Spring 2010, Vol. 8, No. 1

Qualitative & Multi-Method Research

Newsletter of the
American Political Science Association
Organized Section for Qualitative and
Multi-Method Research

Contents

Notes from the Editor

Gary Goertz
University of Arizona
ggoertz@u.arizona.edu

In this issue the section announces a major new award, honoring a key figure in its development and ongoing activities. The David Collier Mid-Career Achievement Award recognizes distinction in methodological publications, innovative application of qualitative and multi-method approaches in substantive research, and/or institutional contributions to this area of methodology (see the next page for the full announcement and details). As most of you will know, David was the driving force behind the formation of the section, authoring the original emails requesting support, guiding the petition through the APSA bureaucracy, and serving as founding President. Since then David has given to the section tirelessly, for example leading short courses and/or working groups at every APSA meeting. David's applied scholarly work is of course widely renowned. David has also made key contributions in multiple areas of methodology, including comparative case analysis, process tracing, typologies, and measurement to name just a few. As a mentor, David has guided the careers of an extraordinary number of scholars. In short, David is the exemplar of the kinds of activities the award seeks to recognize. (An endowment for the award has been established at CQRM by gifts from several scholars, but of course additional support to solidify that fund would be very welcome. Please email me at ggoertz@u.arizona.edu if you would like more details on how to make a gift.)

Multi-method issues continue to provoke much discussion in the section. We have two articles discussing how major projects have dealt with doing research involving dozens of people from all over the world. McBride and Mazur give an overview of the huge RNGS project which is drawing to a close. This project has looked at women's policy issues, has used qualitative, quantitative, and case studies, and has produced an impressive corpus of publications, including at least a half a dozen books, not to mention articles, chapters, etc. Amy Poteete reports on a large project arising from Elinor Ostrom's influential (and Nobel Prize–winning) work on common pool resource institutions. She discusses in particular methods of meta-analysis which I think are particularly relevant to qualitative methods scholars. If one conducts a case study almost inevitably the question arises about generalization. Part of the Poteete, Ostrom, and Jensen project is to con-

Call for Nominations: The David Collier Mid-Career Achievement Award

The APSA Organized Section for Qualitative and Multi-Method Research is pleased to announce the establishment of the David Collier Mid-Career Achievement Award.

The Award honors David Collier's contributions—through his research, graduate teaching, and institution-building—as a founder of the qualitative and multi-method research movement in contemporary political science. The award will be presented annually to a mid-career political scientist to recognize distinction in methodological publications, innovative application of qualitative and multi-method approaches in substantive research, and/or institutional contributions to this area of methodology.

To be eligible for the mid-career award, nominees must have defended their dissertation within fifteen years of the beginning of the year in which the award is presented. For the 2010 award, nominees must have defended their dissertation in or later than 1995.

Each nomination must include a cover letter summarizing specific merits of the candidate (not merely generic praise), as well as an up-to-date curriculum vitae of the nominee, including the date of the doctoral degree. Self-nominations are welcome.

Nominations for the 2010 award should be submitted by June 15, 2010, via email to the President of the APSA Organized Section for Qualitative and Multi-Method Research, currently Colin Elman (celman@maxwell.syr.edu).

The David Collier Mid-Career Achievement Award has been established by the Consortium for Qualitative Research Methods (the co-host of the Institute for Qualitative and Multi-Method Research). By agreement with CQRM, the award is managed by the APSA Organized Section for Qualitative and Multi-Method Research.

duct a meta-analysis of case studies on common pool resource institutions. I think there is much in this volume that could inspire qualitative methods analysis and causal generalizations. Personally, I have been thinking and working a lot on these issues, and reading chapters of their book has been very useful in my own research.

Philosophy of (social) science has always been a core part of qualitative methods. Interpretive scholars in particular have worked a lot in this area, relating the conduct of social inquiry to large philosophical concerns. In international relations, Wendt's work has been very influential and it has a strong philosophy dimension. Hence it is very appropriate to have a symposium devoted to Patrick Thaddeus Jackson's forthcoming book. Robert Adcock has organized a great symposium (originally a panel at ISA) on this volume. The book and the symposium raise a number of central issues that many of us address in our qualitative methods courses. One of the strengths of the section is its diversity: philosophy of social science is one place where our diversities can meet and dialogue.

As has become customary in the Spring issue of the newsletter, we list the section's panels for APSA 2010. Giovanni Capoccia (giovanni.capoccia@ccc.ox.ac.uk) has done a great job in managing and organizing the panels. We have about 20 panels this year. As usual there is considerable diversity in the panels which reflects the diversity of the section membership. Panels may change between now and APSA so you should check the APSA website for the final program. I also encourage members to attend the business meeting and reception at APSA.

The business meeting is usually on Thursday evening. It is a good chance to hear about the section; it is an even better occasion to have some interesting chats during the reception. Check the section website for information about the business meeting, receptions, workshops, and other events.

Symposium: Patrick Thaddeus Jackson's *The Conduct of Inquiry in International Relations*

Introduction: The Voice of Philosophy of Science in the Conversation of QMMR

Robert Adcock
George Washington University
adcockr@gwu.edu

The voice of philosophy—more particularly the philosophy of science—has been heard in the conversation of this section at multiple points: in discussions of ontology and epistemology (Yanow 2003; Chatterjee 2009), causality and explanation (George and Bennett 2005: Chap. 7; Dessler 2006; Lieshout 2007), concepts (Choi 2005; Goertz 2006: Intro.), and more. Patrick Thaddeus Jackson's new book, *The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the Study of World Politics,* vigorously sets out to redirect, broaden, and organize how we hear and deploy the voice of philosophy of science in such discussions. A book of this range and ambition can, as the three symposium contributions below well attest, be engaged at multiple levels in varied ways. I seek here to introduce the newsletter's readers to the general character, organizing structure, and pluralistic spirit of Jackson's endeavor, so as to set the stage for the close critical engagement of the symposium's contributors with more specific aspects and arguments of the book.

To introduce Jackson's philosophical endeavor, we might start on more familiar ground and follow through where he would take us from there. For members of this section such familiar ground can perhaps be found in the influential current of interest in causal mechanisms. During the last two decades attention to causal mechanisms has been variously advocated as an essential complement, corrective, or challenge to the study of empirical patterns of cross-case covariation. This advocacy has presented itself philosophically as drawing on "scientific realism" to supplant (or in hybrid formulations, to supplement) "positivist" constant-conjunction accounts of causality and/or deductive-nomological accounts of explanation. A shift at the level of method has hence been articulated as enacting a change at the level of philosophy of science. At both levels there has, however, been more consensus on the target of dissatisfaction than on what the remedy of attending to causal mechanisms actually involves.[1]

How would Jackson's approach to the philosophy of science redirect this influential, if ambiguous, discussion? First he pushes toward more precision in labeling stances, preferring to talk of *neo*positivism (since "positivism" unqualified loosely gestures at a complex tradition of philosophically diverse views) and *critical* realism (since the implications of realism for social science have been most fully explicated by the school of scholars who identify in these terms). Second, Jackson encourages us not to presume, or attempt to argue for, the general philosophical superiority of neopositivism or critical realism. Third, he points away from using the concept of mechanism to try to capture how these stances contrast in their views of causality. That concept has, at this point in time, both neo-positivist and critical realist meanings, and while each stance may wish to legislate its own meaning as *the* meaning, for Jackson such conceptual power plays are inimical to the pluralistic use of philosophy of science he favors. To explicate a contrast here without privileging either stance, Jackson would direct attention toward a concept less variously used and claimed than mechanism. The concept he spotlights is causal powers (a.k.a. capacities), central to critical realist accounts of causality, but foreign to neopositivism. In calling attention to causal powers Jackson extends a rising interest among political scientists seeking to clarify alternative philosophical views of causality (e.g., Chatterjee 2009; Saleh 2009).

If Jackson's book consolidates incipient trends in certain moves, it comes uniquely into its own as he endeavors to broaden how we talk about philosophy of science. This broadening is exemplified by what could, I expect, become Jackson's most noted intellectual departure: his treatment of "analyticism" as a third stance that debates neopositivism about causality and other matters, but does so along a different philosophical axis than critical realism. Jackson's account of the philosophical basis and methodological implications of analyticism is critically engaged in close detail by Eric Grynaviski's contribution to this symposium. The stance is again highlighted in Banks and O'Mahoney's contribution as they explore what graduate students could learn from Jackson's book. They find its account of analyticism especially eye-opening: spurring them to re-envision the methodological character, both of their own dissertation projects, and of prominent examples of international relations scholarship.

Jackson's broadening push also extends beyond analyticism to identify a fourth stance that he labels "reflexivity." His book's four central chapters—chapters three through six—take up neopositivism, critical realism, analyticism, and reflexivity in turn, explicating each stance's philosophical commitments and their implications for methodological topics such as causality and comparison. Proliferating stances is, of course, no inherent virtue. But Jackson counters its potential pitfalls by pursuing broadening within an ideal-typical organizing framework. Before delving into the four stances, he first introduces them in chapter two as positions in a 2 x 2 table (reproduced on the following page)

Jackson constructs this organizing framework by analytically focusing attention onto two basic philosophical choices, which then provide the axes of his 2 x 2 table. The horizontal axis concerns the relationship between knowledge and observation, with the philosophical choice being between the view ("phenomenalism") that scientific "knowledge is purely related to things that can be experienced and empirically observed" and the competing view ("transfactualism") that "it is possible

		Relationship between Knowledge and Observation	
		Phenomenalism	Transfactualism
Relationship between Researcher and Researched World	Mind-world Dualism	*Neopositivism*	*Critical Realism*
	Mind-world Monism	*Analyticism*	*Reflexivity*

to generate knowledge of in principle unobservable objects" (2010: 36). While some terminology here may be novel to newsletter readers, the philosophical question of the scientific standing of unobservables should be more familiar since it has played a major role in discussions of causal mechanisms (George and Bennett 2005: Chap. 7; Gerring 2008).

Probably less familiar, however, is the philosophical choice that provides the vertical axis in Jackson's table. For Jackson, the debate between neopositivism and scientific/critical realism invoked in causal mechanism discussions leaves unquestioned a shared commitment to "mind-world dualism." Identifying and explicating a competing belief of "mind-world monism" is thus pivotal for his effort to broaden usage of the philosophy of science beyond these more familiar stances. Since the philosophical question at stake may be new for some readers, and part of the terminology is a neologism, it is worth quoting Jackson's formulation here at more length:

> [T]he relationship or connection between the researcher and the researched world... presents an ideal-typical choice between *mind-world dualism* and its opposite, which I will call *mind-world monism*. The former option maintains a separation between researcher and world such that research has to be directed to properly crossing that gap, and valid knowledge must in the end be related to some sort of accurate correspondence between empirical and theoretical propositions on the one hand and the actual character of a mind-independent world on the other. The latter, on the other hand, maintains that the researcher is a part of the world in such a way that speaking of "the world" as divorced from the activities of making sense of the world is literally nonsensical: "world" is endogenous to social practices of knowledge-production, including (but not limited to) scholarly practices, and hence scholarly knowledge-production is in no sense a simple description or recording of already-existing stable worldly objects. (2010: 35–36)

While the distinctions organized in Jackson's table structure much of his book, the spirit of his arguments is decisively shaped by the pluralistic attitude he advocates toward the stances in his table. Jackson carefully labels the choices that provide the axes of his table as "wagers" to reflect his view that they are philosophically irresolvable, and thus political scientists cannot and should not expect philosophy of science to resolve which stance they should adopt. This reflects Jackson's guiding pluralistic commitment: introduced by chapter one's advocacy for a "broad Weberian definition" of science intended to find room within the house of science for all stances explicated in later chapters, and rearticulated by his conclusion's plea for a pluralism that would refuse any stance in his table a monopoly, or even predominance, over how political scientists hear and deploy the voice of philosophy of science. The prescriptive ideal of a pluralistic relation to philosophy of science that motivates Jackson's book is highlighted and critically engaged in John Gunnell's symposium contribution. Drawing on some four decades of experience studying and contributing to political science's evolving interaction with philosophy of science, Gunnell contextualizes and interrogates Jackson's attitude toward this relationship, his use of the –ology triad (ontology, epistemology, methodology), and his plea for pluralism. The symposium closes with a lively response from Jackson to the contributions of Grynaviski, Banks and O'Mahoney, and Gunnell.

Note

[1] In reviewing the by now profuse literature on mechanisms, Gerring (2007: 3, 6) suggests that "the best entrée into contemporary usage of 'mechanism' is by looking at what it is *not*," before then going on to document the "patently contradictory" moves taken by different advocates of causal mechanisms when they go from criticism to the positive work of defining what mechanisms are and how to study them. Evaluated from the perspective of the *sociology* of science this pattern is entirely reasonable: what we have here is an appealing banner phrase both broad and vague enough to attract and hold together a sufficiently large coalition to potentially make a difference in the norms of the discipline. However, evaluated from a perspective prioritizing, as Jackson does, the *philosophy* of science, a banner that serves coalitional ends so ably is perhaps pretty much guaranteed to appear disappointingly lacking in clarity.

References

Chatterjee, Abhishek. 2009. "Ontology, Epistemology, and Multi-Methods." *Qualitative and Multi-Method Research* 7:2 (Fall), 11–15.

Choi, Naomi. 2005. "Crafting Explanatory Concepts." *Qualitative Methods* 3:2 (Fall), 24–29.

Dessler, David. 2006. "Case Studies and the Philosophy of Science." *Qualitative Methods* 4:1 (Spring), 43–45.

George, Alexander L. and Andrew Bennett, eds. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.

Gerring, John. 2008. "The Mechanismic Worldview: Thinking Inside the Box." *British Journal of Political Science* 38:1 (January), 161–179.

Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.

Jackson, Patrick Thaddeus. 2010. *The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the*

Study of World Politics. London: Routledge.
Lieshout, Robert H. 2007. "A Note on Causality and Causal Mechanisms." *Qualitative Methods* 5:2 (Fall), 18–21.
Saleh, Nivien. 2009. "Philosophical Pitfalls: The Methods Debate in American Political Science." *Journal of Integrated Social Sciences* 1:1 (December), 141–176.
Yanow, Dvora. 2003. "Interpretive Empirical Political Science: What Makes this Not a Subfield of Qualitative Methods." *Qualitative Methods* 1:2 (Fall), 9–13.

Do Our Philosophical Commitments Matter?

Eric Grynaviski
Ohio State University
grynaviski.2@polisci.osu.edu

International Relations scholars have produced books on the scientific status of the discipline like Khrushchev claimed the Soviets produced missiles: like sausages. One enticing element of Jackson's *The Conduct of Inquiry in International Relations* is the orthogonal perspective he brings to this growing debate. The point of this cottage industry, argues Jackson, is not simply to provide a foundation for the conduct of inquiry, but all too often to engage in polemics: "when 'science' makes an appearance, it is a pretty good bet that the text in which the term is invoked is more or less explicitly trying to reshape how inquiry is conducted, and doing so by drawing on the rhetorical power of 'science' in order to privilege some modes of inquiry at the expense of others" (2010: 9–10). Jackson compares these debates about science to "bringing out the big guns" and "playing with fire." The success or failure of Jackson's project hinges on his ability to give an account of the logic of inquiry that does not define away important approaches to IR as unscientific. This is a difficult task because the disciplinary stakes are high: debates within the IR community over the meaning of science have major implica-tions for the type of research valued and the standards used to assess work in our community.

The bulk of Jackson's book is a sophisticated breakdown of different philosophical commitments that IR scholars (often unknowingly) make when engaging in empirical analysis. The first "wager" is a stance on the relationship between knower and known. Dualists treat the observer as independent of the observed, whereas monists treat observers as participants in the world in which they work. The second 'wager' concerns the relationship between knowledge and observation. Transfactualists aspire to gain knowledge concerning things that are in principle unobservable, whereas phenomenalists limit their claims to that which can be verified through our experiences in the world. These differences in the philosophy of science create a typology of philosophies of science (reproduced in Adcock's contribution). This typology is a useful—if blunt—way of understanding the differences between alternative philosophical positions in IR, an asset for teaching the varieties of philosophical approaches to scientific questions, and a helpful common vocabulary for debate concerning these positions in IR. First, the axes capture important differences in the philosophy of science. Second, presenting these differences as 'wagers' is intriguing, suggesting the possibility that a tradition of research, at some unlikely moment in the future, might be blotted out by the smudge of a philosopher's pen. Perhaps most importantly, in elaborating philosophical stakes underlying methodological debates and choices, Jackson writes clearly and with a sense of purpose, making an abstract and often dull subject matter lively and interesting.

Jackson's central claim—the definition of science should be broad enough to incorporate different perspectives—will be well appreciated by scholars weary of debates concerning the "scientific" status of different forms of inquiry. The success of *Conduct*, however, hinges primarily on how well it describes different approaches to the conduct of inquiry in IR. I will consider one approach: analyticism. In doing so, I ask two questions related to the usefulness of Jackson's typology. First, does Jackson's analyticism follow from the philosophical wagers of monism and phenomenalism? Second, is his paradigmatic example, Waltz's *Theory of International Politics*, best interpreted as an example of analyticism? Before exploring these two questions, I first summarize the main themes of analyticism.

Analyticism

Jackson situates analyticism in contrast to alternative methodological accounts of science. He argues that neopositivism, for example, stipulates that (true) knowledge is an "unfalsified conjecture." Neopositivists thus commit themselves to the dualist wager because they presume the separation of the researcher from externally existing phenomena (only if a phenomenon is external to the observer can it falsify predictions). By contrast, knowledge seen through an analyticist lens is a "useful account." Here Jackson draws heavily on the pragmatic tradition in philosophy, where the gap between observer and observed is relaxed: the truth of a proposition does not stem from a failure to falsify, but its usefulness to the inquirer (2010: 126). Unlike in his discussion of alternative approaches, however, Jackson's reliance on any single tradition of philosophical argument is less clear in his discussion of analyticism. In sketching the position's intellectual history, he draws on Nietzche; in articulating a concrete methodology for the social sciences, Weber. It is fair to refer to the analyticist enterprise in general as pragmatist because when Jackson philosophically distinguishes analyticism on both dimensions in his typology, he relies on Dewey and James.

The key difference between analyticism and neopositivism is that the former concentrates on the use-value of knowledge. Unfortunately, Jackson's account of use-value is underdeveloped. His claim is explicitly *not* that knowledge is useful if it helps us do things one would usually describe as useful, such as create predictions concerning state behavior, engage in successful foreign policy practice, or fundamentally alter our beliefs about the general course of international affairs. Instead, a concept is useful (practical) if it can be applied in practice by a researcher to illuminate causal relations in a single case. Use here has little to do with practical results, only scholarly process.

Jackson reaches this conclusion through three steps. First, his analyticism rules out, through a philosophical wager, the external world as existing independently from inquiry. Therefore, "it makes little sense to formulate and test hypotheses because the idea of an existing world against which to test them is nonsensical" (2010: 38). This first step is absolutely critical. By ruling out dualism, Jackson claims that all monists (and in particular pragmatists or analyticists) rule out conventional testing. Second, research is a practical activity and as such is governed by intersubjective rules. Different communities of researchers use different theoretical models (instruments) and therefore reach different conclusions. Structural realists and structural Marxists, for example, rely on different theoretical vocabularies, describe the world in drastically different ways as a result, and reach different conclusions. Third, and most important, a community's tools (ideal-type models) constrain knowledge-claims, not the practicality of knowledge as conventionally understood (does it generate novel predictions? can it be relied on in practice?) or in reference to the external world (is my claim debunked through successive testing in the actual world?) (2010: 140–141). This, in sum, is the monist wager: the analyst does not treat the world as a separate and objective fact, but cuts it in new and potentially productive ways by using intersubjective disciplinary rules and theoretical models to explain specific cases.

After laying philosophical foundations for analyticism, Jackson turns to the "concrete research methodology" (2010: 141) that he sees as following from them. This methodology uses ideal-types to organize data, relies on counterfactuals to demonstrate causation, and only does so for single cases (singular causal analysis). Unlike neopositivism, where the researcher is treated as separate from the studied phenomena, analyticism sees the researcher as intimately connected to it. When a researcher approaches a subject of study, her values and beliefs affect her account of the units of analysis. This means that, contrary to neopositivism, a direct representation of reality cannot be generated, rather; one develops "ideal-types," theoretical simplifications that are useful for explanatory goals (2010: 142–143).

The analyticist commitment goes further by ruling out conventional methods of formulating and testing falsifiable general hypotheses. Ideal-types, as abstractions that single out only specific properties of concrete reality, never fully describe a concrete case. Waltz, for example, develops ideal-types of states by abstracting the level of power a state has and its security-seeking motive from the concrete case of any particular state (which may have a broader range of goals and attributes). The result is that no analysis premised on ideal-types can generate causal claims that work across cases: the abstracted factor is only one element of a specific concrete case and alternative causal factors are always at work.

Analyticists, therefore, emphasize "singular causal analysis." Their research explores whether, in a specific case, the causal process identified through the depiction of the ideal-type and its theorized influence are adequate to explain the outcome, or whether there are also "coincidental" causes at work that are not identified in the ideal type. Since an ideal-type is inevitably unable to fully explain a concrete event, focusing on it serves to reveal case-specific factors at work in creating the effect. Thus, there is little ability to generalize because the findings are always case specific; the logic of inquiry does not create the preconditions for a conduct of inquiry that would enable cases to be fruitfully compared. The analyticist is thus committed to creating analytical narratives: accounts of specific events in which ideal-types organize the data (selecting the relevant features of a concrete situation), thereby illuminating causal processes (2010: 146–151).

Making Sense of Philosophical Wagers

Do the philosophical wagers Jackson highlights make a difference? Jackson's overarching claim in *Conduct* is that we need to understand the relationship between wagers in the philosophy of science, made in this case by the pragmatists (along with Nietzsche, Heidegger, Wittgenstein, and others), and then determine what type of methodology is supported by a given combination of wagers. I want to return to the pragmatists, the authors Jackson uses to lay out philosophical foundations of the analyticist position. Pragmatic commitments tend to be very different from Jackson's analyticist commitments, casting doubt on the utility of Jackson's method of differentiating between methodological approaches to IR.

