Keele, D. Tingley, and T. Yamamoto (2011), 'Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies', *American Political Science Review*, **105**, 765–789.
Jansen, R. (2017), *Revolutionizing Repertoires: The Rise of Populist Mobilization in Peru*, Chicago, IL: The University of Chicago Press.
Kadivar, M.A. (2018), 'Mass mobilization and the durability of new democracies', *American Sociological Review*, **83**, 390–417.
Kalyvas, S.N. (2006), *The Logic of Violence in Civil War*, Cambridge, UK: Cambridge University Press.
Kaufman, J. (1999), 'Three views of associationalism in 19th-century America: An empirical examination', *American Journal of Sociology*, **104**, 1296–1345.
King, G. and L. Zeng (2007), 'When can history be our guide? The pitfalls of counterfactual inference', *International Studies Quarterly*, **51**, 183–210.
Knight, C. and C. Winship (2013), 'The causal implications of mechanistic thinking: Identification using directed acyclic graphs (DAGs)', in S. Morgan (ed.), *Handbook of Causal Analysis for Social Research*, Dordrecht: Springer, pp. 275–299.
Kocher, M. and N. Monteiro (2016), 'Lines of demarcation: Causation, design-based inference, and historical research', *Perspectives on Politics*, **14**, 952–975.
Lai, A., A. Noymer, and T. Tai (2019), 'The geometry of mortality change: Convex hulls for demographic analysis', *Revue Quetelet / Quetelet Journal*, **7**, 27–70.
Lebow, R.N. (2000), 'Contingency, catalysts, and international system change', *Political Science Quarterly*, **115**, 591–616.
Lieberson S. (2004), 'Comments on the use and utility of QCA', *Qualitative Methods*, **2**, 13–14.
Lieberson, S. and J. Horwich (2008), 'Implication analysis: A pragmatic proposal for linking theory and data in the social sciences', *Sociological Methodology*, **38**, 1–50.
Lin K. and D. Tomaskovic-Devey (2013), 'Financialization and U.S. income inequality, 1970–2008', *American Journal of Sociology*, **118**, 1284–1329.
Loveman, M. (2014), *National Colors: Racial Classification and the State in Latin America*. Oxford, UK: Oxford University Press.
Lucas, S.R., and A. Szatrowski (2014), 'Qualitative comparative analysis in critical perspective', *Sociological Methodology*, **44**, 1–79.

Macy, M., and R. Willer (2002), 'From factors to actors: Computational sociology and agent-based modeling', *Annual Review of Sociology*, **28**, 143–166.
Mahoney, J. (2010), *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*, Cambridge, UK: Cambridge University Press.
Mann, M. (2005), *The Dark Side of Democracy: Explaining Ethnic Cleansing*, Cambridge, UK: Cambridge University Press.
Manzo, G., S. Gabbriellini, V. Roux, and F.N. M'Mbogori (2018), 'Complex contagions and the diffusion of innovations: Evidence from a small-n study', *Journal of Archaeological Method and Theory*, **25**, 1109–1154.
McLean, P. D. and J. F. Padgett (1997), 'Was Florence a perfectly competitive market? Transactional evidence from the Renaissance', *Theory and Society*, **26**, 209–244.
Mizruchi, M.S., L.B. Stearns, and C. Marquis (2006), 'The conditional nature of embeddedness: A study of borrowing by large U.S. firms, 1973–1994', *American Sociological Review*, **71**, 310–333.
Morgan, S.L. and C. Winship (2015), *Counterfactuals and Causal Inference: Methods and Principles for Social Research*, New York, NY: Cambridge University Press.
Møller J. and S.E. Skaaning (2021), 'The Ulysses principle: A criterial framework for reducing bias when enlisting the work of historians', *Sociological Methods & Research*, **50**, 103–134.
Nee, V. (1989), 'A theory of market transition: From redistribution to markets in state socialism', *American Sociological Review*, **54**, 663–681.
O'Hearn, D. (1994), 'Innovation and the world-system hierarchy: British subjugation of the Irish cotton industry, 1780–1830', *American Journal of Sociology*, **100**, 587–621.
Oliver, P. (1980), 'Rewards and punishments as selective incentives for collective action: theoretical investigations', *American Journal of Sociology*, **85**, 1356–375.