Jackson relies on the pragmatists to develop a monist account of science that collapses the distinction between observer and observed. Then, in developing an account of how a monist conduct of inquiry should proceed, he commits pragmatists to several arguments:

> (1) Pragmatists are not interested in generating useful predictions about the future;
>
> (2) Pragmatists do not believe in cross-case comparison as a useful research technique;
>
> (3) Pragmatists provide no role for falsification in their model of science.

On all three counts, Jackson's interpretation of pragmatism, and perhaps also of analyticism, is inaccurate.

First, pragmatic inquiry is always concerned about predictions. Justifying a belief in the pragmatic tradition, Eric MacGilvray (2004: 40) explains, is a prospective enterprise: "all inquiry whatsoever concerns matters whose resolution lies in the future, and the degree of fit between a given norm of action and the possibilities that the world will admit can only be determined over time." If a causal claim, for example, has been unable to satisfy our doubts because it cannot explain an important case or cases (i.e., fails to generate results when we act upon it), we adjust our belief that a claim satisfies our doubts and leads to predictive success. Charles Peirce (1998: 273), the founder of pragmatism, writes that pragmatism is not about a "dead past," but rather about "what *surely will* happen to everybody in the living future who shall fulfill certain conditions."

Jackson also commits pragmatists to singular causal analysis, that is, to the investigation of singular cases. Here Peirce is again instructive. The aim of an experiment may be to "verify a hypothesis" specific to that experiment (a singular causal analysis), but the pragmatist does not limit herself to singular causal analysis: "the sum of the experimental phenomena that a proposition implies makes up its entire bearing upon human conduct" (Peirce 1998: 272–274). The test of a proposition should include all of its observable implications, across a range of cases. Further, if a belief is only suited to explain a single case, somewhere in the past or present, and bears no lessons for the future, then there is no cash value. And, the only way to determine the projectability of past lessons into the future, presumably, involves a comparison of future situations with those in the past. For Pierce, one must compare the conditions that made events occur in the past to present and future conditions to see if an effect will again follow a cause.

Only because Jackson mistakes the first two pragmatic positions is he able to mistake the third; that pragmatists do not posit generalized propositions. This is far from the case. Dewey's *Logic* holds that the outcome of many investigations is the creation of generalized propositions. One example of progress, according to Dewey, is when a previously thought to be true general proposition is contradicted by a singular negative judgment (Dewey 1998: 196–197). "All states are security seekers," for example, is a general proposition; it is untrue if one believes that U.S. cooperation with Israel undermines U.S. security (Mearsheimer and Walt 2007). The negation of a general proposition leads to further inquiry: why does the United States not pursue its own security in a specific case? After investigation, a new generalized proposition is postulated: states pursue their own security in cases where powerful foreign lobbies do not exist. If this proposition is successful in explaining future cases then it is "true" in the sense that it is instrumentally useful for explaining state behavior. If it is inaccurate, then the general proposition should be continually modified until it is either abandoned or proves useful. While Jackson is right that analysis might focus on single cases, he is wrong to conclude that the aim of inquiry for pragmatists is to understand only those cases: pragmatists' aim is to formulate new knowledge useful in the future.

None of these criticisms show that Jackson's analyticist methodology is wrong, without use, or lacks foundations. Instead, the discussion of pragmatism shows that Jackson's argument that the philosophical wagers one takes determine the concrete research methodology one should pursue is too strong a claim. At least from the pragmatic perspective, it is off base. Scholars committed to the combination of monist and phenomenalist wagers can pursue strategies more akin to conventional neopositivist political science without violating any principled philosophical position. If this is right then a central premise of the book, that different philosophical wagers lead to different methodologies, is inadequate. It suggests that the divisions within the philosophy of science that Jackson spotlights are underdeterminate: they do not enable one to pick specific methodological commitments as the result of philosophical commitments.

Waltz and Foreign Policy Decision-Making

All of these problems feature in Jackson's analysis of Waltz's structural realism. Throughout his treatment of analyticism, Jackson treats Waltz's *Theory of International Politics* as a paradigmatic case of the use of ideal-types in generating singular causal analyses. Jackson argues that Waltz did not intend his *Theory* to suggest a covering law related to balancing, and that Waltz did not believe his core theoretical claims directly represented a deep reality: in the first chapter of *Theory*, Waltz (1979: 9–10) claims that a theory is a simplification of a complex reality into simple abstractions that are of causal significance. This is, certainly, similar to the argument for the formation of ideal-types from Weber.

The problem with Jackson's treatment of Waltz does not arise from the development of ideal-types, but rather his claim that Waltz follows Jackson's analyticist methodology to the point at which it abandons efforts to propose falsifiable general hypotheses or cross-case conditions for outcomes (2010: 151–152). Jackson argues that if Waltz's ideal-types abstract security-seeking motives and levels of power from a more complicated set of motives and attributes that states might have, then other factors (coincidental causes) might arise in any given case that create a deviation from the theory's predicted outcome. There is no meaningful sense in which a falsifiable hypothesis generalizing across multiple cases might be developed because these coincidental causes vary by case. Therefore, the only plausible test of Waltz's theory, using his "own clearly-declared allegiance to such a methodology," (2010: 151) is to focus on specific cases to see if the pressures at work in Waltz's theory help structure explanations. The effort of those engaged in testing Waltz's theories, interpreted as general propositions about balancing behavior over time across different cases, "tells us precisely *nothing* about the utility of Waltz's ideal-type for the analysis of the cases to which he actually applied it. The only thing that would count for or against the use of the Waltzian model would be a concrete, case-specific application of the ideal-type to these and other specific situations." (2010: 153)

The interesting move in Jackson's analysis of Waltz is that he converts what many regard as a problem with the theory—its non-falsifiability—into a theoretical advantage. By abstracting from the concrete situations of states that are pushed toward certain forms of foreign policy behavior by a host of factors (only one of which is power), Waltz leaves open the possibility that the causal process he is interested in (the effect of the distribution of capabilities on the stability of the international system) is always present but never powerful enough to generate predicted outcomes. For many, this would make the theory less useful: without a specification of the conditions under which Waltz's theory becomes a predictive theory, it cannot be applied with any degree of rigor. For Jackson's analyticism, it becomes a model of ideal-typical analysis.

While analyticism captures the importance of ideal-types, Jackson's analysis, unfortunately, also implausibly attributes three claims to Waltz: *Theory* is non-falsifiable, cross-case generalizations regarding the balance of power are unproductive, and generalization is not an important purpose for case studies.

First, Jackson claims that Waltz does not purport to make a falsifiable general claim. This, however, is not Waltz's "self-declared position." Rather, Waltz contends that falsification is not the *only* path to knowledge, particularly in a subject domain in which falsifying hypotheses is very difficult. So, Waltz (1979: 124) argues that one should pursue several strategies of testing that are familiar to the neo-positivist tradition: "Given the difficulty of testing any theory, and the added difficulty of testing theories in such nonexperimental fields as international politics, we should exploit all of the ways of testing I have mentioned—by trying to falsify, by devising hard confirmatory tests, by comparing features of the real and the theoretical worlds, by comparing behaviors in realms of similar and of different structure." Against Jackson's analyticism, which maintains that falsifications are not useful, Waltz argues that falsifications are difficult, and therefore alternative mechanisms for testing general hypotheses are important.

Second, the claim that Waltz does not generate trans-historical claims is hard to sustain. In "Structural Realism after the Cold War," Waltz (2001: 54) writes of the American unipolar moment: "Theory enables one to say that a new balance of power will form but not to say how long it will take" (see also Waltz 1979: 124). The recurrence of balancing is not the only trans-historical prediction that Waltz makes: he makes claims related to systemic stability, the military and economic effects of polarity, emulation, and the death of states that do not pursue security. Each of these general hypotheses can be tested, and Waltz (1979: 124) clearly argues they should be: "Any good theory raises many expectations. Multiplying hypotheses and varying tests are all the more important because the results of testing theories are necessarily problematic. That a single hypothesis appears to hold true may not be very impressive. A theory becomes plausible if many hypotheses inferred from it are successfully subjected to tests."

Third, the reason Waltz gives for in-depth case studies is different from that which Jackson suggests. For Jackson, the aim of generating ideal-types is to explain a case. For Waltz, the reason to explain a case is to confirm a theory. This difference is crucial. Jackson's analyticist draws on ideal-types to explain a historical event: the drawing of ideal-types is not connected to the development of a theory that applies generally because analyticism is not a search for universal historical rules. By way of contrast, for Waltz (1979: 125), the underspecification of his theory of balancing (the possibility of domestic variables) means that one should select a hard case: "if we observe outcomes that the theory leads us to expect even though strong forces work against them, the theory will begin to command belief." The reason to concentrate on specific historical cases is that alternative variables (e.g., domestic politics) should lead one to predict a different outcome than that predicted by Waltz's theory. Waltz argues that the pattern of behavior by the superpowers after World War II—rearmament—is strong evidence that balances recur because domestic explanations would predict disarmament: the costs of war were apparent, as was domestic opposition to peacetime military spending. "These examples tend to confirm the theory. We find states forming balances of power whether or not they wish to" (Waltz 1979: 125).

Waltz's strategy of case selection shows just how different his self-understanding of the purpose of selecting cases and producing theory is from Jackson's analyticism. For Jackson, the analyticist decision to use ideal-typical modeling "is *not about making a direct contribution to empirical generalizations*" (2010: 151). In contrast, Waltz selects hard cases, stacking the deck against his theory, to show that it is empirically generalizable. Waltz's understanding of theory testing, in which one develops a theory to explain a set of pressures that apply in a wide variety of cases, is precisely the type of analysis that Jackson's analyticists oppose (but pragmatists may support).

While reading Waltz as an advocate of singular causal analysis may mistake Waltz's intentions, Jackson's analyticist framework could clarify a nascent debate within the study of foreign policy. Snyder, Bruck, and Sapin's *Foreign Policy Decision-Making*, an analyticist work that inaugurated the field of foreign policy, is rarely treated for its philosophical assumptions. The authors make three arguments that clearly fit the analyticist position.

First, Snyder, Bruck, and Sapin posit that a "frame of reference" is crucial to the study of foreign policy. The concept of a frame of reference draws on Parsons' (1937: 28–34) early work, *The Structure of Social Action*, in which a frame of reference is an idealized model of reality that highlights certain aspects of a concrete agent at the expense of a full description. In the study of foreign policy, this frame of reference orders the field of study and specifies the causal factors that are suspected to be at work (2002: 30–35). Like Jackson's analyticism, the frame of reference develops ideal-types that depict an abstracted element of reality in order to organize (not represent) reality.

Second, the frame of reference, for Snyder, Bruck, and Sapin, depends on the specific values that the researcher brings to bear on a question, and the researcher should make this value commitment explicit: "if there is a Marxist or World Government value cluster which serves as an organizing principle and if the frame of reference is coherent and explicit, that fact should be abundantly clear" (2002: 31). In the same way that a researcher's values lead to the formulation of ideal-types, Snyder, Bruck, and Sapin claim that what one looks for influences how the frame of reference, the ordering of reality via ideal-types, is performed.

Third, while Snyder, Bruck, and Sapin point toward generalizable propositions, the actual course of the field since their seminal work has emphasized intensive case studies. Yuen Foong Khong's (1992) *Analogies at War*, for example, highlights the use of analogical reasoning in Johnson's decision to escalate the Vietnam War, and Elizabeth Saunders (2009) recent article "Transformative Choices" emphasizes the influence of specific clusters of beliefs on Vietnam era–decision-making. In both cases, the selection of analogies and beliefs organizes the decision-makers' choice situations according to the researchers' interests. In both cases, the test of the theory, jargon aside, is to see how satisfyingly it explains the case at hand: singular causal analysis. Moreover, while both authors undertake comparative studies of decision-making, the value of these comparisons is largely that they illuminate how unique the decision-makers and the cases were: had a different decision-maker, who held different beliefs, occupied the position of decision-making authority, a different decision may have been made.

Jackson's account of analyticism, applied to the study of foreign policy decision-making, highlights much that is unique about this area of IR scholarship. The use of single case studies, the creation of ideal-types, and the notion that every policy-maker encounters a complicated and unique situation require a way of understanding scientific method that is distinct from neopositivism, scientific realism, or postmodernism. Jackson's chapter, in this context, might be read with great profit.

Conclusion

Jackson's analyticism neither follows from its assigned philosophical basis in pragmatic theories of scientific inquiry nor is born out in the paradigmatic example of Waltz's *Theory of International Politics*. While these are objections to the manner in which Jackson moves from philosophical wagers to a Weberian methodology, they do not mean that his chapter on analyticism is without merit. It is an eye-opening discussion, both because it is ambitious and because it sheds a different light on a tradition that is under-theorized in IR. While I am not convinced that a combination of monist and phenomenalist philosophical wagers requires us to abandon empirical generalization, I am persuaded that the concrete research methodology that Jackson proposes might be useful if pursued.

References

Dewey, John. 1986. *Logic: The Theory of Inquiry*. Carbondale: Southern Illinois University Press.

Jackson, Patrick Thaddeus. 2010. *The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the Study of World Politics*. London: Routledge.

Khong, Yuen Foong. 1992. *Analogies at War: Korea, Dien Bien Phu, and the Vietnam Decisions of 1965*. Princeton: Princeton University Press.

MacGilvray, Eric. 2004. *Reconstructing Public Reason*. Cambridge: Harvard University Press.

Mearsheimer, John and Stephen Walt. 2007. *The Israel Lobby and U.S. Foreign Policy*. New York: Farrar, Straus, and Giroux.

Parsons, Talcott. 1937. *The Structure of Social Action: Volume One*. Boston: Free Press.

Peirce, Charles. 1998. *Charles S. Peirce: The Essential Writings*. Amherst, NY: Prometheus Books.

Saunders, Elizabeth. 2009. "Transformative Choices: Leaders and the Origins of Intervention Strategy." *International Security* 34:2 (Fall), 119–161.

Snyder, Richard, H.W. Bruck, and Burton Sapin. 2002. *Foreign Policy Decision-Making (Revisited)*. New York: Palgrave Macmillan.

Waltz, Kenneth. 1979. *Theory of International Politics*. Boston: McGraw Hill.

Waltz, Kenneth. 2001. "Structural Realism after the Cold War." In *America Unrivalled*. G. John Ikenberry, ed. (Ithaca: Cornell University Press), 29–67.

The Conduct of Inquiry in International Relations: The View from Graduate School

David E. Banks
George Washington University
debanks@gwmail.gwu.edu

Joseph O'Mahoney
George Washington University
omahoney@gwmail.gwu.edu

Jackson's book, *The Conduct of Inquiry in International Relations*, is most likely to be assigned or recommended in graduate classes addressing the philosophy of science, qualitative methodology, and research design. It might then be useful to ask two graduate students whether this is a good idea. How helpful is yet another book on the meta-theoretical status of International Relations? Our answer to this question has four parts. First, we ask whether and how Jackson's ordering scheme clarifies debates in IR. Second, we discuss the consequences of the scheme for understanding the state of research in IR. Third, we outline the prescriptive consequences of the scheme for graduate students and our own research. Finally, we present three limitations on the usefulness of Jackson's book to budding scholars.

A Clarifying Contribution

At its most basic, this book helps to map the contours and confusions of many debates in IR; ably drawing the links between them, while also debunking much of what scholars think is at stake here. Perhaps the greatest debunking Jackson provides is his stress that the meaning of "science" is still unsettled, and that authors like Lakatos, Kuhn, and Popper—names that all IR graduate students are familiar with—spent much of their time manifestly disagreeing with each other, not building towards a consensus position. This fact alone helps the graduate student breath a sigh of relief. If these philosophers of science never agreed, it is no wonder that IR research is so confused about its scientific underpinnings.

To explicate what science can mean in IR, Jackson introduces his ordering idea of philosophical ontology and the four ideal-type categories that make up the book's 2 x 2 table of philosophical-ontological positions (see the table in the introduction to this symposium). In an arena of debate already cluttered with difficult concepts, this table actually provides welcome relief. A common barrier to graduate student attempts to navigate the waters of the philosophical underpinnings of inquiry is that those waters are often muddy. Opacity of prose can deter a student from spending valuable time on these matters since their empirical topic of interest or methods might seem to have a more direct payoff. It can also mean that time spent on philosophy of science or social science is less profitable than it might be. By contrast, Jackson's writing in this book is very accessible, especially considering the subject matter.

When engaging philosophical matters graduate students can find confusing the forest of terms designating positions that are often only comprehensible in relation to a specific debate. Jackson alleviates this problem in two ways. First, he generates neologisms for some of his conceptual categories. There are startup costs to this strategy since readers must deal with even more new terms. However, once past this hurdle, the cost is outweighed by the benefits of clarity. The reader is told what transfactualism, analyticism, etc., are and is unlikely to confuse them with other more well-known terms that mean something different. Jackson's second way of mitigating potential confusion is placing the ideas he discusses in context, both of work in the philosophy of science and metatheoretical debates in IR. For example, learning that Lakatos' conception of progressive and degenerative research programmes sought to account for the demonstrated empirical success of physics helps indicate how this conception may or may not be applicable to IR. For one thing, physicists were successful long before there were philosophers of science to tell them what to do. The theme that philosophy of science should not be a narrowly prescriptive exercise is accompanied by an exhortation to philosophical awareness of what it is that one is trying to do with one's research and of whether that makes sense. Most critically, the four varieties of philosophical ontology explicated by Jackson give the graduate student eager to engage the underpinnings of IR inquiry a clear sense of what's what in the discipline: both practically and theoretically. We would warn readers, however, that the initial presentation of Jackson's central concept of philosophical ontology is not especially clear. As graduate students socialized into a particular vernacular, we kept waiting to see how philosophical ontology linked into IR debates about ontology and epistemology. It does not become clear until later why this traditional dichotomy was ignored—i.e., it presupposes a dualist position—but if mentioned earlier this would have made Jackson's framework understandable sooner.

"Science" in IR: Rhetoric and Practice

Jackson ably demonstrates that complaints that a certain piece of research is not "scientific"—due, for example, to it being unamenable to falsification—are too often nothing more than disciplining moves rather than substantive criticisms. Jackson takes a "broad Weberian" stance on what counts as science: it is "empirical inquiry designed to produce knowledge." Science is not differentiated from the category of pseudo-science, then, but from partisan political action. One of the liberating functions of Jackson's position is that it offers a basis for deflating knee-jerk denunciations of work that is different from one's own as "unscientific" or "not political science." Graduate students cannot help but be desperately concerned that their work be taken seriously as political science, and in order to avoid the charge of being unscientific, they might be motivated to conform to standards of methodology, and of method, that are widely regarded as scientific rather than those that make sense to them. The overriding theme of Jackson's view of science is a call to intellectual honesty that encourages researchers not to allow themselves to be forced down certain paths. Acknowledging that the function of the commonplace "science" is often to discipline—that is to try and "reshape how inquiry is conducted" by "drawing on the rhetorical power of 'science'"—can take the sting out of charges that some modes of inquiry are not worthwhile.

A significant contribution of Jackson's book is his discussion of the dominance of neopositivism in IR, and how it is frequently considered the only real definition of science in the discipline. Once one recognizes this, it becomes clear that the big debate about quantitative vs. qualitative methods is, as Jackson stresses, just that: a discussion about method that sidesteps broader ontological or epistemological concerns. Moreover, although constructivism tends to be treated as a "post-positivist" position in the discipline, much of it often professes to adhere to the mind-world dualism and phenomenonalist position of the most committed neopositivist. Constructivism vs. rationalism is thus commonly a battle over what Jackson calls scientific ontology, not underlying philosophical principles.

This dominance of neopositivism is somewhat ironic once the actual practice of IR research is considered. As Jackson occasionally points out, IR as practiced rarely adheres strictly to the tenets of neopositivism and frequently veers toward analyticist or critical realist positions, even when authors do not necessarily think this is what they are doing. Highlighting this fact provides both good and bad news to the budding IR practitioner. The good news is that one may be able to dress research in neopositivist clothing without necessarily following its implications to the letter. Consider the attention to causal mechanisms in current IR, for example. Although any truly neopositivist mechanism should in principle be reducible to empirically observable intervening variables and thus subject to falsification (King, Keohane, and Verba 1994: 86), the significance of game theoretic models shows that this is not a necessary requirement for work to be considered scientific—as game theoretic models require the consideration of unobservable off-the-equilibrium-path outcomes. The bad news is that if one is committed to full intellectual honesty and defending one's work on its philosophical merits, this can be hard to do. However unfair it may seem, the dominance of a particular view of science in IR is not something a young scholar can ignore.

Jackson's philosophical-ontological distinctions also put front and center the meaning and use of concepts—in principle, for neopositivists, concepts should only be treated instrumentally. Critical realism, on the other hand, treats many common concepts in IR—such as the state, institutions, social forces etc. —as real-but-unobservable entities, which allows for much more complicated concepts. Yet in IR scholarship, concepts often have an ambiguous status: their place in many theories is far more fundamental than simple placeholders, but they are not necessarily claimed to be real entities. Consider Snyder's (1991) log-rolling coalition of imperial expansion, Finnemore and Sikkink's (1998) norm life cycle model, or Fearon's (1995) war model. These conceptual frameworks do not comfortably fit either the critical realist or neopositivist stances, yet have made significant contributions to the field, and may, we suggest below, be best understood in light of

Jackson's analyticist stance. Philosophical-ontological wagers thus influence the assumptions that scholars implicitly or explicitly make in their theories. Unpacking such assumptions, as Jackson advocates, can clarify what would count as persuasive arguments against the concepts a scholar uses from *within* that scholar's own methodological point of view.

Applying Philosophical Ontology: Consequences for Research

This book was not written as a strategy guide to tell PhD students how to avoid disciplining moves, but as a call to researchers to understand that there is no consensus about what philosophical commitments are scientific and, instead of trying to solve this puzzle, to get on with conducting research in a manner aware of one's own philosophical commitments. For this insight alone the book should be required reading for any graduate student.

Jackson offers much in the way of practical usefulness by showing the significance of philosophical "wagers" for producing coherent scholarship. His presentation of alternative philosophical-ontological positions as "wagers" is especially useful as it emphasizes the unresolved nature of the puzzles about what science is, and suggests that at some point a researcher must make a leap of faith, but should do so with awareness of what they are doing. As Jackson's discussion of monism vs. dualism illustrates, there is no way a dualist can claim there is a distinction between an objective world "out there" and a subjective world "in here" without first subscribing to a mind-dualist position to begin with. In other words, arguments that "facts speak for themselves" fundamentally presuppose the very thing they try to demonstrate. One of the primary lessons from the book is that researchers should not only be clear about what their epistemic aim is, but also what it means to consistently pursue that aim.

If graduate students are able to manipulate the philosophical-ontological categories in Jackson's scheme, that is, if they are able to articulate the link between fundamental epistemic warrant and their research in a way that is accepted by their audience, then the scheme is potentially useful in organizing their thinking. The scheme provides ways to identify logical inconsistencies in the work of others, and avoid them in one's own work. For example, Jackson links the use of case comparison to the epistemic warrant sought for one's claims. Recognizing that there are non-positivist types of epistemic warrant, and that much of how IR is conducted is far from professed neopositivist ideals (for example, Waltz's monist explanation of balancing), liberates scholars from the belief that falsification is a gold standard by which to judge research. Continuity across empirical (actually existing) cases is indeed pivotal for a neopositivist knowledge claim. However, if you are making such a claim, then you cannot consistently resort to unobservables or the idea that your model is just heuristically useful. If you do not intend to rest a knowledge claim on a neopositivist basis, then empirical generalization may be irrelevant to the value of your research. This point is different from the neopositivist idea that single cases or small-n research designs are defensible as part of a broader attempt at establishing causal laws (Rogowski 2004). Neither an analyticist pursuing singular causal analysis nor a critical realist trying to "elucidate the variety of ways that causal properties and the complexes into which they are arranged play out in practice" (Jackson 2010: 111) need be concerned with evaluating the status of a general causal law. The importance of this point cannot be overstated for a graduate student looking to defend their work to a seminar, dissertation committee, or peer group. The ability to defend against methodological criticisms is vital to being a competent scholar. There may be a temptation to resort to one's own stock of buzzwords, like process-tracing, instead of carefully thinking through the kind of epistemic warrant one is seeking, in order to be able to clarify the goals and conduct of one's research. Another temptation is to give in and respond to charges of lack of generalizability with a nominally comparative case design or an "inconsequential data analysis...tacked on as the final one-tenth of the paper" (Clarke and Primo 2007: 749).

It is in the presentation of analytic ideal-types that Jackson is perhaps most interesting. Prior to reading his book, we had not conceived of analyticism as a distinct coherent approach to the philosophical underpinnings of knowledge claims. Some confusion might be expected here insomuch as both neopositivists and analyticists are commited to phenomenalism. However, the monist element of analyticism means that instead of agonizing over falsification, the researcher is encouraged to focus on a theory's internal logical coherence. Indeed, this approach assumes ideal-types will never completely map onto the empirics as all analyticist models fundamentally make equal that which is unequal (Jackson 2010: 124). The value of theory comes not from how well it appears to be empirically true in general, then, but rather how useful it is in a given instance; not from its representational truth, but its explanatory usefulness. If we think again about authors such as Fearon or Finnemore and Sikkink, or models such as the tragedy of the commons or prisoners' dilemma, we see that their importance comes from precisely this kind of analyticist explanatory utility. In effect then, much of the best work in IR may be analyticist at least in part, and Jackson's illumination of this is of great value to the graduate student.

Indeed, insights gleaned from Jackson's account of analyticism have had practical implications for how we are conducting our own dissertation research. Banks' dissertation is focusing on how diplomatic symbols and rituals are manipulated by states. Prior to reading Jackson, his theoretical reflection on this had tried to steer an uncomfortable path somewhere between critical realism and neopositivism. In addition, Banks had worried about how to falsify and test against the effects of symbols and rituals in diplomacy. Now, his focus has shifted to building a logically coherent ideal-type of the diplomatic game that he sees states playing. The analyticist approach directs Banks to understand the value of his theoretical framework in terms of how well it illuminates specific empirical cases, where such illumination may come from the way deviation from the ideal type calls attention to case-specific causes explaining this deviation.

Similarly, Jackson's account of how analyticist research methodology combines ideal-typical concept construction and singular causal analysis has provided O'Mahoney with a more coherent means toward his explanatory end. His work on changes in the rules of state reaction to interstate war outcomes is essentially directed toward explaining a single major historical transformation in these rules. This is the epistemic goal of the research, rather than the use of a single case to test a general causal proposition about institutional change. Analytically general models of institutional change are useful in constructing potential explanations of the single transformation and crucial to a disciplined use of counterfactuals as a way of imagining alternate pathways to those observed. However, the truth value of those models is independent of the empirical findings of the current research. Similarly, the theoretical implications of O'Mahoney's project are potentially more general if they are useful for analytically modeling other historical transformations. But they are not necessarily general in the specifically neopositivist sense of establishing causal laws about what states' reactions to war outcomes are going to be.

Limitations of *The Conduct of Inquiry in IR*

As much as we find to praise in *The Conduct of Inquiry* in IR, we also saw three major limitations in the book. First, Jackson unfortunately confuses a key issue in contemporary IR discourse: the status of hypothesis testing. He associates "hypothesis testing" narrowly with the falsification or verification of empirical generalizations. However, this is not how we have found the phrase used, either in the literature or in discussions amongst graduate students or with professors. A broader conception of hypothesis testing can include specifying what data might be relevant to one's research question before doing the research, or simply being clear about the claims that one is making. It seems that, in Jackson's terms, hypothesis specification and testing is irrelevant to any research that is not neopositivist. Making this claim explicit, especially in the conclusion's prescriptive lexicon, would have been helpful. It would also have been interesting to hear what Jackson would object to about broader uses of the phrase that encourage scholars to be more explicit in such activities, for example, as the forming, appraisal, and revision of hunches in a singular causal analysis.

Second, while the practice of IR rarely fits neatly into one of his four philosophical-ontological boxes, Jackson does not address this issue head-on. We would have liked him to be clearer about if mixing commitments from different philosophical-ontological positions is a mistake. A prominent example of such mixing is research in which a formal model is used to generate predictions and then a statistical analysis is done to empirically test those predictions. Apart from a few asides, Jackson does not address how such research fits in terms of his philosophical wagers. A reader may infer that the formal model is analyticist and the statistical test neopositivist, and therefore there is a possible inconsistency in marrying them. But Jackson does not say exactly what the problem would be. He claims that singular causal analysis is the goal of an analyticist, and says that empirical generalization is logically independent of ideal-types, or that it does not make sense to test an ideal-type against evidence. Indeed he states that an ideal-type is "not available for any kind of direct empirical verification or falsification, in virtue of its roots in a set of value commitments on the part of the...researcher" (2010: 142). This would seem to imply that the conclusions that come out of formal models and statistical analysis have no bearing on each other. If Jackson does reject the widely lauded practice of statistically testing a formal model, there are ways in which he could have debated this practice more explicitly: stating the implicit premise that underlies using large-n covariation regularities to test an ideal-type and explaining why this premise is unfounded, or providing an example of a claim that tries to meld analyticism and neopositivism and showing how to frame a rejection of the claim.

Jackson is especially strong when discussing the analyticist tradition, so much so that it has influenced how we are conducting our own research. Moreover, he uses it himself in setting up the two-by-two typology that structures the bulk of his book. But in doing so he fails to fully apply to himself the standards he ascribes to analyticism. According to Jackson, the "value of an ideal-type lies precisely in its being 'entirely' used as a means for the comparison and measurement of actuality" (2010: 144). Yet as just noted, his failure to analyze the manner in which traditions are mixed in *practice* means the reader is left unclear how one should assess much of the work that is actually *conducted* in the discipline. By title this is a book about the conduct of inquiry in international relations, but there is not enough attention to how research is actually conducted. More explicitly attending to the fact that IR as practiced is more mixed than Jackson's ideal-types, and explaining the trade-offs or implications of mixing philosophical ontologies, would have been of great use to graduate students as they nervously undertake dissertation research.

Our third and final criticism approaches Jackson in light of the concern with audience that he so convincingly explicates as central to a reflexivist philosophical ontology. Jackson employs the disassociated stance of scholarship with which most of us are familiar, and the result is a lucid and intellectually honest book. But in being so dispassionate has Jackson perhaps done himself and the discipline a disservice? Consider his critiques of critical realism and neopositivism. Jackson makes a cogent case against critical realism's treatment of real-but-unobservables. Although he notes that these unobservables are often proclaimed to be provisional in their use, he follows the logic of such a position to its practical conclusion and argues that knowing the true nature of such entities is likely a fruitless endeavor. However, in discussing neopositivists' instrumental use of concepts he is less philosophically cutting.

From a reflexivist position this is troubling because one must ask: who is the audience Jackson is trying to reach? Although there is certainly an IR "cottage industry" in philosophy of science (Grynaviski, this symposium), it does not have many writers in the neopositivist camp. The rigorous philosophical unraveling that Jackson pursues against critical realism may have the effect of alienating the very group most likely, due to their own engagements in the philosophy of science, to be interested in his book! This is a shame, as Jackson probably only seeks here to encourage critical realists to avoid the temptation of playing the science card themselves. Such a goal may have been better achieved by analyzing their position a bit more sympathetically while putting a little more edge into his account of neopositivism.

Conclusion

Jackson certainly shows that philosophy of science should be treated seriously, and this clarifying and inspiring book should be read by all students in the discipline. Yet it is not clear that it will be. As Jackson notes, there is an absence of sincere philosophy of science training in IR. Given the prominence of graduate students' concerns with making their work acceptable to key constituencies, Jackson displays a surprising reticence toward the practical feasibility of adopting his scheme for IR graduate students writing their dissertations. Success in a prospectus defense, conference presentation, or job talk relies upon the extent to which the audience will accept, or at least take seriously, the claims being made. It may, however, be quite difficult to stare down a dissertation committee or a job talk audience member and say, "Well, you are ignoring the philosophical-ontological contradictions implicit in your criticism."

Yet, unless held to such a standard—one, it should be noted, he does not set for himself—Jackson ultimately cannot be held responsible for how seriously the discipline will treat the very substantive issues this book raises. Although this book does not help us as graduate students to navigate the waters of the discipline *as a discipline* as much as we might hope, it has certainly helped us to steer our own thoughts more steadily. Being able to understand that many debates in IR already presuppose the same philosophical wager, and that others often mix and match from different underlying understandings of the hook-up between theory and the world, has helped us as scholars become more clear and confident about the standards which would establish if our own research counts as "science."

References

Clarke, Kevin A. and David M. Primo. 2007. "Modernizing Political Science: A Model-Based Approach." *Perspectives on Politics* 5:4 (December), 741–753.

Fearon. James. 1995. "Rationalist Explanations for War." *International Organization* 49:3 (Summer), 379–414.

Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 52:4 (October), 887–917.

Jackson, Patrick Thaddeus. 2010. *The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the Study of World Politics*. London: Routledge.

King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Rogowski, Ronald. 2004. "How Inference in the Social (But Not the Physical) Sciences Neglects Theoretical Anomaly." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Henry Brady and David Collier, eds. (Lanham, MD: Rowman and Littlefield), 75–83.

Snyder, Jack. 1991. *Myths of Empire: Domestic Politics and International Ambition*. Ithaca, NY: Cornell University Press.

Making Sense of the Study of International Relations: Seeking a Guide for the Perplexed

John G. Gunnell
State University of New York, Albany
jgg@albany.edu

Patrick Jackson's book on *The Conduct of Inquiry in International Relations* offers graduate students, younger scholars, and, indeed, many specialists a useful map for charting the often inhospitable terrain of scholarship in the field of International Relations (IR). This is particularly the case as far as illuminating the awakening to issues in the philosophy of science that has taken place in IR during the last two decades. Jackson presents a typology for sorting the debates about the nature and demands of scientific inquiry, which have often been conducted, either explicitly or implicitly, in terms of diverse and complicated philosophical arguments. As opposed to many previous analyses, which have tended to be couched in terms of dichotomies and biased toward a particular philosophical persuasion, Jackson's scheme is remarkably neutral, but, in some respects, maybe too neutral.

In exploring these matters, Jackson sometimes becomes entangled in the puzzles he seeks to resolve, such as the relationship between philosophy and social science, and, at certain points in his presentation, the historical context and genealogy of this relationship, as well as that between natural science and the philosophy of science, seems obscured.[1] My purpose is not to quarrel with his attempt to sort out what is going on in IR, but rather to add a little historical and critical gloss in the hope of joining in the kind of constructive contention that he so strongly advocates as essential to the conduct of inquiry.

Jackson's basic destination, after a long journey through a wide range of philosophers, reaching from Descartes to Roy Bhaskar, as well as a representative number of scholars in IR, is a condition of "healthy pluralism" in matters methodological. This general stance is certainly not novel. There is no doubt that pluralism has once again become the dominant ethic in political science, as well as democratic theory, and it now seems nearly as awkward to find fault with pluralism as it is to criticize the norm of eating a balanced meal. By securing objectivity while allowing variety, they seem both to solve the problem of relativism and to secure authority, but even some balanced meals are not easily digested.

The Genealogy of the "Science Card"

To put Jackson's work into perspective, it is helpful to consider briefly the intellectual genealogy from which his argument has emerged. Despite the fact that, from its earliest stages, and especially after the end of the nineteenth century, what Jackson refers to as "playing the science card" was the pivot of political science's rhetoric of inquiry, the image of science was both inchoate and largely unchallenged. There was little in the work of individuals such as Charles Merriam that resembled an articulate meta-science, and even among philosophers such as John Dewey, the concept of science had little distinct content beyond a purported affiliation with liberalism. This was, in part, because scientists did not tend to give an account of their own practices; science as a basic value was seldom questioned; and there was yet no disciplined field of the philosophy of science upon which to draw. During the creation of the American Political Science Association, the science card primarily functioned as a basis for claiming the kind of cognitive authority that would sustain the perennial hope for practical purchase in matters of liberal reform.

Although the reform concern would persist through mid-century in a latent manner, and sometimes explicitly in the work of individuals such as Harold Lasswell, the meaning of "science" increasingly for the most part surfaced in disputes internal to the discipline. By the early 1950s, the issue of what constituted science had become central to the behavioral movement and the controversy surrounding it. Although behavioralism is often viewed today as initiating a program for emulating the natural sciences, it was to a large extent a reaffirmation of the discipline's traditional commitments to the value of science and an image of pluralist liberal democracy, which together had been forged a generation earlier. As early as 1950, these commitments were, for the first time, confronted with a substantial philosophical challenge. What was particularly egregious was the fact that this challenge was mounted from within political theory which heretofore had been the source of the discipline's self-understanding of both science and liberalism and of the assumed integral relationship between them. An ideologically and philosophically diverse group of émigré scholars, ranging from Leo Strauss to members of the Frankfurt School, focused on science and liberalism as both causes and manifestations of a fundamental "crisis of the West" in which, they claimed, contemporary political science was deeply implicated.[2]

American political scientists had little in the way of indigenous resources on which to draw for a defense of science, but, serendipitously, another group of émigrés also arrived in the United States and brought with them a complex but systematic account of the logic and epistemology of scientific explanation, which became, in the American context, the basis of the philosophy of science as a distinct field of study. The logical positivists, and their intellectual evolution as the founders of logical empiricism, represented by figures such as Carl Hempel, formulated a meta-science which, in its secondary and tertiary renditions, was appropriated by social scientists. The formulations of logical empiricism were not intended as a guide to the practice of science, but its normative character was the residue of its historical origins and particularly of its ideological purpose of propagating a "scientific view of the world" in the European context. American social scientists were largely unaware of this patrimony, and although Hempel famously noted that if his account of scientific explanation did not conform to the practice of science, so much the worse for science, social scientists viewed it not only as an account and justification of science but as the basis of a technique for creating and deploying scientific theories and empirically verifying claims to knowledge.

While admirably attentive to the political context in which logical positivism took shape in Vienna, Jackson's account of mid-century shifts in the concept of science among scholars of international relations loses sight of such contexts. The notions of science deployed by Hans Morgenthau and E.H. Carr were, he notes, philosophically vague, but he overlooks their parallels, respectively, with Strauss and the Frankfurt School's critical interrogations of science, liberalism, and the purported affinity between them. Seen against this backdrop, the 1960s "great debate" to which Jackson traces the linking of "science" with quantification in IR might be interpreted as a defensive effort by liberals to reclaim ownership over the mantle of science by yoking it to new techniques they were mastering. It might, in turn, then appear as no accident that the democratic peace finding with which Jackson introduces his chapter on neo-positivism is not only the proudest empirical finding of the quantification agenda, but a cornerstone of contemporary liberal theory in IR.

While missing such more politically-charged resonances, Jackson is well attuned to nuanced philosophical contrasts between the view of science at play in IR's second "great debate" and the views deployed in more recent decades, especially the impact here of Kuhn and Lakatos. As early as the 1960s, Kuhn had begun to be enlisted by both critics and defenders of behavioralism, and subsequently mainstream political scientists increasingly became aware of the implications of his and Lakatos's arguments for their mantra of science. By the mid-1970s, debates about behavioralism had to a large extent become conducted in the surrogate language of opposing philosophies of science.[3] The critique of the positivist image of science was also complemented by a growing literature derived from the philosophy of social science advocating what was often referred to as an interpretive or hermeneutic approach. All of this discussion shared an assumption that the practice of science was grounded in philosophy.

Jackson correctly indicates that these debates did not significantly spillover into IR until after Martin Hollis and Steve Smith (1990) embraced what had become a common attempt at resolution among certain political theorists, that is, the idea that scientific and interpretive approaches were more complementary than mutually exclusive (Moon 1976). Hollis and Smith also voiced another concern that had gained a foothold in political theory. Although there were few critics of logical empiricism who did not find support in the work of Kuhn and although arguments such as that of Peter Winch provided new grounds for a neo-Weberian image of social inquiry as understanding meaningful action, this literature seemed to many to have relativistic implications and undercut the congenital idea of social science as a critical discourse (Bernstein 1976). One line of argument that evolved in political theory was that the attack on positivism signaled a need to reconstitute the image of science on a sounder philosophical basis, which would not only underwrite empirical inquiry but also provide a foundation for a critical social science. This was to be accomplished, it came to be claimed in the 1980s,[4] by turning to one of the principal successors to logical empiricism in the philosophy of science. This was scientific realism, which in IR, during the subsequent two decades, has become the most enthusiastically pursued meta-theory of science, even though, as Jackson makes clear, it is only one among a number of philosophical options.

Can Engaging the Philosophy of Science Help IR?

Jackson persuasively maintains that, in the end, all of these philosophical claims about science have not achieved a great deal and that science remains largely a rhetorical concept to which diverse content is attributed. After also exploring the philosophical literature devoted to the issues of how to demarcate science and how to determine what kinds of claims are authoritatively scientific, Jackson argues that this work, as well, does not provide much traction for getting on with the job of determining how a science of politics should be conceived and practiced. Jackson's first suggestion is that rather than seeking answers to these questions from philosophy, we should look within the tradition of social science itself and particularly at the work of Max Weber, which, among other things, indicated that science should be defined more by its goal than by a particular methodology or approach. As opposed to partisan politics, science should be devoted to the search for credible empirical knowledge, and this search should accommodate both (what came to be called in the 1970s) "explanation and understanding" or what today is often parsed as quantitative and qualitative research.

In some respects, Weber's argument seems an odd choice as a model. The work to which Jackson refers was a masterpiece of rhetoric in the service of reconciling academic ideological and philosophical polarities and demonstrating to political actors the epistemic authority of science.[5] One might reasonably argue that what prompted Weber's brand of ecumenicism regarding both the relationship between science and interpretation and the acceptance of many forms of ideal typification, ranging from marginal utility to the Protestant ethic, was precisely his conclusion that there was no authoritative resolution to these conflicts. And although Jackson ends his first chapter suggesting that we release ourselves from remaining intellectually mortgaged to the vagaries of the philosophy of science and, instead, seek social scientific autonomy, he opens the very next chapter arguing that philosophy of science can in fact aid us in clarifying and improving our research practices. Jackson leaves no philosopher's stone unturned, but some terminological problems emerge, in part because everyone from Descartes to the present is treated somewhat as if they belonged to the same club. In taking up this line of argument, he employs a somewhat odd terminology. The principal problem in this case revolves around Jackson's use of the term "methodology," and at this point, it is important to say something about the history of the philosophy of science.

"Methodology" and the Philosophy of Science

Although the philosophy of science is today often considered a branch of philosophy, it did not emerge as simply part of the evolution and differentiation of philosophical specialties. Departments and subfields in this area are often designated as some combination of the methodology and philosophy of science. Methodology typically refers to the study of the principles underlying a practice of knowledge such as science, and it is often considered a branch of logic. The origins of methodology, however, were largely within the practice of science itself, and the field of the philosophy of science evolved from those origins. Although today we may be inclined to think of someone such as Descartes as a philosopher, a strong case can be made for classifying him as a scientist who made "methodological" arguments to justify his substantive scientific claims at a time when "science" was neither paradigmatic nor socially authoritative. And the same could be said for a number of other people, including Newton and Darwin. Eventually, however, scientists, increasingly institutionalized, did not feel the need to defend themselves against rival authorities, and the theories of individuals such as Newton became internally dominant.

Methodological claims about such matters as induction tended to drop out of the discourse of science and to be taken up by individuals in philosophy or on the cusp of philosophy and science, who were in many cases concerned with raising the status of social science. John Stuart Mill was a typical methodologist who attempted to extend a logical account of scientific procedures to clarify and promote the practice of social inquiry. In the case of the Vienna Circle, most of the members were not easily defined as either scientists or logicians, but their principal arguments were about the underlying logic of science and about how to apply that logic to many areas of human practice. By the time that logical positivism had been re-established in the United States and lodged in departments of philosophy, the methodological emphasis, with its normative focus on the principles of scientific practice, began to give way to claims about the general transcendental structure of scientific explanation such as that represented in the "covering-law" model. Others, such as Popper, and later Imre Lakatos, never relinquished the normative methodological concern, which was at the center of what estranged Popper from Kuhn and, later, Feyerabend. When Feyerabend turned "against method," he might have more felicitously said "against methodology." "Method" typically referred more to particular techniques of research than to verification, falsification, induction, deduction, and the like, which was what Feyerabend rejected as specifying the essence of science.