Paschel, T.S. (2016), *Becoming Black Political Subjects: Movements and Ethno-racial Rights in Colombia and Brazil*, Princeton, NJ and Oxford, UK: Princeton University Press.
Padgett, J.F. and P.D. McLean (2006), 'Organizational invention and elite transformation: The birth of partnership systems in Renaissance Florence', *American Journal of Sociology*, **111**, 1463–568.
Przeworski, A., M.E. Alvarez, J.A. Cheibub, and F. Limongi (2000), *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*, New York, NY: Cambridge University Press.
Ragin, C. (2008), *Redesigning Social Inquiry: Fuzzy Sets and Beyond*, Chicago, IL: University of Chicago Press.
Ragin, C. (2014), *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, Berkeley, CA: The University of California Press, 1987.
Redding, K. and J.S. Viterna (1999), 'Political demands, political opportunities: Explaining the differential success of left-libertarian parties', *Social Forces*, **78**, 491–510.
Rogowski, R. (1995), 'The role of theory and anomaly in social-scientific inference', *American Political Science Review*, **89**, 467–470.
Roy, W.G. (1997), *Socializing Capital. The Rise of the Large Industrial Corporation in America*, Princeton, NJ: Princeton University Press.
Seawright J. (2014), 'Comment: Limited diversity and the unreliability of QCA', *Sociological Methodology*, **44**, 118–121.
Sharkey, P. and F. Elwert (2011), 'The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability', *American Journal of Sociology*, **116**, 1934–1981.
Simmons, M.C. (2020), 'Becoming wards of the state: Race, crime, and childhood in the struggle for foster care integration, 1920s to 1960s', *American Sociological Review*, **85**, 199–222.
Simon, H.A. (1986), 'Rationality in psychology and economics', *The Journal of Business*, **59**, S209–224.
Slez, A. (2020), *The Making of the Populist Movement: State Market, and Party on the Western Frontier*, Oxford, UK: Oxford University Press.
Sohrabi, N. (2011), *Revolution and Constitutionalism in the Ottoman Empire and Iran*, Cambridge, UK: Cambridge University Press.
Soule, S.A. and S. Olzak (2004), 'When do movements matter? The politics of contingency and the equal rights amendment', *American Sociological Review*, **69**, 473–497.
Steinmetz, G. (1993), *Regulating the Social: The Welfare State and Social Politics in Imperial Germany*, Princeton, NJ: Princeton University Press.
Stinchcombe, A.L. (1968), *Constructing Social Theories*, Chicago, IL: The University of Chicago Press.
Su, Y. (2011), *Collective Killings in Rural China during the Cultural Revolution*, Cambridge, UK: Cambridge University Press.
Traugott, M. (2002), *Armies of the Poor: Determinants of Working-Class Participation in the Parisian Insurrection of June 1848*, New Brunswick, NJ: Transaction Publishers.
Vaughan, D. (1996), *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago, IL and London, UK: University of Chicago Press.
Vicari, S. (2010), 'Measuring collective action frames: A linguistic approach to frame analysis', *Poetics*, **38**, 504–525.

Wallerstein, I. (2011), *The Modern World-System IV: Centrist Liberalism Triumphant, 1789–1914*, Berkeley, CA: The University of California Press.
Weber, M. (1949), *The Methodology of the Social Sciences*, Glencoe: The Free Press.
Weeks, J.L.P. (2014), *Dictators at War and Peace*, Ithaca, NY: Cornell University Press.
Wickham-Crowley, T.P. (1992), *Guerillas and Revolution in Latin America: A Comparative Study of Insurgents and Regimes since 1956*, Princeton, NJ: Princeton University Press.
Whitt H. (2010), 'The civilizing process and its discontents: Suicide and crimes against persons in France, 1825–1830', *American Journal of Sociology*, **116**, 130–186.
Wimmer, A. and Y. Feinstein (2010), 'The rise of the nation-state across the world, 1816 to 2001', *American Sociological Review*, **75**, 764–90.
Zaret, D. (1996), 'Petitions and the 'invention' of public opinion in the English Revolution', *American Journal of Sociology*, **101**, 1497–1555.
Ziegler, R. (2008), 'What makes the Kula go round? A simulation model of the spontaneous emergence of a ceremonial exchange system', *Social Networks*, **30**, 107–126.