Jackson argues that although we should not turn to the philosophy of science to find out what science is, it can nevertheless help us clarify our research practices, and he devotes much attention to the philosophy of science as a source for clarifying what he refers to as "methodological" questions. He claims that methodology is neither "method" nor "epistemology" and that it is more important than "ontology," in the sense that critical realists in IR use that term. Jackson equates methodology with what he labels "philosophical ontology," and he argues that the core methodological principle that we should adopt is to engage in "wagers" or "provisional commitments" about general background matters such as the nature of the "world" and the relationship between mind and world. Jackson, along the way, seeks to provide a philosophical lineage for such assumptions as well as an account of how they feature in contemporary philosophy. A basic problem with Jackson's discussion at this point is his use of and distinctions among terms such as "methodology," "epistemology," and "ontology." Although he distinguishes method from methodology, he oddly opposes methodology and epistemology which have historically been closely related. It was largely the traditional epistemological project that Feyerabend was arguing against. What the term "philosophical ontology" would typically evoke would be something such as metaphysical realism, which in some version is central to most accounts of scientific realism in IR. Jackson might have been better off if he had simply spoken of something such as "strategies of inquiry" rather than to define methodology in what seems to be a somewhat odd manner. In any event, he isolates a set of commitments that define such strategies and in terms of which he constructs a matrix of ideal typifications for specifying and comparing them and which are, in various respects, reflected in four general categories of approaches in IR: neopositivism, critical realism, analyticism, and reflexivity. He examines these approaches successively in chapters three to six.

One difficulty with the structure of the book is that this scheme is not fully elaborated until the concluding chapter, and it seems that it might have been more straightforward, and generated less potential misunderstanding, if Jackson had presented the typology more completely before proceeding to examine the similarities and differences among these philosophical premises and their implications for analyzing approaches to IR. In doing so, he might also have made clearer the extent to which the "methodological" assumptions manifest in these approaches simply resembled claims in philosophy, were the actual historical residue of such claims, or, as in the case of critical realism, were explicit attempts to draw directly upon the philosophical literature.

Conclusion: Some Payoffs and Problems

One of the most helpful sections in Jackson's book is his conclusion's discussion of the use of the term "constructivism" in IR. This usage has been very ambiguous, and even scholars who subscribe to the label or ascribe it to others often seem unclear about exactly what is entailed. Jackson recognizes that what is behind constructivism is, although often not well articulated, a theory, or what Jackson refers to as a scientific ontology regarding social facts. It is thus, in his terms, neither a method nor methodology even though it may entail both. I am, however, less sanguine about Jackson's account of particular aspects of the work of certain philosophers such as Kuhn;[6] his discussion of the relationship between IR and the philosophy of science; and his avowal of ecumenicism as the ethic of inquiry.

The model that Jackson proposes for thinking about disparate approaches in IR is derived from Abraham Heschel, who argued that although no one religion could demonstrate that it possessed special access to religious truth, they all, despite their differences, shared a number of basic concerns and beliefs. Jackson suggests that we extend this attitude to the study of IR and abandon the "holy war" among approaches based on different methodologies. Jackson argues that each approach believes that science should be systematic, open to public discussion, and devoted to the search for empirical knowledge and that this should be sufficient for finding common ground and constituting a scientific community. His message is that an examination of the philosophy of science reveals that accounts of science are irreducibly pluralist, so instead of chasing the meaning of "science," scholars, each in their own way, should get on with the business of producing knowledge about world politics. But although he insists that there is no general philosophical answer to the foundations of science, this does not mean that any approach can proceed without philosophical foundations, which are specific to a particular line of inquiry. And we must, as Weber claimed, be explicit about these basic commitments that inevitably inform inquiry.

The first problem with this formulation is that it tends to mix up two senses of "philosophy." To say that there is plurality in the philosophy of science would be like saying that there is plurality in the philosophy of religion and that because there is no answer to the "religion card," we should instead pay attention to what might be generically referred to as the philosophical (that is, theoretical) assumptions indigenous to particular religions. The second problem is that while Heschel was talking about the relationship among religions, the problem that Jackson confronts is not relationships among sciences but approaches within a science. He argues that in order to engage in an ecumenical dialogue, it is necessary to have a common vocabulary which allows "translation" and a "mediated" form of contestation.

Jackson notes that Kuhn's later work relinquishes his prior focus on paradigms in favor of speaking of a "lexicon" which defines a specific scientific community, and Jackson claims that his typology could be the basis of a vocabulary for discussion and mediation within the field of IR. The difficulty with this analogy, however, is that Kuhn's lexicons reflected the scientific theories or ontologies that bound a community together but which, despite some limited possibility for translation, ultimately not only distinguished one science from another and one historical form of a particular science from earlier and subsequent forms but rendered theories and their lexicons incommensurable. Kuhn never backed off from this claim of incommensurability. Jackson's vision of the state of contemporary IR is far from that of a scientific community bound together by common theoretical commitments. Jackson's stresses that his goal is not to urge a "synthesis" of approaches but rather "agonism" without "antagonism" or an "*engaged pluralist attitude*" in the context of "contentious conventions" which respect all approaches as valid but yet rejects "relativism." Nothing could be further from Kuhn's image of a scientific community.

It is significant that of all the approaches that Jackson examines, it is only in the literature of critical realism that there is much focus on the philosophy of science. While Jackson may detect in other approaches indirect connections and certain generic "philosophical" ideas such as "mind-world dualism," it is only in the case of realism that there is an explicit attempt to constitute inquiry in IR on the basis of a distinct literature in philosophy. What is characteristic of the work of Alexander Wendt, Colin Wight, Heikki Patomaki, and others of this persuasion is that not only are quite diverse varieties of philosophical realism enlisted as a foundation of inquiry, but they are usually combined with elements of other philosophies such as hermeneutics and constructivism. "Scientific" and "critical" realism are intellectual conglomerates of derivations from diverse philosophical positions, and the primary assumption is that philosophy, in the formal sense of the term, is the foundation of social scientific inquiry and normative judgment. Even someone such as Fred Chernoff who opposes realism in favor of a Duhemian theoretical conventionalism agrees with this basic premise. The crucial point here is that although Jackson begins by rejecting the philosophy of science as the key to defining science, he continues to assume in many respects that we must turn to the philosophy of science in order to both understand inquiry and learn how to conduct it. But the pluralism of philosophy with respect to these issues is as great as it is in the case of seeking to isolate the essence of science. This is not the place to engage this issue, but suffice it to say that (historically, conceptually, and logically) it is a dubious proposition and assumption.

When Feyerabend suggested that "anything goes" in science, he was not arguing that this had been characteristic of the history and current practice of science but only that no philosophical account of science could either capture an underlying secret to its evolution or provide a guide to its success. This was primarily directed against his former mentor Popper, as well as his friend Lakatos, and even though by this point he had largely come to agree with Kuhn, he still worried that there was a normative methodological message built into Kuhn's narrative. What Jackson's ecumenicism fails to note is that science is not grounded so much in what he calls "philosophical ontology"—general background assumptions about such matters as the relationship between mind and world—but in what he calls "scientific ontology"—substantive theoretical claims about what constitutes the "world." The poverty of social science resides less in a failure to be tolerant of diverse methodologies than in the liberal assumption that truth will emerge in the marketplace of ideas, just as Heschel assumed that the major religions were all in some basic way on the same page. It is doubtful Heschel's recommendation has much relevance to either the past history of religions or the basic character of most religious practices—except maybe for some aspects of contemporary Unitarianism. It certainly has little application to natural science. Heschel was confronting neither the kind of cognitive dissonance that Darwin experienced in choosing *between* science and religion nor, to use my favorite example, which Kuhn's work was too early to acknowledge, the mid-twentieth century conflict within geology between plate tectonics and the geosyncline theory with respect to issues such as the nature of mountains. What lies explicitly at the core of natural science, and what is more unreflectively manifest in social science, are theoretical claims and assumptions about what kinds of things exist and the manner of their behavior.

The "ontological" claims typically defended in critical realism are quite different from ontologies in natural science. The former involve, first, a philosophical faith in a transcendental but unrepresented reality which somehow provides the basis for adjudicating specific empirical propositions and, second, the assumption that positing categories of social entities such as states, structures, and agents is comparable to theoretical claims in natural science. The actual ontological issue in social science, however, is, on the contrary, the theoretical issue of the basic nature of social phenomena, and this issue, like all theoretical issues, does not allow ecumenicism. From the theories advanced by science flow the epistemologies and in turn the methodologies and the methods, and when the theories change these principles and techniques of practice are often transformed as well. And there is no general philosophical explanation for those transformations.

If there is any common insight into science that can be derived from the work of Kuhn, Popper, Feyerabend, and the like, it is that the dynamics of science depend on competition among theories which in the end is settled by persuasion, based on the application of various means, followed by a period of consensus before conflict arises once more. Where these philosophers differed was with respect to how that competition was, and should be, conducted and resolved. Despite what some at times have viewed as the hegemonic designs of approaches such as systems analysis, structural functionalism, and rational choice analysis, the dominant ethic in political science has typically been tolerance and the proliferation of models, conceptual frameworks, and strategies of inquiry, which in the field are often passed off as "theories." This ethic has presupposed the pragmatic liberal assumption that pluralism in ethics, politics, and science is necessary because of the methodological principal of "fallibilism" which is the secret of progress and dictates the logic of never "blocking the road to inquiry."[7] It seems that in the end Jackson's argument is very much a reflection of that spirit.

Notes

[1] For a fuller discussion of some of these issues, see Gunnell (1998).

[2] For a fuller discussion of this period, see Gunnell (1993).

[3] See, for example, Gregor (1971), Gunnell (1975), Ball (1976).

[4] For an example, see Isaac (1987); for a fuller discussion of the evolution of realism in social and political science, see Gunnell (1998: Ch 4).

[5] For a detailed discussion of Weber in this respect, see Gunnell (2007).

[6] See Gunnell (2009).

[7] For a classic statement, see Smith (1957).

References

Ball, Terence. 1976. "From Paradigms to Research Programs: Toward a Post-Kuhnian Political Science." *American Journal of Political Science* 20:1 (February), 151–177.

Bernstein, Richard. 1976. *The Restructuring of Social and Political Theory*. New York: Harcourt, Brace Jovanovich.
Gregor, A. James. 1971. *An Introduction to Metapolitics*. New York: Free Press.
Gunnell, John G. 1975. *Philosophy, Science, and Political Inquiry*. Morristown: General Learning Press.
Gunnell, John G. 1993. *The Descent of Political Theory: The Genealogy of an American Vocation*. Chicago: University of Chicago Press.
Gunnell, John G. 1998. *The Orders of Discourse: Philosophy, Social Science, and Politics*. Lanham, MD: Rowman and Littlefield.
Gunnell, John G. 2007. "The Paradoxes of Social Science: Weber, Winch, and Wittgenstein." In *Max Weber's "Objectivity" Revisited*. Laurence McFalls, ed. (Toronto: University of Toronto Press).
Gunnell, John G. 2009. "Ideology and the Philosophy of Science: An American Misunderstanding." *Journal of Political Ideologies* 14:3 (October), 317–337.
Hollis, Martin and Steve Smith. 1990. *Explaining and Understanding International Relations*. Oxford: Oxford University Press.
Isaac, Jeffrey C. 1987. "After Empiricism: The Realist Alternative." In *Idioms of Inquiry*. Terence Ball, ed. (Albany: State University of New York Press).
Jackson, Patrick Thaddeus. 2010. *The Conduct of Inquiry in International Relations: Philosophy of Science and its Implications for the Study of World Politics*. London: Routledge.
Moon, J. Donald. 1976. "The Logic of Political Inquiry: A Synthesis of Opposed Perspectives." In *Handbook of Political Science, Vol. 1*. Fred I. Greenstein and Nelson S. Polsby, eds. (Reading MA: Addison-Wesley).
Smith, James Ward. 1957. *Theme for Reason*. Princeton: Princeton University Press.

Pluralizing Social Science

Patrick Thaddeus Jackson
American University
ptjack@american.edu

The Conduct of Inquiry in International Relations (*C of I*) was not a book that I had any long-standing plans to write. The manuscript did, however, grow out of two related and long-standing frustrations that I had with discussions in Political Science in general and International Relations in particular about research design, causation, and the basic contours of knowledge-production. First of all, people seemed to invariably conflate questions of *method* or technique with questions of *methodology* or strategy of inquiry. Thus we had and continue to have rather problematic contrasts between "qualitative" and "quantitative" ways of doing social research as though the decision to use or not to use numbers had any determinate bearing whatsoever on the epistemic status of particular empirical claims. But whether or not one uses numbers is a question of technique, not a question of strategy, and as such *cannot* have any such profound impact; this means that in conducting these debates about how to do our work, we are working with impoverished and misleading terminology. Second, and related, people drew on extremely thin and partial conceptions of "science" as a way of warranting their positions; this was equally true of scholars contrasting "explaining" and "understanding" as ways of knowing, and of scholars reducing the entire panoply of the philosophy of science to the triumvirate Popper-Kuhn-Lakatos as though those were the only three people to have ever intervened in the debate about how science worked. When I taught my Ph.D. seminar on the production of valid empirical knowledge—entitled "The Conduct of Inquiry in International Relations"—I tried to allay both of these frustrations by equipping my students with a broader set of conceptual tools for thinking about these fundamental issues and articulating a defensible position with which they felt comfortable. This book derives from that seminar and from the frustrations that animated my pedagogy in that seminar.

In responding to the excellent critical engagements with my book provided by John Gunnell, Eric Grynaviski, and David Banks and Joseph O'Mahoney, I felt it appropriate to begin with this bit of context so as to clarify the book's aims and social location with respect to ongoing discussions. Because the book grew out of my frustrations with the narrowness of existing terminology and conceptual vocabulary, an important goal of the book is to broaden the discussion by casting a wider net and bringing in authors and notions that do not yet have as much currency in our field as they do elsewhere. Because the book grew out of a seminar in which I invited students to develop their own position on certain fundamental issues, an important goal of the book is not to take a strong stand for or against any particular articulation of how knowledge is to be produced scientifically. And because the book grew out of my extreme dissatisfaction with dichotomies like quantitative/qualitative and explaining/understanding, an important goal of the book is to replace those dichotomies with a more nuanced vocabulary that is still concise enough to be useful.

As such, *C of I* is neither directed against nor advocating for any particular kind of social-scientific methodology; it is instead inveighing against the narrow and biased ways that we have been talking about these issues in Political Science and International Relations over the past few decades. Narrow ways, in that the starting-point for many of our discussions seems to be a fairly unreflective commitment to a deductive-nomothetic hypothesis-testing model of "science," and accordingly the discussions descend all too quickly to the technical level of particular tools that can help to advance that unquestioned epistemic goal. Biased ways, in that the very terms that we use to frame and characterize the logic of social-scientific inquiry incline toward one way of proceeding—neopositivism—and generate an uphill battle for anyone wishing to advocate a different variety of social science. Chief among these biased terms, in fact, is the term "epistemology," since the traditional project of epistemology was almost entirely wrapped up with a particular way of conceptualizing the relationship between the mind and the world or between the knower and the known (Taylor 1995: 3–5, 14–17); that is why I am at such pains in the book to redirect the discussion towards methodology *broadly* understood, and away from a more or less exclusive focus on ways of increasing our confidence in general claims about cross-case covariation.

My interlocutors raise a variety of trenchant points, too many for me to exhaustively deal with here. But in general let me point out that the issues that they raise signal *precisely* the kinds of broad conversations that I hope that the book provokes and continues to provoke: conversations about what we are doing when we engage in social-scientific inquiry, conversations about what we ought to be doing, and conversations about how we can do it better. Conversations that do not take as their starting-point a specious notion of "*the* scientific method" or "*the* scientific way to study social life," but instead recognize that there are multiple ways of proceeding, multiple ways that are not reducible to one another. Conversations that do not start off with the common-practice fallacy—"this is how lots of people do things, therefore this is how we should do things"—but instead seek to provide positive warrants for diverse approaches to research design and the evaluation of substantive claims. If people read and react to the book in the way that my interlocutors have, then I will count the book a successful contribution to a richer discussion of these and related questions.

Of course, one can't focus on everything at once, so in the remainder of this response I'm going to engage three issues raised by my interlocutors: why I distinguish between the argument presented in *C of I* and an argument about the sociology of our scholarly field; why I think a reconstruction of diverse commitments in philosophical ontology and the methodologies to which they give rise at the present time is a useful exercise; and why I am opposed to efforts to combine methodologies.

Locating the Text

I am delighted to hear that Banks and O'Mahoney found my discussion of methodology in *C of I* helpful for their efforts to think through their own projects; that is the primary use that I hope individual scholars will make of the book. Their concluding reflection that "[a]lthough this book does not help us... to navigate the waters of the discipline *as a discipline* as much as we might hope, it has certainly helped us to steer our own thoughts more steadily" is one that I take not as a criticism, but as simple observation that one cannot do everything at once. Had I spent more time in the book in the crevices and crannies of contemporary scholarship, it might have been more difficult to achieve the kind of broad-brush depiction of different methodologies that forms the core of the book's argument. The kind of pressures that more or less compel scholars to emphasize their differences from one another (brilliantly discussed in Abbott 2001) makes it extremely difficult to get a clear view of the whole scholarly landscape and the implicit conceptual scaffolding undergirding it. Faced with a choice between writing a detailed account of what people are saying at the moment, and advancing a broader account of the basic categories with and within which they are operating, I chose the latter course.

Besides which, all too much contemporary discussion in the field about logics of scientific inquiry is sometimes so confused that the best way to "navigate" it is probably to steer clear of it as much as possible. For example, as Daniel Nexon and I have argued (2009), the use of terms like "paradigm" and "research programme" within International Relations and Political Science often bears little resemblance to the actual use of those terms by philosophers of science to assess scientific progress in fields like physics. For another example, the broad use of a phrase like "hypothesis-testing" to refer, as Banks and O'Mahoney observe, to "specifying what data might be relevant to one's research question before doing the research, or simply being clear about the claims that one is making" strikes me as obfuscation. We have less philosophically freighted terms for these operations; to my mind, "specifying relevant data" and "being clear about claims" seem like perfectly reasonable pieces of advice on their own, so I'm unsure what good attaching a phrase like "hypothesis-testing" to that advice would do.[1] Hypothesis-testing *does* carry philosophical baggage with it, and it should *not* be used simply as a generic term for making clear claims and using evidence to evaluate them, because there are *other ways* of using evidence to evaluate claims—ways that I endeavor to elucidate in the book. In such situations, the best way to deal with conventional use is, I think, just to walk away from it.

In some ways I have a very similar reaction to Gunnell's entirely accurate observation that I downplay sociology of science considerations about the political context of philosophical claims as my discussion moves into the heart of 20th century scholarship on world politics. While part of the reason that I do this is because of the fact that there is already some extremely good work—including his own—on these considerations, perhaps a larger part of the reason why I downplay sociological factors is because my target is not to explain the present shape of the field, but to intervene in the field and hopefully disclose some of the ongoing tensions and strategic misunderstandings within our conversations about methodology. While "how we got here," like "precisely what people are saying at the moment," can be a helpful way of getting at those bigger issues, I regard them (at least for the purpose of this book) as means to an end, to be pursued only as far as they help both author and reader to make sense of the conceptual issues tacitly in play.

It is for this reason that I privilege the philosophy of science (and not the sociology of science) in the book: if one wants to mount an internal critique of conversations in a field where the notion of "science" enjoys widespread currency, then taking "science" seriously seems like the basic ante for the game. As it happens, even a cursory examination of debates in the philosophy of science quickly reveals the poverty of the vocabulary current in Political Science and International Relations for discussing methodological questions. Sociology of science, which seeks to historicize and contextualize such vocabulary along with the concrete research practices to which it is linked, is not necessarily the best tool for improving our discussions. This is especially true insofar as "science" in a self-proclaimed scientific field functions as a *foundational* claim, meaning that it declares itself capable of grounding or warranting concrete research practices; while I am not persuaded that one need therefore be a foundational*ist* about science (see Jackson 2009 and Chernoff 2009 for an elaboration of this distinction), I do think that it is important to engage that foundation on its own terms if one wants to open some "thinking space" (as in George and Campbell 1990) within it.

In this way, I have endeavored to position *C of I* in about the same place that Max Weber was standing when formulating and delivering his famous lectures on science and politics (Weber 2004). Weber sought to work through some contemporary debates in a way that would give him a clearer view of the basic conceptual issues involved, particularly the distinctions and transactions between the areas of science, politics, and religious faith. It would, I think, be inaccurate to read Weber's empirical statements in those lectures as the point of the exercise; rather, the lectures were intended to serve a *hortatory* function, and to advise his listeners about the boundaries of realms of practice by developing a conceptual apparatus adequate to the task.[2] Similarly, if nowhere near as sweeping in its scope, *C of I* is intended to take a critical look at our contemporary discussions of how to do research and then to propose ways to improve those discussions. The book thus stands—deliberately—at the border of those discussions, taking no specific position within them so that it can address the shape of the discussion as a whole.

Reconstruction and Diversity

The non-standard terminology that Grynaviski and Gunnell note in the book is, as Banks and O'Mahoney rightly point out, an integral part of my argumentative strategy. Because certain words (like "epistemology" or "positivism") have acquired conventional meanings in our field that prevent them from being particularly useful tools for conducting a broad and pluralist debate, I chose to avoid them, and to formulate my central 2x2 map with the lesser-known axes of "mind-world dualism vs. mind-world monism" and "phenomenalism vs. transfactualism." I prefer these terms in part because they do not have the baggage associated with our conventional terminology—terminology which obscures and devalues both mind-world monism and transfactualism, and as such biases the debate toward the combination of mind-world dualism and phenomenalism familiar to us in "neopositivist" methodology. (Actually, it's probably more familiar to us as "positivism" or "the scientific method," but those are even less useful labels.) By introducing novel terminology, even at the cost of forcing the reader to work through it in the first couple of chapters, I seek to avoid simply saying what has already been said, and instead focus on organizing what has already been said (and what is currently being said) into a more useful set of categories.

Thus, the point of the exercise is to reconstruct diverse logics of inquiry in a way that allows us to think more systematically about what it might mean to do research in different ways. In practice, given the rather unreflective prevalence of neopositivism in the official pronouncements both of our leading research design textbook (King, Keohane, and Verba 1994) and its erstwhile critics (Brady and Collier 2004; George and Bennett 2005), this means thinking more systematically about what it might mean to do non-neopositivist research. My set of distinctions, which I call a "metamethodological lexicon" in the last chapter of the book,[3] is designed to do just this, both by foregrounding the commitments that tacitly support neopositivist research practices and by exploring the methodological entailments of other commitments. The reason that I call this work "metamethodological"—dealing with the philosophical foundations of methodology—and not, say, "epistemological," is because my target here is broader than traditional questions about how subjects acquire valid knowledge of objects, and is instead concerned with the more general question of how factual knowledge is produced. Gunnell quite rightly notes that methodology is usually thought to follow from epistemology, but I would say that precisely this "following from" supports the traditional epistemological project, and biases the whole conversation in favor of mind-world dualism and its concern with validly crossing the gap between the mind and the world. Admittedly, I could have generated a neologism here too, but the distinction between method and methodology already had some presence in the existing conversation (e.g., Sartori 1970: 1033; Waltz 1979: 12–13; Schwartz-Shea and Yanow 2002: 459–460), so I elected to retain it.

The distinctions that I draw in order to help us get a handle on the relevant methodological issues—a distinction between different ways of conceptualizing the mind-world interface, and a distinction between different ways of relating knowledge to experience—are ideal-typical distinctions in philosophical ontology. Both of those aspects of my typology require brief elaboration, in the light of some of my interlocutors' comments. Philosophical ontology, a notion I adapted from Mario Bunge via Heikki Patomäki and Colin Wight (2000), refers to that which is logically prior to particular substantive claims and theories—that is, issues pertaining to our "hook-up" to the world (Shotter 1993: 73–79). Methodology, understood as distinct from and prior to "method," operationalizes or enacts philosophical ontology, standing on a set of often-tacit philosophical commitments as it delineates concrete strategies of inquiry. I am thus suggesting, along with many other methodologists, that we should think about issues of research design and knowledge-production *separate from* any particular substantive account of the world and the things in it; to modify the critical realist catchphrase, we should indeed put ontology first, as long as it's philosophical ontology we're talking about. But I am simultaneously suggesting that we should think about these issues as *irresolvable* on purely logical or philosophical grounds; there is no definitive argument for or against any particular commitment of (or combination of commitments of) philosophical ontology, and so—like the skeptical humanists treasured by Stephen Toulmin (1992)—we have no defensible alternative but tolerance, both of alternate methodologies and of the commitments of philosophical ontology that underpin them.

It is important to keep in mind the ideal-typical character of the distinctions I am drawing. Because of this character, the categories that I flesh out in the book's central chapters are, *by definition*, artificially pure; they perfectly describe neither concrete actual authors nor concrete actual research strategies. But it is their abstract logical purity that constitutes their value as conceptual devices for helping us to think through the implications of our commitments, thereby clarifying our methodological stances. Each of the four central chapters in *C of I* takes one combination of philosophical-ontological commitments and discusses some of the methodological implications of those commitments, and I have endeavored to do so in such a way that the result is not an "unraveling" (*contra* Banks and O'Mahoney) of any position but rather a delineation of what that position logically implies. To respond to Banks and O'Mahoney's concern about my treatment of critical realism: you can't coherently be a critical realist unless you take pains to vet posited causal powers of objects either in a laboratory or via transcendental argument, because otherwise you'd just be positing things and regarding them as true without any evidence. But this is not a *problem* with critical realism, but is instead the basic *point* of a critical realist approach to science. Whether self-proclaimed critical realists in our field actually do either of these two things is another matter; my point is that their very methodology and the commitments of philosophical ontology on which it stands directs them to do so.

That said, I would never claim that any of the authors discussed in the central chapters of the book are somehow perfectly located within a particular box in my typology. Indeed, it would be surprising if they were, since the typology itself inhabits an idealized conceptual realm, and like all ideal-types would be quickly falsified in practice if it were treated as a description (see Weber 1999: 192–194). The point is not just that no actual work perfectly matches all of the standards logically entailed in a particular quadrant of the typology. Rather, the point is that actual authors and their works are a good deal more ambiguous than *any* abstract delineation of their main points. This is precisely why the purpose of closely reading a text is not to generate a definitive and incontrovertible summary of its argument, but instead to generate a defensible account of the work as a whole—an ineluctably interpretive process. Hence, locating an author with respect to her or his methodology and philosophical ontology can never be a simple matter of proving beyond a reasonable doubt that author X belongs in box Y; instead, the relevant questions are: Is this reading sustainable, especially for the text or the author as a whole? and does reading author X through category Y help to illuminate the point at issue, whether that point pertains to the text in particular or whether it pertains to the broader argument?

Applying those standards to Grynaviski's argument that I have misread both the American pragmatists and Kenneth Waltz in explicating the analyticist methodological stance, I must admit to being somewhat puzzled. There are certainly lines in Dewey and Pierce that can be read as consistent with the neopositivist procedures of hypothesis-testing and the quest for nomothetic generalizations, but *sustaining* that reading is more difficult in the face of both authors' pronounced reluctance to make either hypothesis-testing or nomothetic generalization the key warrant for valid knowledge. Indeed, Dewey argued (1920: 169) that the point of abstract systematization through scientific inquiry was to create analytically general claims that could serve as "tools of insight; their value is in promoting an individualized response to the individual situation." There is a world of difference between treating a claim about whether states pursue security as this kind of analytically general claim, and (as Grynaviski suggests) treating such a claim as an empirically general proposition: the former is a model that can be useful or not useful and can also be calibrated or updated, while the latter is a hypothetical conjecture that can only be falsified.[4]

Similarly, although one might read Waltz as a neopositivist interested in the testing of hypothetical generalizations—and I freely admit that this is the usual way that Waltz is read in the field—this reading can only be sustained if one downplays or ignores both Waltz's structural-functionalist roots (Goddard and Nexon 2005, 17–18) and his "theory of theory" (Wæver 2009: 206–208). Both of these aspects, by highlighting the importance of conceptualization and imagery, sharply differentiate Waltz's own efforts from those of scholars articulating general, falsifiable empirical propositions. Indeed, the pages of *Theory of International Politics* that Grynaviski cites (Waltz 1979: 124–125) do not, when read in context, unequivocally support a reading of Waltz as a neo-positivist, for at least two reasons:

> (1) Waltz speaks of "confirming" a theory and designing evaluations that, if passed, will help a theory begin to "command belief," but these are operations that a neopositivist can never consistently perform. For a neopositivist, knowledge is only ever an unfalsified conjecture, liable to falsification at any time; belief in confirmation, as Popper (1970; 1996) might have said, provides an obstacle to scientific progress by immunizing certain propositions from testing.[5] So Waltz, by deploying language of this sort, sounds less like a neopositivist and more like something quite different.[6]
>
> (2) Waltz suggests that a theory demonstrates its worth by helping us make sense of events "within a given area and over a number of years," which does not sound like the kind of covering-law explanation sought by neopositivists. Admittedly, to a neopositivist this might initially sound very much like "scope conditions" or some other kind of a call for middle-range theorizing, but note that Waltz never proposes that the *theory* be modified by introducing such empirical boundaries; rather, he suggests that the theory—itself analytically general, and unmodified—helps us make sense of a particular case (since a case is, of course, a unit of observation and not a concrete entity like a state).

Admittedly, Waltz's "penchant for ambiguity" (Goddard and Nexon 2005: 22), or at any rate his ambiguous use of "terminology about hypotheses…allowed easy assimilation" to prevalent neopositivist understandings (Wæver 2009: 211). Part of the point of my reconstruction of Waltz, like those of Wæver and Goddard and Nexon, is to clear up some of that ambiguity. By linking together the tantalizing hints in Waltz's seminal book—and some of his very disparaging comments about prediction and hypothesis-testing in subsequent publications (e.g. Waltz 1996; 1997)—a picture of Waltz as an analyticist emerges, despite Waltz's sometimes unclear use of methodological terms.

Beyond the Semblance of Pluralism

But given the ambiguities of textual interpretation and the alternative rules that might frame interpretative strategies (favoring charity versus suspicion, consistency versus contradictions, and so on), a debate about whether to read Waltz as an analyticist or a neopositivist soon hits diminishing intellectual returns. What is more important to recognize is that, even if Waltz is read as an analyticist, there is nothing to stop any neopositivist from reading Waltz, extracting a claim, converting it into a falsifiable proposition, and proceeding to test it. This is, in fact, precisely how neopositivism is supposed to work. But to then turn around and claim that the results of that test should have some bearing on what Waltz was doing in the first place, or to claim that the testability of that proposition would somehow "prove" that the original source was a neopositivist, is to overstep the boundaries of a pluralist approach to methodology and to at least implicitly legislate one methodology as exhausting the boundaries of "science" per se.

This is the flaw in the Empirical Implications of Theoretical Models approach to the use of formal models, in the effort to evaluate feminist claims about patriarchy by gathering evidence about gender discrimination, and in the attempt to read social mechanisms as intervening variables: the problem is not that advocates of one methodology take insight and inspiration from others, but that the advocates of one methodology (neopositivism, for the most part) claim exclusive rights to evaluate all empirical claims on their philosophical-ontological terms. Testing a hypothesis derived from a formal or an informal analytical model (for example) tells us precisely *nothing* about the worth of the model because the epistemic standards appropriate to an analytical model are distinct from those appropriate to falsifiable empirical generalizations. So my opposition to combining methodologies, simply stated, is that I think that it is impossible to combine methodologies, because every logically coherent piece of social science will end up having a dominant epistemic warrant for its claims even if it derives some of those claims from other sources. A neopositivist testing a hypothesis derived from a formal model is not engaging in "mixed methodology" or "multiple methodology" research; she or he is engaging in *neopositivist* research while testing that hypothesis.[7]

In addition to this logical and conceptual barrier to combining methodologies, there is also a practical reason to refrain from doing so: in a field marked by the dominance of neopositivism, a "mixed methodology" is likely to be neopositivist. Although virtually no one is *officially* against pluralism nowadays, many scholars are in effect against methodological pluralism in their research practices and in their engagement with other scholars. The clearest example of this that I know of is David Laitin's (2003) "tripartite methodology," in which formal models pass claims to the neopositivists who use both large-n and small-n hypothesis-testing to evaluate them.[8] Declarations of tolerance for practitioners of reflexivist fieldwork as long as they provide systematic data that can be used to code variables of interest (e.g., King, Keohane, and Verba 1994: 37–43) also provide the same kind of specious pluralism. And an exchange in the journal *International Theory* (2:1, 2010) between Andrew Moravcsik and Beate Jahn about the character of liberal theory was unproductive precisely to the extent that the fundamental methodological differences between the authors were not even acknowledged by Moravcsik; instead, calls for the systematic testing of empirical propositions stood in as a substitute for genuine methodological discussion.

All of this leads me to conclude that even though "pluralism" is the sort of thing that everyone claims to be for, it is in fact not the kind of thing that everyone actually practices. It is, as Colin Elman (2009) stressed in the last issue of this newsletter, "a hard choice" that calls for greater learning about and tolerance of other approaches, and greater care in our claims about the character of our field, than has usually been recognized.[9] A pluralism of mere method within a single methodological framework is not the same thing as a genuine methodological pluralism that would embrace and celebrate fundamentally different ways of producing knowledge. The starting point for any such pluralism, I think, has to be a richer vocabulary for discussing methodological issues, and a vocabulary that begins from the position of important and consequential differences obtaining between methodologies. *The Conduct of Inquiry in International Relations* is intended to contribute to the crafting and to the refinement of such a vocabulary, so that we can continue to do our work—in all of its varied forms—without constantly having to defend the legitimacy of what we are doing against dismissive critics operating with an overly narrow view of science. It is against such critics that the book is directed; it is for the rest of us that the book was written.

Notes

[1] Of course, a usage like this might be deployed strategically—even cynically—to help a dissertation prospectus or grant proposal pass muster with neopositivist referees. But any gains to the individual researcher here come at the collective expense of helping to prolong the fiction that something called "hypothesis-testing" is at the heart of all social science—comforting to neopositivists, perhaps, but not especially helpful to the rest of the field.

[2] Whether Weber's broader goal in his discussions of the politik/wissenschaft divide was to preserve the epistemic authority of science in the face of partisan considerations, as Gunnell maintains in his contribution, seems to me to be a slightly different issue. I think it is equally plausible to read Weber as seeking not to preserve science as a non-partisan force in politics, but as seeking to free science and particularly social science for fulfilling the very different social role of formalizing cultural values—but this is a subtle matter of "Weber studies" that we don't have to get into here.

[3] Gunnell is quite right that my use of the term "lexicon" diverges somewhat from Thomas Kuhn's, even though I borrow the term and some of the sensibility from Kuhn's later work (collected in Kuhn 2000). Kuhn remained focused on *substantive* vocabularies throughout his career, while my concern here is with *methodological* issues; however, the emphasis on the historicity and indexicality of key terms, plus the logical untranslatability of certain claims between lexicons, is common to us both.

[4] At the risk of turning this into a discussion about subtle points of Dewey interpretation, I should point out that the section of Dewey's *Logic* that Grynaviski cites is contained in a discussion about why the idea of falsification has to be replaced with "the institution of a contradictory negation" as part of a process of revising a general claim to account for seemingly discrepant evidence (Dewey 1938: 196–

198). The sequence that Grynaviski describes—general claim about state behavior, discrepant evidence provided by Mearsheimer and Walt, reformulated general claim that takes discrepant evidence into account—is a pragmatic procedure rather than a neopositivist one, precisely because neopositivist hypothesis-testing provides no logical way to link a falsified proposition with a successor proposition (except, perhaps, through Lakatosian retrospective reconstruction, and that raises a whole different set of concerns). Pragmatic analyticism, which never treated the general claim as a falsifiable proposition in the first place, has no such problem.

[5] Lakatosian language about "hard cores" is no help here, since Lakatos is very explicit that his philosophical procedure gives no advice to the practicing scientist about which propositions to believe (see, in particular, Lakatos 1970: 178–179).

[6] In fact, Waltz's language here sounds quite strikingly like the language characteristic of pre-Popperian, old-school Vienna Circle logical positivism—which inclined in a decidedly *monistic* direction when it came to the mind-world interface, and was accordingly very much at odds with Popperian notions of falsification (see the discussion in Jackson 2010: chap. 3). Systems theorists prior to Waltz—Talcott Parsons, Morton Kaplan, *et cetera*—had similarly monistic/logical positivist inclinations.

[7] If she or he was also responsible for building the model in the first place, then during model-construction and calibration she or he was engaged in analyticist research; the unit of analysis here is the *argument*, not the *person*. Grynaviski seems to be reading Waltz as doing precisely this: operating as an analyticist when developing his model of the international system, and then operating as a neopositivist when evaluating the model. As I've said, I disagree with this as an interpretation of Waltz, but even if I agreed with it, the fact would remain that what Grynaviski would read as the two different parts of Waltz's argument would be *logically* distinct, and we'd have two methodologies, not a single "mixed methodology."

[8] For my criticisms of Laitin's tripartite methodology, see Jackson (2006); for a variety of views see the symposium on Laitin in the Spring 2006, Vol. 4, No. 1 issue of this Newsletter.

[9] Although, when push comes to shove, I'm considerably more skeptical than Elman seems to be about what he calls "the limits of association and commensurability between several equally valid epistemes," since I don't think that there's any meaningful kind of "commensurability" to be had between different commitments of philosophical ontology and the methodologies that they entail; in this I agree with some of the authors skeptical of "multi-method" research in that issue of the newsletter, such as Ahmed and Sil (2009), and Chatterjee (2009). I also fear that searching for such "commensurability" would end up leading us back into methodological univocality in tacit support of neopositivism.

References

Abbott, Andrew. 2001. *Chaos of Disciplines*. Chicago: University of Chicago Press.

Ahmed, Amel, and Rudra Sil. 2009. "Is Multi-Method Research Really 'Better'?" *Qualitative and Multi-Method Research* 7:2, 2–6.

Brady, Henry E. and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham, MD: Rowman & Littlefield Publishers.

Chatterjee, Abhishek. 2009. "Ontology, Epistemology, and Multi-Methods." *Qualitative and Multi-Method Research* 7:2, 11–15.

Chernoff, Fred. 2009. "Defending Foundations for International Relations Theory." *International Theory* 1:3, 466–477.

Dewey, John. 1920. *Reconstruction In Philosophy*. New York: Kessinger Publishing, LLC.

Dewey, John. 1938. *Logic: The Theory of Inquiry*. New York: Henry Holt and Company, Inc.

Elman, Colin. 2009. "Letter from the Section President: Pluralism as a Hard Choice." *Qualitative and Multi-Method Research* 7:2, 1–2.

George, Alexander L. and Andrew Bennett, eds. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

George, Jim and David Campbell. 1990. "Patterns of Dissent and the Celebration of Difference: Critical Social Theory and International Relations." *International Studies Quarterly* 34:3, 269–293.

Goddard, Stacie E. and Daniel H. Nexon. 2005. "Paradigm Lost? Reassessing Theory of International Politics." *European Journal of International Relations* 11:1, 9–61.

Jackson, Patrick Thaddeus. 2006. "A Statistician Strikes Out: In Defense of Genuine Methodological Diversity." In *Making Political Science Matter: Debating Knowledge, Research, and Method*. Sanford Schram and Brian Caterino, eds. (New York: New York University Press), 86–97.

Jackson, Patrick Thaddeus. 2009. "A Faulty Solution to a False(ly Characterized) Problem: A Comment on Monteiro and Ruby." *International Theory* 1:3, 455–465.

Jackson, Patrick Thaddeus. 2010. *The Conduct of Inquiry in International Relations*. London: Routledge.

Jackson, Patrick Thaddeus and Daniel H. Nexon. 2009. "Paradigmatic Faults in International-Relations Theory." *International Studies Quarterly* 53:4, 907–930.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Kuhn, Thomas S. 2000. *The Road Since Structure: Philosophical Essays, 1970–1993*. James Conant and John Haugeland, eds. Chicago: University of Chicago Press.

Laitin, David. 2003. "The Perestroikan Challenge to Social Science." *Politics and Society* 31:1, 163–184.

Lakatos, Imre. 1970. "Replies to Critics." PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1970, 174–182.

Patomäki, Heikki and Colin Wight. 2000. "After Postpositivism? The Promises of Critical Realism." *International Studies Quarterly* 44:2, 213–237.

Popper, Karl. 1970. "Normal Science and its Dangers." In *Criticism and the Growth of Knowledge*. Imre Lakatos and Alan Musgrave, eds. (Cambridge: Cambridge University Press), 51–58.

Popper, Karl. 1996. *Myth of the Framework: In Defence of Science and Rationality*. London: Routledge.

Sartori, Giovanni. 1970. "Concept Misinformation in Comparative Politics." *American Political Science Review* 64:4, 1033–1053.

Schwartz-Shea, Peri and Dvora Yanow. 2002. "'Reading' 'Methods' 'Texts': How Research Methods Texts Construct Political Science." *Political Research Quarterly* 55:2, 457–486.

Shotter, John. 1993. *Cultural Politics of Everyday Life*. Toronto: University of Toronto Press.

Taylor, Charles. 1995. *Philosophical Arguments*. Cambridge: Harvard University Press.

Toulmin, Stephen. 1992. *Cosmopolis: The Hidden Agenda of Modernity*. Chicago: University of Chicago Press.

Waltz, Kenneth N. 1979. *Theory of International Politics*. New York: McGraw-Hill.

Waltz, Kenneth N. 1996. "International Politics is not Foreign Policy." *Security Studies* 6:1, 54–57.

Waltz, Kenneth N. 1997. "Evaluating Theories." *American Political Science Review* 91:4, 913–917.

Wæver, Ole. 2009. "Waltz's Theory of Theory." *International Relations* 23:2, 201–222.

Weber, Max. 1999. "Die 'Objektivität' Sozialwissenschaftlicher und Sozialpolitischer Erkenntnis." In *Gesammelte Aufsätze zur Wissenschaftslehre*. Elizabeth Flitner ed., Potsdam: Internet-Ausgabe, http://www.unipotsdam.de/u/paed/Flitner/Flitner/Weber.

Weber, Max. 2004. *The Vocation Lectures*. Indianapolis: Hackett Press.

Two Cultures: Hume's Two Definitions of Cause

Gary Goertz
University of Arizona
ggoertz@u.arizona.edu

James Mahoney
Northwestern University
james-mahoney@northwestern.edu

That and no other is to be called cause, at the presence of which the effect always follows, and at whose removal the effect disappears.

Galileo

A famous quote from David Hume provides a useful way to introduce two different approaches to causation in the social sciences:

We may define a cause to be *an object followed by another, and where all the objects, similar to the first, are followed by objects similar to the second*. [definition 1]...Or, in other words, *where, if the first object had not been, the second never would have existed*. [definition 2] (David Hume in *Enquiries Concerning Human Understanding, and Concerning the Principles of Morals* 1975 [1777])

As many philosophers have suggested, Hume's phrase "in other words" is misleading, if not completely incorrect.[1] The phrase makes it appear as if definition 1 and definition 2 are equivalent, when in fact they represent quite different approaches. Lewis writes that, "Hume's 'other words'—that if the cause had not been, the effect never had existed—are no mere restatement of his first definition. They propose something altogether different: a counterfactual analysis of causation" (Lewis 1986a: 160).

Following Lewis, we shall call Hume's definition 2 the "counterfactual definition." By contrast, we shall call definition 1 the "constant conjunction definition," to highlight Hume's idea that causes are always followed by their effects.[2] In this short essay, we consider how these two definitions have informed understandings of causation in the qualitative and quantitative research traditions in political science. Following our earlier work, we characterize these traditions as representing contrasting cultures marked by diverse beliefs, norms, and values (Mahoney and Goertz 2006).

It bears emphasizing that we are not arguing that our interpretations should be attributed to Hume himself. Hume's views on causation have been the source of enormous debate among philosophers, and we make no claim to resolving that debate. Rather, our purpose is to use Hume's definitions, which are widely reproduced in discussions of causation, as a device for discussing the different ways in which political scientists understand the concept of a cause.

The Quantitative Tradition

Before the rise of the Rubin approach (see Morgan and Winship [2007] for a good survey), statistical discussions of causation focused on Hume's constant conjunction definition (definition 1) within a probabilistic framework. For example, Suppes, in an early and prominent analysis, wrote that, "Roughly speaking, the modification of Hume's analysis I propose is to say that one event is the cause of another if the appearance of the first event is followed with a high probability by the appearance of the second" (Suppes 1970: 10).[3] Under this probabilistic approach, it seems natural to understand the constant conjunction definition in terms of correlation. Thus, definition 1 suggests that causation occurs when there is a strong, or at least statistically significant, correlation between X and Y. While all know the mantra "correlation is not causation," in practice statistically significant correlations are very central in identifying causal relationships.

One can also develop a statistical interpretation of Hume's counterfactual definition (definition 2). Doing this requires some work, however, because Hume's counterfactual definition implies a single case. Unlike definition 1, which states "all objects [plural] are followed...," definition 2 states "if the first object [singular] had...." To interpret definition 2 in a constant conjunction fashion, therefore, requires expanding Hume's idea to multiple cases.

The quantitative tradition accomplishes this move by interpreting both definitions 1 and 2 in terms of constant conjunction across many cases. A correlation of 1.00 means that there is a constant conjunction of $X = 1$, $Y = 1$ and $X = 0$, $Y = 0$. Definitions 1 and 2 can thus be fused together into one statistical interpretation. Definition 1 holds that when the cause is present, the outcome will be present (probabilistically). Definition 2 holds that when the cause is absent, the outcome will be absent (probabilistically). Since it makes no statistical sense to just look at cases of $X = 1$ without cases of $X = 0$ (or vice versa), the two definitions become joined as one. Neither definition can stand alone and make statistical sense. But when fused together, they offer a coherent symmetrical understanding of causation, one in which the emphasis is on what follows different values on the independent variable.

As of 2010, it seems safe to say that the dominant statistical view on causation in political science and sociology is the Rubin model.[4] Perhaps its most important innovation within statistical circles was the emphasis on the counterfactual basis of causation. For example, Morgan and Winship's excellent summary is called *Counterfactuals and Causal Inference*. Earlier statistical and probabilistic accounts are understood to have ignored or underappreciated this crucial aspect of causation.

The Rubin approach starts with the individual case and then builds a full-blown statistical model of causation. Using the basic experimental setup, an individual, i, is subject to a treatment. The counterfactual is then what *would have happened* if i had received the control. Since the individual cannot receive both the treatment and control at the same time, one of the two possibilities must always remain a counterfactual. This reality leads to a fundamental problem:

> *Fundamental Problem of Causal Inference*. It is impossible to *observe* the value of $Y_t(i)$ [t = treatment, c = control] and $Y_c(i)$ on the same unit and, therefore, it is impossible to *observe* the effect of t on i. (Holland 1986: 947)

Since using statistics to estimate or evaluate causal effects requires relatively large amounts of actual data, the best the statistician can do is estimate the *average causal effect*, or, to use the more popular terminology, the average treatment effect (ATE) in the sample or experiment.

> The important point is that the statistical solution replaces the impossible-to observe causal effect of t on a specific unit with the possible-to-estimate *average* causal effect of t over a population of units. (Holland 1986: 947)

The arrival at ATE as the basis for a counterfactual theory of causation completes what we call the "causal inference circle" (in analogy to the hermeneutic circle) in the quantitative culture:

(1) One starts with Hume's definition 2, which stresses the counterfactual for subject i.

(2) One interprets the definition using algebra and statistics: the counterfactual is the *difference* between treatment and control, $Y_t(i) - Y_c(i)$ (Holland 1986: 947).

(3) One applies definition 1 in its constant conjunction form for treatment and control separately, i.e., for $X = 1$ and $X = 0$.

(4) One calculates the average difference between treatment and control in all the cases, i.e., ATE.

(5) The ATE then provides the individual case counterfactual for subject i.

In this circle, Hume's constant conjunction definition 1 is doing the heavy lifting, even though the starting point is his counterfactual definition 2. The counterfactual starting point raises an impossible-to-resolve problem. As a result, the scholar must quickly turn to definition 1 and use notions of constant conjunction for both treatment and control to make any headway. The consequence is, however, that the Rubin model follows earlier statistical approaches in reducing the counterfactual definition 2 to the constant conjunction definition 1 in the actual practice of estimating causal effects.

The Qualitative Tradition

In the qualitative tradition, Hume's definitions are understood in terms of set theory and logic. Philosophers and social scientists who use set theory focus on the logical form of his constant conjunction definition: "If $X = 1$, then $Y = 1$." Reading the if-then clause as a statement of logic, definition 1 defines "cause" as a relationship of *sufficiency* between X and Y. This sufficiency interpretation calls attention to the $X = 1$ cases (i.e., cases where the cause is present). The researcher starts with the cause being present, and then looks to see if there is a corresponding effect. In this sense, the qualitative interpretation of definition 1 is similar to the quantitative one.

At this point, however, the two traditions part company. Qualitative researchers do not make the further inference that Hume's constant conjunction definition implies a correlation between $X = 0$ and $Y = 0$ cases. Rather, they assume that if the cause is *not* present, the outcome could be either present or absent. They treat definition 1 as a claim about sufficiency that can be investigated in its own right. Thus, unlike the statistical interpretation which requires definitions 1 and 2 to be fused, definition 1 when interpreted as a sufficient condition stands completely on its own, and can be valid independent of the conclusions reached by using definition 2.

Another key difference between the two traditions related to definition 1 concerns how they address the fact that effects rarely always follow individual, single causes (at least in the social sciences). As we saw above, quantitative approaches have long addressed this issue with probabilistic assumptions. Although these assumptions are also sometimes incorporated into qualitative research, another standard solution is to link causal sufficiency with "multiple, conjunctural" causation (Ragin 1987; see also Mackie 1980, and Baumgartner 2008, 2009 for recent analyses). In qualitative research, causal sufficiency is "conjunctural" in the sense that several *different* causes must combine together, e.g., $X_1 * X_2 * X_3$, to generate an outcome. Individual causes, e.g., X_1, are neither necessary nor sufficient for the outcome by themselves. When taken together as a *package*, they are jointly sufficient for the outcome. But in terms of the logic of definition 1, nothing fundamental has changed: one replaces X with a package of Xs such as $X_1 * X_2 * X_3$.

At the same time, qualitative researchers treat causal sufficiency as "multiple" in the sense that there are almost always *different combinations* of factors that are each sufficient for the same outcome. This is the general principle of equifinality: no one single package of causes generates all $Y = 1$ outcomes. One must try to discover the different packages that each lead to the same result. This is critical in the context of definition 2; the implication is that a value of zero on package 1 does not imply that $Y = 0$ because other causal packages might yield Y=1.

This logic-based approach generates its own chain of reasoning for starting with Hume's definition 1 and arriving at the individual case:

(1) One starts with Hume's definition 1, which stresses constant conjunction.

(2) One interprets this definition to mean that $X = 1$ is sufficient for $Y = 1$.

(3) One treats X as consisting of a package of causal factors, e.g., $X_1 * X_2 * X_3$.

(4) One establishes a generalization that all causal pack-

ages $X_1 * X_2 * X_3$ "are followed by" $Y = 1$.

(5) If case *i* has $X_1 * X_2 * X_3$, then this package is interpreted as the cause of $Y = 1$.[5]

We thus end up with a causal claim about case *i*, just as we did for the statistical causal inference circle. Despite this similarity, the process of reasoning is fundamentally different, and the nature of the causal claim about *i* at the end is different. The statistical approach uses both definitions 1 and 2, while this version of the qualitative approach uses just definition 1.[6]

The qualitative tradition also analyzes Hume's counterfactual definition 2 in terms of set theory and logic. Within philosophy, counterfactual aspects of causation have long received attention. Arguably the most influential account of causation within philosophy for decades was that of David Lewis. His book *Counterfactuals* was published in 1973, well in advance of the rediscovery of counterfactuals in statistics. Consistent with Hume's definition 2, Lewis develops his counterfactual definition in terms of the *individual* case:

> My analysis is meant to apply to causation in particular cases. It is not an analysis of causal generalizations. (Lewis 1986a: 161–162).
>
> Event *e* depends causally on the distinct event *c*, if *c* had not occurred, *e* would not have occurred. (Lewis 1986b: 242)

Other literatures outside of statistics also emphasize causation in individual cases, including Max Weber's famous analysis of counterfactuals (1949) and Honoré and Hart's (1985) analysis of causation in the law. When the focus is on individual events, the counterfactual account is the natural choice (even in the Rubin model). The key move in the qualitative tradition is to interpret definition 2 as a counterfactual in terms of logic: if $\neg X_i$ then $\neg Y_i$. This seems completely natural since Hume says "if the first object had not been." Much of the philosophical tradition that stresses counterfactuals remains at the single-case level, but general explanations are a central goal of much of social science, and hence it is of particular interest to social scientists to see how definition 2 can be reformulated in terms of causal regularities.

To arrive at a causal generalization, the qualitative culture reformulates the counterfactual in terms of a necessary condition. Generalizing about necessary condtions actually occurs in both qualitative and quantitative research. For example, Goertz (2003) provides 150 necessary conditions hypotheses by prominent scholars in political science, sociology, economics, and economic history. However, necessary condition hypotheses are highly controversial in the large-N statistical culture. In contrast, qualitative scholars find them quite natural, particularly in fields such as comparative historical analysis (Mahoney 1999). Hypotheses about necessary causes bring us back to Hume's constant conjunction definition 1 in the sense that the focus returns to many cases and general patterns.

The process through which qualitative researchers generalize counterfactuals suggests another causal inference circle beginning with definition 2 and returning to the individual case:

(1) One starts with Hume's definition 2, which stresses the counterfactual.

(2) One interprets this definition in terms of logic: if X had not occurred, then Y would not have occurred, i.e., if $\neg X_i$, then $\neg Y_i$.

(3) One generalizes the individual case counterfactual to all cases (within scope conditions), i.e., if $\neg X$ then $\neg Y$ for all *i*.[7]

(4) One converts this counterfactual into a generalization, using definition 1, about a necessary cause; that is, X is necessary for Y.

(5) If X is present in case *i* then X is a cause of Y.

In this circle, the key move is the conversion of the individual case counterfactual into a regularity statement about a necessary cause. In effect, the analyst stays with definition 2 throughout the circle, bringing in definition 1 to produce a generalization across cases. The retention of definition 2 is accomplished by assuming that the definition can be directly extended to many cases, and that, with this extension, testable hypotheses are generated.

Conclusion

Hume's famous quotation suggests two definitions of causation. Definition 1 suggests a constant conjunction between cause and effect, such that effects always follow causes. This definition assumes many cases. Definition 2 suggests a counterfactual view of causation, in which the absence of a cause leads to the absence of an outcome. This definition is built around a single case.

In quantitative research, which is normally interested in generalizing across many cases, it seems natural that scholars would gravitate more toward definition 1. Yet in recent years, as attention has turned to counterfactuals, definition 2 has become the starting point for defining causation. Nevertheless, the quantitative approach quickly sets aside the counterfactual notion of causation as applied to individual cases out of a conviction that it is impossible to estimate this kind of causation. For the statistical culture, there are really not two different definitions, because each one individually would make no sense. The statistical approach fuses the definitions into one in moving from the impossible to estimate definition 2 to the possible to estimate ATE.

In the qualitative tradition, the two definitions remain separate. Definition 1 is understood to represent a claim about causal sufficiency, whereas definition 2 is understood to represent a claim about necessary conditions. As a result, different sets of scholars may gravitate toward one definition rather than the other. Qualitative scholars who use methods, such as QCA/fs (e.g., Ragin 1987, 2000), for testing causal sufficiency may gravitate more naturally toward definition 1 and the sufficiency circle. By contrast, qualitative scholars who explore hypotheses about necessary causes more naturally embrace the counterfactual definition 2 (Goertz and Starr 2003). Nevertheless, these two causal inference circles easily coexist since they pose different, but related, questions.

Both the statistical and qualitative traditions use regularities to make causal claims. Indeed, their individual case counterfactuals are based on some idea of a causal regularity. But the two cultures look for different kinds of regularities. The statistical tradition looks for symmetrical regularities suggested by the idea of a correlation. When viewed in terms of Hume, this approach fuses his two definitions into a single claim about symmetrical causation. By contrast, the qualitative tradition looks for asymmetrical regularities suggested by the ideas of necessary and sufficient causation. It leads to an interpretation of Hume's two definitions as representing separate claims about causal sufficiency and causal necessity.

Is there a right interpretation of Hume's two definitions? Although we are not historians of philosophy, we think that one's view of the most useful interpretation of Hume will be strongly influenced by one's methodological background and approach. Our own view, consistent with the two cultures argument, is that each interpretation is useful and makes good sense within the overall approach within which it is embedded.

Notes

[1] We would like to thank Robert Adcock, Michael Baumgartner, and David Owen for comments on earlier drafts.

[2] This view of causation underpins the covering law model formalized in mid-twentieth century social science; for example: "A [covering, scientific, Hempel (1965)] law has the form 'If conditions $C_1, C_2, ..., C_n$ obtain, then always E'" (Elster 1999: 5).

[3] Obviously Suppes's account is more complex. Particularly when dealing with observational, as opposed to experimental, data one must be concerned with spurious relationships and the like. For the purposes of this essay one can interpret his view in terms of a randomized experiment.

[4] Sometimes this is called the Neymann-Rubin, or Neymann-Rubin-Holland model. Neymann is included because he published an early paper that foreshadows Rubin's (1990) ideas. Holland is sometimes added because he published a very influential account (1986) of Rubin's ideas. On substantive grounds one should discuss Pearl (2000), but his account has not had an impact equivalent to Rubin's within political science and sociology. For example, the extremely influential King, Keohane, and Verba (1994) based their account of causation on Holland.

[5] This final inference about *i* assumes that the generalization in step 4 is valid. This is analogous to our assumption above that there is a significant average treatment effect.

[6] In discussions of trivialness (Goertz 2006) or coverage (Ragin 2008) of a sufficient condition, the absence of the causal factor plays a key role.

[7] Of course, "if $\neg X$ then $\neg Y$" is equivalent to "if $Y = 1$ then $X = 1$". However, this formulation reverses the causal direction of the counterfactual version. The "$Y = 1$ then $X = 1$" equivalent nonetheless remains important for empirical testing, since often one uses this version for case selection, see, e.g., Dion (1998).

References

Baumgartner, Michael. 2008. "Regularity Theories Reassessed." *Philosophia* 36:3 (September), 327–354.

Baumgartner, Michael. 2009. "Inferring Causal Complexity." *Sociological Methods & Research* 38:1 (August), 71–101.

Dion, Douglas. 1998. "Evidence and Inference in the Comparative Case Study." *Comparative Politics* 30:2 (January), 127–145.

Elster, Jon. 1999. *Strong Feelings: Emotion, Addiction, and Human Behavior*. Cambridge: MIT Press.

Goertz, Gary. 2003. "The Substantive Importance of Necessary Condition Hypotheses." In *Necessary Conditions: Theory, Methodology, and Applications*. Gary Goertz and Harvey Starr, eds. (New York: Rowman & Littlefield), 65–94.

Goertz, Gary. 2006. "Assessing the Trivialness, Relevance, and Relative Importance of Necessary or Sufficient Conditions in Social Science." *Studies in Comparative International Development* 41:2 (June), 88–109.

Goertz, Gary and Harvey Starr, eds. 2003. *Necessary Conditions: Theory, Methodology, and Applications*. New York: Rowman & Littlefield.

Hempel, Carl G. 1965. *Aspects of Scientific Explanation*. New York: Free Press.

Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:396 (December), 945–960.

Honoré, Tony and H. L. A. Hart. 1985. *Causation in the Law*, 2nd ed. Oxford: Oxford University Press.

Hume, David. 1975 [1777]. *Enquiries Concerning Human Understanding, and Concerning the Principles of Morals*. Oxford: Oxford University Press.

King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.

Lewis, David. 1973. *Counterfactuals*. Cambridge: Harvard University Press.

Lewis, David. 1986a. "Causation. Postscripts to 'Causation.'" *Philosophical Papers*, Vol. II. Oxford: Oxford University Press.

Lewis, David. 1986b. "Causal Explanation." *Philosophical Papers*, Vol. II. Oxford: Oxford University Press.

Mackie, John L. 1980. *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press.

Mahoney, James. 1999. "Nominal, Ordinal, and Narrative Appraisal in Macrocausal Analysis." *American Journal of Sociology* 104:4 (January), 1154–1196.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14:3 (Summer), 227–249.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2008. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago: University of Chicago Press.

Rubin, Donald. 1990. "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5:4 (November), 472–480.

Suppes, Patrick. 1970. *A Probabilistic Theory of Causality.* Amsterdam: North-Holland.

Weber, Max. 1949. *Max Weber on the Methodology of the Social Sciences*. New York: Free Press.

Multiple Methods in Practice

Amy R. Poteete
Concordia University, Montreal
apoteete@alcor.concordia.ca

A few years ago, reflecting a growing appreciation for the diversity of methods in use and the value of combining multiple methods, the Qualitative Methods section decided to change its name to Qualitative and Multi-Method Research. Valuable though it may be, multi-method research also places new demands on scholars and is not easy to do well.[1] The spring 2007 symposium on "Multi-Method Work: Dispatches from the Front Lines" recognized some of the opportunities and challenges associated with this sort of research.

In this contribution, I am particularly concerned with combinations of methods that contribute to both internal validity and external validity, by making it possible to trace causal processes *and* evaluate their generality. My interest in external validity and the prevalence of causal relationships does *not* reflect a commitment to the covering law model of causal relationships. There are many reasons to expect contextually contingent relationships. Nonetheless, an expectation of causal heterogeneity does not imply a lack of concern with the prevalence of relationships or their distribution. Not all forms of multi-method research achieve a balance of internal validity and external validity, or of process-tracing and the evaluation of generality. We should be more attentive to variable contributions of different combinations of methods.

This essay makes three contributions. First, it notes some discrepancies between methodological goals and methodological practices. Second, it identifies practical challenges to engaging in multi-method research and discusses their implications for methodological practice. Third, it discusses possibilities for enabling multi-method research through meta-analysis and collaboration. Neither suggestion represents a cure-all. Meta-analysis can only evaluate relationships that are explored in existing studies; theoretical and conceptual innovations demand new primary research. Collaboration does not guarantee the use of multiple methods and raises a host of new challenges. Nonetheless, both strategies have a positive role to play. The profession would benefit from thinking more seriously about how to facilitate forms of multi-method research that contribute to both internal validity and external validity, and to process tracing and the evaluation of general patterns.

Methodological Principles and Practices

The complementarities of qualitative and quantitative research are well known. In multi-method research, the strengths of one method can compensate for the limitations of another. Furthermore, confidence in findings increases when multiple types of evidence and analytical techniques converge.

Some mixing of methods is fairly commonplace. In an analysis of articles published in *Comparative Political Studies*, *Comparative Politics*, and *World Politics* between 1989 and 2004, Munck and Snyder (2007a: 13) found that nearly a third of the articles (32.1%) combined qualitative and quantitative methods of empirical analysis. Because they code any empirical analysis that relies on numbers as quantitative (p. 27, fn. 9), an otherwise qualitative study that includes tables or charts depicting, for example, counts or percentage breakdowns would be categorized as using mixed methods. These forms of multi-method research mostly enhance the internal validity of analysis.

To balance process-tracing with the evaluation of general patterns, it would be more helpful to combine qualitative field-based studies, experiments, or agent-based models (ABMs) with more broadly comparative analyses, perhaps involving statistics or qualitative comparative analysis (QCA). This sort of multi-method research appears to be somewhat less common. A survey of ten top-ranked political science journals found that the share of articles that combined at least two different methods—case studies, statistics, or formal modeling—fluctuated between 15% and 19% during the period from 1975 to 2000 (Bennett, Barth, and Rutherford 2003: 377).

The use of statistical methods or a large number of observations does not guarantee the sort of broadly comparative analysis needed to evaluate the prevalence or distribution of relationships. In fact, the scope of cross-national analysis remains relatively limited in political science. Munck and Snyder (2007a) report that only a quarter of the articles in their sample examined data for more than five countries. The scarcity of cross-national analyses extends beyond comparative politics. Between 25% and 39% of the articles published in the top-ranked political science journals since 1975 involve case studies (Bennett, Barth, and Rutherford 2003: 374). Broadly comparative analysis based on qualitative, field-based research appears to be particularly difficult. Marco Janssen, Elinor Ostrom, and I conducted a meta-analysis of methodological practices in articles published between 1990 and 2004 that reported on field-based research concerned with collective action for natural resource management and were indexed in the Academic Search Premier database (Poteete, Janssen, and Ostrom 2010). Only 5.2% of these articles were cross-national in scope. Furthermore, we found an inverse relationship between the number of observations and the spatial scale of the analysis. Studies that analyzed more than 30 observations on the main unit of analysis were more likely to have a *sub*-national scope. This is because most of the large-N studies analyzed survey data. A survey of individuals or households in a particular location cannot speak to the prevalence or spatial distribution of patterns beyond that location. Thus, it does not enhance the external validity of the analysis.

The evidence suggests that, despite the enthusiasm for multi-method research and the growing interest in the identification of multiple causal paths, relatively little political science research involves the *sorts* of combinations of methods required to trace causal processes and evaluate their prevalence or spatial distribution. Why is this variety of multi-method research relatively rare?

Practical Challenges in Multi-Method Research

Multi-method research is constrained by a number of practical considerations. Practical challenges in multi-method research range from physical limitations on what any single individual can do to resource constraints and professional incentives that reward specialization. It is not simple to disentangle these factors.

What is required to conduct multi-method research that balances internal validity and external validity in the study of causal processes, and to balance process-tracing and an evaluation of the prevalence and distribution of hypothesized relationships and processes? Qualitative field-based research, experiments, and ABMs all lend themselves to the study of causal processes. Whereas quantitative research focuses primarily on the analysis of correlation, qualitative research is characterized by the sorts of close observations of interactions among various actors, institutions, and other features of a situation that are required for process-tracing.[2] But the social world is complicated, and it is often impossible to isolate the influence of particular factors. Triangulation of different types of evidence raises confidence in findings, but considerable uncertainty remains.

In experiments, comparisons of behavior in a control group with one or more treatment groups make it possible to isolate the influence of specific elements of a situation. Replication of experiments raises confidence in observed relationships between behavior and differences in experimental design. Poteete, Janssen, and Ostrom (2010, Ch. 10) describe how experiments helped make sense of sanctioning behavior observed by Elinor Ostrom and colleagues in Nepal. During a field visit to an irrigation canal, the researchers and the farmers who were guiding them discovered that somebody had dug a hole into the system, so that water was pouring downstream. The farmers jumped into action immediately. Some ran downhill to identify the culprit(s), while others repaired the canal. Ostrom was impressed. These rural farmers were willing and able to defend their shared resource (the irrigation system) through monitoring and sanctioning. But how widespread is such behavior? Back at Indiana University, Ostrom worked with Roy Gardner and James Walker to develop a formal model of sanctioning, which they then evaluated using laboratory experiments. Experimental evidence confirmed that, when individuals engaged in a common-pool resource (CPR) game are given the opportunity to sanction defection by others, sanctioning behavior occurs regularly (Ostrom, Walker, and Gardner 1992). These findings have been confirmed in many subsequent experiments by a variety of scholars (see Poteete Janssen, and Ostrom, Ch. 6). This example shows how the generality of field observations can be evaluated using experiments.[3]

Despite its advantages, experimental research has limited external validity, especially if the experimental design has no obvious parallel in ordinary social settings. Furthermore, it is often the case that observed differences in behavior across treatments can plausibly be attributed to more than one causal or behavioral process. With ABMs, causal and behavioral processes can be evaluated directly. ABMs involve the generation of simulated data based on algorithms that reflect alternative assumptions about human behaviors in particular types of situations. Alternative assumptions about alternative processes can be evaluated against data generated through multiple iterations of algorithms associated with each set of assumptions. Of course, ABMs have even less external validity than do experiments.

As discussed further below, some recent scholarship seeks to overcome the limited external validity of experiments and especially ABMs by combining these methods with field-based research. These efforts include field experiments and the involvement of people in field settings in the design and evaluation of ABMs. There is considerable scope for combining other forms of qualitative field-based research with experiments and ABMs. While qualitative field-based research has more external validity than either experiments or ABMs, qualitative studies cannot support an evaluation of the prevalence and distribution of observation relationships. Broadly comparative research involving a large number of cross-national observations is needed to achieve these goals.

How realistic is it to expect researchers to run ABMs and experiments to evaluate assumptions about human behavior under particular conditions, conduct qualitative field research to trace processes in actual social settings, and collect cross-national observations to support a cross-national analysis? The first practical challenge concerns skill development. After all, mastery of each method requires considerable initial training as well as ongoing effort to stay familiar with innovations. Good scholarship demands mastery of at least one set of methods. It also seems reasonable to expect scholars to gain enough familiarity with a variety of methods to support critical consumption of research using those methods. But how reasonable is it to expect a scholar to master two, much less three or more, distinct methods? For any single researcher, such an agenda would be overwhelming over the short term. Although a multi-method research program along these lines might seem more realistic if pursued over the course of an entire career, it would still be quite demanding.

Skill development is not the only obstacle to multi-method research, and probably not the most important. Resource constraints and professional incentives also influence methodological practices. Academia rewards evidence of research productivity, measured in terms of the number of publications and their perceived quality. Questions are raised when graduate students take longer than usual to finish their theses. Doctoral students who start to build a publication record before completing their degree are at an advantage. Evaluations for tenure and promotion focus on the pace of publications as well as their placement.

The emphasis on the pace of publication discourages more time-consuming forms of research. There are several reasons why multi-method research may be more time-consuming than single-method research, at least for an individual researcher. First, developing skills in multiple methods requires more time. Concerns about limiting the time to complete a degree mean that doctoral students are discouraged from taking the extra coursework needed to gain skills in multiple methods, even if relevant courses are offered at their universities. Once in a faculty position, a scholar faces tradeoffs between spending non-teaching time to produce publications using familiar methods or to gain expertise in a new method. Second, the execution of research involving multiple methods *may* take more time if each step of the project must be conducted sequentially.[4] Third, it may be more difficult—and thus take longer—to publish multi-method research. According to Lohmann (2007), most peer reviews are prepared by specialists who do not understand the rationale for multi-method (or interdisciplinary) research and tend to be suspicious of unfamiliar methods (or disciplines). Consequently, she argues, multi-method research is more difficult to get published. From a practical standpoint, the peer-review process may result in a slower pace of publication based on multi-method research, or placement of such research in less prestigious outlets. Similar dynamics can be expected whenever evaluations rely on peer reviews, such as the awarding of research grants and decisions about tenure and promotion.

The concerns raised above affect a wide variety of research topics. When data are either readily available or easily generated, decisions about methodological practices will focus primarily on how best to generate or access data (e.g., existing database versus experiment versus ABM versus collection of new field-based data), and how to analyze data (e.g., qualitatively, statistically, or with QCA). The challenges are even greater when data are scarce and difficult to access, as is the case with research on informal institutions and practices, non-elite actors, and historically marginalized people (see also Schedler forthcoming). Accessibility refers to research locations, the willingness of people to interact with the researcher, and the ability to observe the phenomenon of interest. Informal institutions and practices, by their very nature, cannot be recognized by reviewing official documents. Access requires trust, to be able to observe the informal and learn about the historically marginalized. Access also requires local knowledge to recognize the cultural coding that gives these observations meaning. When data are scarce and difficult to access, intensive field-based research is unavoidable and generally involves multiple methods. The use of multiple methods, such as interviews, household surveys, archival research, and participant observation, contribute to local knowledge, facilitate process-tracing, and create opportunities to develop trust. The analysis may involve a mix of qualitative and quantitative methods. However, since prepackaged cross-national data are not available for these topics and there are physical limits on the number of locations in which any single researcher can conduct this sort of intensive field research, cross-national analyses based on qualitative field-based research rarely compare more than a handful of cases. The absence of broadly comparative research is a problem because it is difficult to distinguish idiosyncratic conditions from more general patterns, evaluate the prevalence of more general patterns, or—when causal heterogeneity is an issue—identify the conditions under which specific processes operate.

Responding to the Challenges of Multi-Method Research

I have shown that, while multi-method research is fairly common, not all combinations of methods balance internal validity and external validity, or process-tracing and an evaluation of the prevalence of relationships. To address these goals, methods with strengths in process-tracing and the isolation of causal relationships—qualitative field-based research, experiments, and ABMs—should be combined with broadly comparative field-based research. I have identified a variety of practical considerations that make it very difficult for any single researcher to succeed in this sort of multi-method research, particularly within a short time horizon. I now consider two strategies for addressing these challenges: (1) greater use of meta-analysis to synthesize research findings and (2) collaborative research. Neither strategy can replace the need for the development of skills in particular methods or local knowledge about particular research locations. Instead, these strategies seek to accommodate specialization by individual scholars without giving up on the sorts of multi-method and broadly comparative research that is beyond the reach of many individual scholars.

Evaluating the Prevalence of Patterns through Meta-Analysis

Meta-analysis is a method for synthesizing research that involves systematic coding of data and characteristics derived from existing studies. Meta-analysis makes it possible to synthesize findings from a large number of studies and can be particularly helpful as a technique for making sense of seemingly contradictory findings. Since its recognition as a distinctive strategy in the mid-1970s (Glass 1976), the use of meta-analysis has become widespread in the clinical sciences (e.g., medicine, psychology) and education. There is such a volume of studies, and so many slight differences in the designs of clinical trials and experiments, that scholars and medical professionals have grown wary of relying too heavily on the findings of any single study. In these disciplines, meta-analysis typically involves quantitative analyses of quantitative data, such as the effect sizes of specific treatments. Because of its widespread adoption in the clinical sciences, meta-analysis has become closely identified with the specific quantitative techniques used to integrate findings from clinical trials and experiments. Meta-analysis, however, can be used to synthesize other types of research as well.

Within political science, meta-analysis may be particularly valuable for synthesizing findings from qualitative research. Meta-analysis of qualitative case studies would make it possible to retain the benefits of intensive field research, including the focus on process-tracing and heightened internal validity, while gaining some insight into the prevalence and distribution of observed patterns. A meta-analysis of case studies requires the collection, screening, and coding of a body of case studies related to a phenomenon of interest. It does not require the use of any particular method of analysis. While a scholar interested in general patterns might conduct statistical analyses, another scholar might use QCA to check for the existence of multiple patterns. Simple though it sounds in the abstract, meta-analysis of case studies presents important challenges at every step. How does one identify, collect, and screen relevant studies? What criteria should be used in coding?

Meta-analysis begins with the identification of a body of qualitative case studies concerned with the phenomenon to be studied. While it is tempting to rely on searches of bibliographic databases, one must be aware of possible sources of bias. For example, we might be interested in comparing the influence of social and economic heterogeneity on the prospects for collective action. Studies on this topic have appeared in diverse forms: academic articles, books and book chapters, working papers, reports produced by NGOs and donor agencies, conference papers, and as draft papers in file cabinets and computer hard drives. There are reasons to suspect that the form of research might correspond with the nature of the findings reported. Publication bias—a systematic relationship between characteristics of the findings reported in a paper and its publication—is of particular concern. For instance, findings that are deemed significant are easier to publish, and easier to place in highly ranked journals. The perceived "significance" of a study may depend on its inconsistency with the conventional wisdom, so seemingly anomalous findings are seen as more significant than studies that confirm widely held expectations. Consequently, confirmatory studies may be underrepresented in peer-reviewed journals and overrepresented in the gray literature, filing cabinets, and personal hard drives. If so, a meta-analysis based only on publications indexed in bibliographic databases will yield misleading results. Of course, the body of unpublished research is immense and difficult to access. The emergence and growing population of online archives for working papers reduces but does not eliminate this problem. At a minimum, then, it is important to evaluate and report likely differences between the sources included in the meta-analysis and those that were inaccessible.

Scholars should take measures to avoid exacerbating possible biases during the screening process. Studies should *not* be excluded from the meta-analysis solely on the basis of limited novelty or sophistication in the analysis. Rather, screening should focus on the inclusion and reliability of the information reported for key variables.

In qualitative field research or when working with experiments or ABMs, a scholar chooses a particular set of focal variables, decides how best to measure those variables, and organizes data collection and generation activities accordingly. In meta-analysis, a scholar must work with studies that consider somewhat different sets of variables and might differ in the conceptualization or measurement of common variables.[5] Some source materials will focus on social heterogeneity but not economic heterogeneity, while others report on economic but not social heterogeneity. These differences across studies limit the ability to synthesize findings through meta-analysis. When source material does not include enough information to support coding of a variable, that variable must be coded as missing and the study will fall out of the meta-analysis. As a consequence of missing data, our eventual meta-analysis would encompass only a fraction of the body of research concerned with, for example, the influence of heterogeneity on collective action.

A meta-analyst must decide whether to adopt broadly or narrowly defined criteria for coding variables. Staying with the same example, social and economic heterogeneity have been conceptualized and measured in a variety of ways. Social heterogeneity, for example, might refer to religious or ethnic differences, or to differences in attitudes or values, and economic heterogeneity may refer to differences in sources of livelihood or income or wealth. Broadly defined criteria for coding source materials would group together diverse indicators of each concept. While this strategy reduces the number of missing variables, it raises questions about the validity of the resulting measures. Does it really make sense to code cultural heterogeneity in a manner that groups together religious, ethnic, attitudinal, and value differences? Do differences in wealth and sources of livelihood really influence collective action in the same way? Narrow criteria distinguish between these different conceptualizations and measurements by coding them as separate variables. This makes it possible to evaluate whether there are differences in the influence of religious and attitudinal heterogeneity, or in the role played by differences in wealth and sources of livelihood. Of course, the ability to hone in on more specific relationships comes at the cost of increased missing data. A third strategy tries to strike a balance through an iterative process of coding and analysis. Initial rounds of analysis use broadly defined coding criteria, while subsequent rounds of analysis probe the sensitivity of the initial results to more precise specifications.

There are relatively few examples of meta-analysis in political science or the social sciences more generally. Scholars associated with the Workshop on Political Theory and Policy Analysis at Indiana University developed meta-databases for the Common-Pool Resource (CPR) and the Nepal Institutions and Irrigation Systems research programs. Both databases were designed to identify conditions under which people overcome collective action problems in the management of common-pool resources. These meta-databases have supported analysis of a wide variety of topics related to collective action, including the relative importance of different types of property rights (Schlager 1994; Tang 1994) and resource characteristics (Schlager, Blomquist, and Tang 1994), and the different roles played by social, economic, and locational heterogeneity (Ruttan 2006, 2008; Lam 1996, 1998). Ostrom (1990) involved a re-analysis of a sub-set of the cases from the CPR meta-database, using a modified coding instrument. The value of meta-analysis extends well beyond the study of collective action to topics as diverse as the relationship between democracy and economic growth (Doucouliagos and Ulubasoglu 2008), the influence of foreign aid on economic growth (Doucouliagos and Paldam 2008), decentralization and centralization (Pagdee, Kim, and Daugherty 2006; Nugent and Sanchez 1999), forest destruction and regeneration (Rudel 2005), policy coordination in metropolitan areas (Sager 2006), the effects of negative campaigning (Lau, Sigelman, and Rovner 2007), and voter turnout (Geys 2006).

Political scientists should use meta-analysis more widely because it enables the synthesis of large bodies of research. For topics for which data are scarce and difficult to access, meta-analysis may be the most realistic technique for evaluating the prevalence and distribution of relationships found in qualitative studies. It can also facilitate the synthesis of the increasingly voluminous bodies of statistical or experimental research.

On the other hand, meta-analysis *cannot* eliminate the need for primary analysis, nor should it be seen as somehow superior to primary analysis. Precisely because meta-analysis is constrained by the variables and measures found in the existing literature, it has limited potential for theoretical innovation or conceptual refinement. Meta-analysis might suggest revisions of hypotheses or measures but, since these innovations will by definition be absent from most existing studies, they must be evaluated through new primary research. Thus, meta-analysis is complementary to primary analysis. Ideally, there should be an alternation between the two forms of analysis.

Collaborative Research: Opportunities and Challenges

Meta-analysis does not eliminate the need for primary analysis, nor does it eliminate the value of using multiple methods in primary research. Thus, we come back to the practical limits on what any single researcher can be expected to do, particularly within the framework of a short-term research project. Collaboration seems to offer the perfect solution. Individual scholars can respond to the incentives for specialization while working with others in research teams to produce multi-method research. Collaborators with different methodological skills can work together to expand the range of methodologies used in any given study. Collaboration among scholars working in different geographical locations makes possible a more broadly comparative analysis of observations drawn from qualitative field-based research.

Collaborative research has proven its value in a wide variety of research areas within political science. Examples range from the Research Network on Gender Politics and the State (RNGS) to the Constituency-Level Election Archive (CLEA) and the Afrobarometer.[6] Given the advantages that collaboration seems to offer, however, it is not as common as might be expected. Evidence of collaboration can be seen in co-authorship and in references to larger projects. In our analysis of methodological practices in articles concerned with collective action related to shared natural resources, Janssen, Ostrom, and I found that both forms of collaboration have increased (Poteete, Janssen, and Ostrom 2010). The portion of articles with more than two authors climbed from 5.4% in the 1990s (1990–1999) to 20.7% of the articles published in the subsequent five years (2000–2004). Over the same period, the share of articles that referred to a larger research project increased from 16.1% to 35.3%. Others have found similar patterns. Only 25.7% of the articles published in *Comparative Politics*, *Comparative Political Studies*, and *World Politics* between 1989 and 2004 had more than one author (Munck and Snyder 2007b). Most multi-authored articles in these journals had two authors and none had more than three. By comparison, co-authored articles account for approximately half of the articles published in the *American Political Science Review*, *American Journal of Political Science*, and *Journal of Politics* (Munck and Synder 2007b: 339). Sigelman (2009) suggests that there is no publication bias (positive or negative) associated with co-authorship. Out of a sample of 500 manuscripts submitted to the *American Political Science Review* for review between 2001 and 2007, 54.9% of the manuscripts had one author and 30.6% had two authors; there was no statistically significant difference in the acceptance rates for single-authored and multi-authored manuscripts.

I argued above that not all forms of multi-method research contribute to a balancing of external validity and internal validity, or of process-tracing and evaluation of the prevalence and distribution of relationships. The same holds true for collaboration.[7] In our review of methodological practices, Janssen, Ostrom, and I found that collaboration was associated with *more intensive* field research (Poteete, Janssen, and Ostrom 2010). A research team might bring together, for example, social scientists and natural scientists, thereby expanding the range of methods and types of data collected in study sites. Most of the collaborative projects we identified brought together scholars with a shared interest in a particular research site; only a few supported broadly comparative research. These patterns are consistent with Munck and Snyder's (2007a) observation that, at least in the comparative politics journals that they examined, collaboration between political scientists and non-political scientists is most common in area studies research.

What influences the prevalence and forms of collaborative research? To draw out the advantages and challenges of collaborative research, I briefly describe two examples with which my co-authors and I have been directly involved.

First, consider the International Forestry Resources and Institutions (IFRI) research program, which involves an international network of collaborative research centers in the study of interactions between patterns of forest resource use, institutional arrangements, and forest conditions.[8] Members of the IFRI network conduct field research in forest sites using a common set of data-collection protocols and contribute data to a common database. Repeat studies are conducted every five years or so to assess changes over time. The data-collection protocols encompass socioeconomic data related to social structure and patterns of resource use, institutional arrangements related to forest management, and biophysical indicators of forest condition. In this manner, IFRI is building a cross-national database based on intensive, multi-method field research. It took time for IFRI scholars to reap the benefits of organization as a research network. Although there are now twelve collaborating research centers, the network was established in 1992 with only four centers, in Bolivia, Nepal, Uganda, and the United States. Even with a network of scholars, intensive field research takes time, and it took time for the database to grow. Not only did IFRI scholars need to familiarize themselves with other countries in the network, they also had to figure out strategies for comparing forests across ecological zones. The first cross-national analyses of IFRI data began to appear about ten years after the network was established. Now, however, IFRI scholars are producing cross-national analyses at a steady pace. These studies make it possible to discern general patterns, as well as regional differences, in the relationship between socioeconomic conditions, institutional arrangements, and forest conditions. If IFRI did not exist, these sorts of broadly comparative studies simply would not be possible.

The second example concerns a research project initiated by Marco Janssen (Arizona State University), who sought to ground formal models of collective action more firmly in empirical evidence of how groups create and adapt institutional arrangements. He decided to pursue this goal by combining ABMs, laboratory and field experiments, and role games in field settings. While the ABMs simulate human behavior arising from particular assumptions about human behavior and the structure of the choice situation, they lack external validity. Experiments and field-based research would make it possible to evaluate the fit between the formal models and actually observed behaviors. The project brought together an applied mathematician (Janssen), an experimental economist (Juan-Camilo Cardenas), an ecologist (François Bousquet), a cognitive scientist (Robert Goldstone), a computer scientist (Filippo Menczer), and a political scientist (Elinor Ostrom). The project builds on Cardenas' past work comparing laboratory experiments with field experiments involving Colombian villagers and Bousquet's development of role games involving Thai villagers as a mechanism for empirically grounded ABMs. Precisely because team members came from different disciplinary and methodological backgrounds, they had to first familiarize themselves with one another's approaches before they could identify the most fertile strategies for integrating those approaches. The project's innovations in strategies for combining these methodologies in field settings, and the associated insights into the dynamics of collective action, would not have been possible in the absence of this preparatory work. In light of our earlier discussion of professional incentives that prioritize rapid production of results, the longer lead time involved in such multi-method research does represent a challenge. On the other hand, an individual scholar seeking to bring together diverse methods in this manner would face an even more daunting preparatory period.

Both examples demonstrate the real value to be gained from collaborative research in terms of balancing internal validity and external validity, and in gaining comparative breadth without sacrificing qualitative depth. At the same time, these examples also give some indication of the very real challenges associated with collaborative research. The fruits of collaboration may be most pronounced when partners bring complementary skills and expertise to a project, and take the time to gain a meaningful level of familiarity with one another's areas of expertise. This means, however, that the benefits of collaboration do not flow immediately or easily. To understand one another and work productively together, partners must overcome differences in assumptions and vocabulary. Misunderstandings will arise and must be worked through. This process requires considerable effort and patience. As discussed above, however, academia rewards scholarly productivity measured in terms of the pace of publications and their placement. Furthermore, because it is difficult to discern individual contributions to co-authored publications, they may be discounted in performance evaluations. Aware of these considerations, scholars fear that they may not be given enough credit for publications with more than two or three authors to make the work worthwhile.

The longer start-up time for collaborative initiatives also presents challenges related to funding. Many agencies offer three-year grants and expect to see results—publications—within that time period. As we have seen, however, collaborative research that crosses methodological or disciplinary boundaries may only begin to bear fruit toward the end of this period; it may take five to ten years to realize the potential of these sorts of collaborative efforts. The emphasis on short-term productivity has the perverse effect of discouraging scholars from gaining familiarity with the methods and disciplinary approaches of their colleagues. Under these conditions, scholars face a strong temptation to exploit the resources associated with a formally collaborative project to support their personal research agendas instead of working together to achieve methodological synergies. These dynamics could account for at least part of the discrepancy between references to larger collaborative projects and the prevalence of co-authorship in the articles reviewed by Poteete, Janssen, and Ostrom (2010).

Concluding Thoughts

I have argued that not all forms of multi-method research contribute to a balancing of internal validity and external validity, or of process-tracing and evaluations of the prevalence and distribution of relationships. I have shown that combinations of methods that could contribute to these goals are relatively rare, and argued that the relative rarity of such research reflects practical constraints on methodological practices.

It is sometimes suggested that external validity has limited relevance in light of causal heterogeneity. I disagree. In the presence of causal heterogeneity, it would be helpful to understand the relative prevalence of distinctive causal patterns and the conditions under which each occurs.

I have considered two strategies for overcoming these challenges: meta-analysis of case studies and collaborative research. I am not recommending that other forms of research should be abandoned in favor of meta-analysis or collaborative research. In fact, neither approach can stand on its own. Meta-analysis can glean patterns from a large body of existing studies that might suggest theoretical revisions or new ways of conceptualizing and measuring key variables. New primary research is required, however, to evaluate new ideas. Likewise, the power of collaborative research comes from bringing together scholars with different methodological skills, area expertise, and/or disciplinary perspectives, each of whom has a deep grounding in his or her own approach. Both meta-analysis and collaborative research present their own distinctive sets of challenges, several of which we have discussed.

None of these approaches is superior to the other; they are instead complementary.

Poteete, Janssen, and Ostrom (2010) underline the influence of professional incentives on methodological practices, both at the individual level and with reference to collaborative research. The profession rewards speed and evidence of individual productivity. While these criteria make sense to some extent, they become a problem when they influence methodological practices in ways that interfere with the accumulation of knowledge. In particular, we are concerned that the emphasis on the pace of research productivity measured in terms of publications discourages investments in learning new methods, while the emphasis on individual work discourages multi-method collaboration. Professional incentives are not easily changed, but neither are they immutable. I encourage efforts to develop mechanisms by which investments in multi-method and collaborative research could be rewarded, without interfering with the development of a solid foundation in particular methods. Such mechanisms might include an expansion of postdoctoral fellowships linked to multi-method research teams, recognition in performance evaluations of investments in new methodological skills, and extensions of the timetables granting agencies use in evaluating the productivity of multi-method collaborative research projects. There are undoubtedly many other possibilities.

Notes

[1] Thanks to Lin Ostrom and Marco Janssen for helpful suggestions, and to Joanna Broderick for her careful editing.

[2] Some qualitative studies, such as those based on Millian methods, also analyze correlations.

[3] Of course, while experiments are useful for phenomena related to individual and group behavior, they are either inappropriate or not feasible for many other research questions.

[4] A scholar could choose to publish a series of articles, however, with each article relying on a different method or mix of methods.

[5] Similar challenges arise in the analysis of existing databases, which may not include all variables of interest or may use measures that do not correspond well to underlying concepts.

[6] For more information, see the websites for RNGS (http://libarts.wsu.edu/polisci/rngs/), CLEA (http://www.electiondataarchive.org/), and Afrobarometer (http://www.afrobarometer.org).

[7] On a somewhat different point, Sigelman (2009) found that co-authorship was associated with higher acceptance rates when teams of co-authors included both political scientists and non-political scientists.

[8] There are currently twelve centers in Africa, Asia, Latin America, and the United States. See the IFRI website for more information (http://sitemaker.umich.edu/ifri/home).

References

Bennett, Andrew, Aharon Barth, and Kenneth R. Rutherford. 2003. "Do We Practice What We Preach? A Survey of Methods in Political Science Journals and Curricula." *PS: Political Science and Politics* 36:3 (July), 373–378.

Doucouliagos, Hristos and Martin Paldam. 2008. "Aid Effectiveness on Growth: A Meta Study." *European Journal of Political Economy* 24:1 (March), 1–24.

Doucouliagos, Hristos and Ali Mehmet Ulubasoglu. 2008. "Democracy and Economic Growth: A Meta-Analysis." *American Journal of Political Science* 52:1 (January), 61–83.

Geys, Benny. 2006. "Explaining Voter Turnout: A Review of Aggregate-Level Research." *Electoral Studies* 25:4 (December), 637–663.

Glass, Gene V. 1976. "Primary, Secondary, and Meta-Analysis of Research." *Educational Researcher* 5:10 (November), 3–8.

Lam, Wai Fung. 1996. "Improving the Performance of Small-Scale Irrigation Systems: The Effects of Technological Investments and Governance Structure on Irrigation Performance in Nepal." *World Development* 24:8 (August), 1301–1315.

Lau, Richard, Lee Sigelman, and Ivy Brown Rovner. 2007. "The Effects of Negative Political Campaigns: A Meta-Analytical Reassessment." *Journal of Politics* 69:4 (November), 1176–1209.

Lohmann, Susanne. 2007. "The Trouble with Multi-Methodism." *Qualitative Method*s 5:1 (Spring), 13–17.

Lam, Wai Fung. 1998. *Governing Irrigation Systems in Nepal: Institutions, Infrastructure, and Collective Action*. Oakland, CA: ICS Press.

Munck, Gerardo L. and Richard Synder. 2007a. "Debating the Direction of Comparative Politics: An Analysis of Leading Journals." *Comparative Political Studies* 40:1 (January), 5–31.

Munck, Gerardo L. and Richard Synder. 2007b. "Who Publishes in Comparative Politics? Studying the World from the United States." *PS: Political Science and Politics* 40:2 (April), 339–346.

Nugent, Jeffrey B. and Nicholas Sanchez. 1999. "The Local Variability of Rainfall and Tribal Institutions: The Case of Sudan." *Journal of Economic Behavior and Organization* 39:3 (July), 263–291.

Ostrom Elinor, James M. Walker, and Roy Gardner 1992. "Covenants with and without a Sword: Self-Governance is Possible." *American Political Science Review* 86:2 (June), 404–417.

Pagdee, Adcharaporn, Yeon-Su Kim, and P. J. Daugherty. 2006. "What Makes Community Forest Management Successful: A Meta-Study from Community Forests throughout the World." *Society and Natural Resources* 19:1 (January), 33–52.

Poteete, Amy R., Marco A. Janssen, and Elinor Ostrom. 2010. *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*. Princeton, NJ: Princeton University Press.

Rudel, Thomas K. 2005. *Tropical Forests: Regional Paths of Destruction and Regeneration in the Late Twentieth Century*. New York: Columbia University Press.

Ruttan, Lore M. 2006. "Sociocultural Heterogeneity and the Commons." *Current Anthropology* 47:5 (October), 843–853.

Ruttan, Lore M. 2008. "Economic Heterogeneity and the Commons: Effects on Collective Action and Collective Goods Provisioning." *World Development* 36:5 (May), 969–985.

Sager, Fritz. 2006. "Policy Coordination in the European Metropolis: A Meta-Analysis." *West European Politics* 29:3 (May), 433–460.

Schedler, Andreas. 2012. "The Measurer's Dilemma: Coordination Failures in Cross-National Political Data Collection." *Comparative Political Studies* 45:2 (February), forthcoming.

Schlager, Edella. 1994. "Fishers' Institutional Responses to Common Pool Resource Dilemmas." In *Rules, Games, and Common-Pool Resources*. Elinor Ostrom, Roy Gardner, and James M. Walker, eds. (Ann Arbor: University of Michigan Press), 247–266.

Schlager, Edella, William Blomquist, and Shui Yan Tang. 1994. "Mobile Flows, Storage and Self-Organized Institutions for Governing Common-Pool Resources." *Land Economics* 70:3 (August), 294–317.

Sigelman, Lee. 2009. "Are Two (or Three or Four…or Nine) Heads Better than One? Collaboration, Multidisciplinarity, and Publishability." *PS: Political Science and Politics* 42:3 (July), 507–512.

Tang, Shui Yan. 1994. "Institutions and Performance in Irrigation

Systems." In *Rules, Games, and Common-Pool Resources*. Elinor Ostrom, Roy Gardner, and James M. Walker, eds. (Ann Arbor: University of Michigan Press), 225–245.

Integrating Two Cultures in Mixed-Methods Research: A Tale of the State Feminism Project

Dorothy E. McBride
Florida Atlantic University
dmcbrid6@fau.edu

Amy G. Mazur
Washington State University
mazur@wsu.edu

As the growing literature on multi-methods research (MMR) shows, there are distinct advantages and disadvantages to conducting such studies. While some assert that MMR should be the industry standard in political science and is becoming the norm, particularly in the subfield of Comparative Politics (Coppedge 2009), others take a "cautionary perspective," arguing that the trend toward MMR may actually undermine good political science scholarship.[1] Interestingly, both perspectives play on the differences between qualitative and quantitative approaches, differences that originate from divergent worldviews, concepts and analytical logics—the "two cultures" (Mahoney and Goertz 2006). For advocates, combining qualitative and quantitative methods brings more analytical leverage to studies through, for example, addressing omitted variable bias, identifying causal mechanisms, or developing more valid concepts (e.g., Bennett 2007 or Collier, Brady, and Seawright 2004). For detractors, the fundamental methodological differences undermine the accuracy and validity of mixed-methods studies through unsophisticated and often incorrect use of different methods, differentiated measurement of key concepts and general "epistemological incommensurability" (Ahmed and Sil 2009).

The goal of this article is to show how the differences between the "two cultures" have been used to methodological and theoretical advantage in one large-scale comparative study, the State Feminism Project. Drawing from the study's process and results, we illustrate how an integrated approach to the "two cultures" can enhance empirical research and develop a theory of state feminism about women's policy agencies, women's movements and the state in Western postindustrial democracies. In doing so, we show how some of the traps or "speed bumps" (Coppedge 2009) of MMR can be overcome by conducting an "integrated concurrent" (Creswell 2003) approach to the results of different methods to reach a productive combination of the divergent methodological traditions.

We first present the State Feminism Project in terms of how it sought to "choose not to choose" (Mazur and Parry 1998) between the two cultures from its beginnings in 1995 to the recent completion of the concurrent and integrated mixed-methods capstone analysis in McBride and Mazur (2010). Next, we take a closer look at the effect of this integrative approach first on conceptualization and then on theory development by illustrating how fitting the findings from different methods advanced understanding of the puzzle of state feminism. Our essay concludes with a brief discussion of the lessons learned from bringing the two cultures together in one study.

The State Feminism Project: A Pragmatic and Integrated Approach to MMR

In many ways the State Feminism Project is unique. It has a single set of descriptive cases collected by a research group of 43 people, the Research Network on Gender Politics and the State (RNGS), that benefited from generous research funding. The group also developed a common theoretical framework that guided data collection and analysis through an integrated analytical logic, published in five issue-specific books (McBride Stetson 2001; Mazur 2001; Outshoorn 2004; Lovenduski et al. 2005; Haussman and Sauer 2007). RNGS then worked together to transpose the qualitative concepts, measurements, and data into a quantitative data set.[2] Following the completion of the comparative issue–area books and the data set, the capstone analysis addressed research propositions from the state feminism framework developed by RNGS using three methods, seeking an integrated concurrent mixed-methods analysis. Thus, we are not suggesting that others can conduct a study of the same magnitude, but that the strategy and principles used in the State Feminism Project can expand the MMR agenda, open up the menu of effective practices, and perhaps assuage some of the recent criticisms.

Founding Principles: Empirical Feminism, Methodological Pragmatism, and Integration

RNGS was founded in 1995 as a response to the weaknesses of an initial study of women's policy agencies in 14 western postindustrial democracies (McBride Stetson and Mazur 1995). Many contributors to this edited volume and other experts agreed that the country-case studies of agencies were insufficient to assess the effectiveness of the relatively new government structures assigned to promote women's rights and gender equality. From the first founding meeting, members of RNGS shared a set of common principles about research. First, we opted for an empirical feminist approach where hypothesis testing, standards of replication, and transparent measurements and indicators are combined with a focus on gendered processes and an effort to bridge the gap between feminist and non-feminist scholarship related to gender, movements, and the state.[3]

Second, RNGS members shared a methodological pragmatism with respect to research methodologies, arguably the bedrock of good mixed-methods research (Creswell 2003). That is, we were willing to consider whatever methods would help answer the core question of the study: if, how, and to what degree do women's policy agencies achieve *state feminism* through bringing women's movement interests into government affairs and policy? Part and parcel of this open-mindedness toward methodology was an understanding that qualitative and quantitative approaches could be useful in developing a systematic cross-national and longitudinal study of the dynamics and drivers of state feminism. In fact, from the beginning, inspired by King, Keohane, and Verba's, *Designing Social Inquiry*, RNGS agreed to incorporate elements of both approaches in the original research design; in other words "choosing not to choose" (Mazur and Parry 1998) one approach over the other.

The following decisions for the research design, made by the network as a whole, guided the project researchers through the collection of data about and analysis of the interactions among women's policy agencies, policy debates, and women's movements in 16 countries pertaining to five policy areas from the 1960s to the early 2000s, finally resulting in the five issue-based books. It is important to note the degree to which these decisions intertwine both qualitative and quantitative logics. RNGS was, therefore, wedded to an integrative strategy prior to the development of an explicit mixed-methods plan.

(1) *Comparative Method with Policy Debate as the Unit of Analysis*—Overall, RNGS followed a most-similar-systems design to control for levels of economic and political development, studying agency activity only in Western postindustrial democracies. At the same time, the group wanted to maximize the number of observations, deciding that the unit of analysis would not be a country, but a specific policy debate within a country. Researchers agreed to study the influence of agencies on between one and five debates in each policy area for each country, for a final N of 130 policy debates in the data set.[4]

(2) *Quantitative Universe Construction and Sample Selection*—Given the group's goal of doing a systematic analysis of agencies, RNGS sought to enhance the representativeness of the cases, following ideas based in quantitative analysis and sampling: (a) to expand the range of issue areas to cover policies that have gender dimensions—job training (work and family), prostitution (sexuality), abortion (reproduction), and political representation (citizenship); (b) to add a fifth issue of national importance regardless of gender dimension; (c) to establish criteria to guide researchers toward a systematic selection of debates in each gender dimension area to enhance coverage over time, salience, and institutional arenas. Each policy, therefore, was a stand-alone analysis of agency-movement interactions over the course of the policy process while at the same time a case or observation providing information about a common theoretical framework.

(3) *Qualitative Case Analysis*—In the interest of trying to understand the specific role of agencies in affecting policy debates and bringing women's movement actors and their interests into each debate, RNGS agreed that researchers would follow standard methods of process-tracing to analyze the dynamics and outcomes of each debate using archival research, interviews, and consultation of secondary and primary resources. Worksheets were used to ensure uniform debate selection and analysis, to standardize as much as possible the data collection process, and to provide the maximum potential for replication.

(4) *Model Specification*—RNGS designed an analytical model, informed by both feminist and non-feminist comparative politics literature, setting forth dependent, independent, and intervening variables to generate hypotheses about the dynamics and determinants of women's movement success with the state and the intervening influence of agencies in those outcomes. Three clusters of independent variables covered women's movement resources, policy environment characteristics, and left support. At the same time, RNGS followed the tradition of small N comparative studies with two typologies based on nominal measures. These define women's movement impact—the dependent variable—and women's policy agency activity—the intervening variable.

Moving from Descriptive Cases to a Numerical Dataset

In the study's fifth year, RNGS members decided to transpose the qualitative case analyses published in the issue books into a numerical dataset that could be used to test hypotheses about state feminism across all the debates in the study as well as become a resource for researchers outside the project. This process involved a complete review and enhancement of conceptualization, setting forth operations for measurement that were valid and reliable, securing agreement of the network researchers, and finally asking them to complete additional worksheets to provide the necessary information for the measures. Released in 2007, the RNGS data set has information on 28 concepts from the state feminism framework with measurements for 120 variables that pertain to these concepts, including numerical indicators (98 variables) and descriptive information (22 variables) for each policy debate. The data set suite includes an SPSS file, a 130-page codebook, and 700 pages of text appendices with the descriptive information on each debate. Thus, even the quantitative data set includes both qualitative and quantitative information. The RNGS data set corresponds with what Lieberman identifies as "Historically-oriented and Integrated Replication Data Bases": "systematically collected and theoretically informed containers of facts and observations for a consistent set of units over time" (2010: 39).

The Concurrent Integrated Approach to MMR in the Capstone Analysis

The final step in the State Feminism Project was to explore the propositions from the RNGS framework across all the issues, countries, and decades of the study, using the numerical data set, the text appendices, and the case studies. After ten years of research and reporting, the central question remained the same: to what extent does state feminism exist and what are its causal drivers? From this question the capstone analysis extracted several propositions and examined them using appropriate methods for each. With RNGS data and studies as a launching pad, we thus conducted a mixed-methods analysis that developed systematic understanding about state feminism and its component parts—representation, women's movements, debate framing, and feminist institutions—in terms of both description and a larger theory of state feminism. By providing a composite descriptive and theoretical picture of agencies in Western postindustrial democracies from the 1960s to the early 2000s, the capstone analysis and book, *The Politics of State Feminism: Innovation in Comparative Research*, is distinctive from the more general RNGS study.

Given the overall approach of the RNGS project, adopting an integrated approach to the capstone analysis was necessary if not inevitable. Typically, MMR in political science uses two methods in sequence, e.g., an in-depth case study to validate findings from statistical analysis (e.g., Lieberman 2005) and, more recently, case studies in relation to formal theory (e.g., Dunning 2007). Unlike that sequential approach, we took a more "iterative" approach, which "leverages the distinctive but complementary strengths of different research methods to make progress on substantively important topics" (Dunning 2007: 22). This is also referred to as "triangulation" (Collier, Brady, and Seawright 2004). We prefer to use the term offered by Creswell—a concurrent integrated approach—since it suggests bringing different logics of qualitative and quantitative approaches together for more accurate and theoretically meaningful results, not just "adding case studies and stirring." We combine three different methods that cut across the qualitative–quantitative divide—descriptive and inferential statistics, comfigurative qualitative analysis, and causal mechanisms case studies. As will be illustrated in the examples below, we integrate findings, recognizing the analytical logic of each approach, to develop a new theory of state feminism. We found that, rather than seeking mutual validation for each method, this integrative approach is similar to the "alternative logic" of MMR of "fitting-together of a puzzle" identified by Ahram (2009: 9).

Methodological Integration in Conceptualization

In this section, we discuss the ways that the decision to use multiple methods led to a necessary and valuable exploration of the central state feminism concepts leading to new discoveries and greatly expanding the significance of the work. When we started planning the RNGS project, we wanted to "do science"; we belonged to the culture that values comparative analysis, causal inference, replicable methods, and empirical validity. Yet, we had chosen a topic where information was so limited that we had to start from scratch to gather data through descriptive research, case by case. In developing the research plan in the 1990s we were heavily influenced by KKV; thus, we intended to have enough cases to provide the basis for empirically valid findings and contribute to cumulative knowledge and theory building.

We thought long and hard about the challenges: (1) many researchers from different countries and backgrounds were needed to study the cases and write up the results; (2) we were studying policy debates across five different issue areas (abortion, job training, prostitution, political representation and priority issues of the 1990s); and (3) the span of the study—from the 1960s to the early 2000s—covered a period of change in just about every aspect of the topic. We knew that if we were to be able to deliver an empirically valid theory of state feminism that we had to pay careful attention to conceptualization: from nominal definitions to clear steps for gathering the information about those concepts, i.e., operationalization.

The first collaborative meetings of RNGS focused on the research plan—specifically, conceptualization of the major components of the RNGS model. For example, *gendering* was a key concept in connecting the women's policy agencies, women's movements, and policy debates. Influence of women's movements and women's policy agencies was determined by the extent to which their activities led to the gendering of policy debates—bringing explicit gender language and ideas into the issue definitions used by policy actors. We spent a good deal of time figuring out how to define gendering and how to determine whether or not the debates were gendered. We also clarified what we meant by other terms such as *compatible policy content* and *procedural access*. The RNGS project description, an ever-evolving document, was the "handbook" for this work, and the worksheets filled in by each researcher showed the data results.

It was not until we moved into what we called the "quantitative phase" of the project, developing the data set containing numerical measures of all the concepts of the model, that we learned the limits of the conceptualization up to that point. The chapters of the issue books contained the information on each of the cases. We discovered, as we started to put them together, that the researchers at times did not use the same definitions and indicators even within the same country. Editors of the issue books worked with these chapters to count and sort into cross-tabs; they did the best they could to draw comparative conclusions using the case descriptions. However, in developing a quantitative version of these studies, we confronted the bad news that we had not met our goal of empirical validity, let alone reliability in the cases.

But the news got even worse. We also saw that the researchers had a great range of opinion as to what entities were parts of women's movements and which actors and goals were feminist. On reflection, it seems that we all thought we knew what the women's movement was and that there was no need to go further. However, we were long aware that the term *feminism* when uttered was likely to provoke a prolonged argument, not only about what feminism meant, but whether the term could be used across time and cross-nationally. Some researchers did not want to be labeled "feminist" and they did not want to take the responsibility to call political actors in their countries "feminist."

We could not leave it there, however. What a women's movement is—that is how to observe the movement in scientific research—could no longer be ignored once the quantitative phase began. It was too central to the RNGS model and later, the state feminism framework for the capstone analysis. The central research question was whether women's policy agencies promoted women's interests in policy and helped women gain procedural access to policy arenas. We used *women's movement* as the indicator of women's interests and participation. Thus, we had to determine the demands of the movement in every policy debate and whether such demands were picked up by the agencies and coincided with the gendering of debates. It was essential to identify those entities that were (and were not) part of the women's movement in each country.

While a thorough understanding of the women's movement concept was important to answer the central research question, the meaning of feminism was even more important: the overall framework for analysis was *state feminism*. How could we do a study of state feminism if we could not agree on a definition of feminism? Of course the use of the term "feminism" in public discourse has always been contentious and certainly imprecise. But the rigorous criteria of validity and reliability so central to quantitative analysis gave us no excuse to just let it ride. We had to solve that conceptual problem as well.

This story of these conceptualization nightmares and their resolution are well described elsewhere (McBride and Mazur 2008; Mazur and McBride 2008). The analytical distinction between two parts of the women's movement—ideas and actors who present them—was the key. Researchers could thus identify women's movement actors (as opposed to a collective notion of a movement) by their gendered ideas articulated in the public sphere. The feminist ideas became a subset of the movement ideas and could also be identified in policy debates and outcomes. The distinction between feminist movements as a subset of the women's movement, however, led to an awkward situation: what to call "state feminism"? The RNGS project defined it in terms of the relation between the agencies, the movement demands, and policy outcomes. But, with this more refined and rigorous conceptualization of feminism, the question arose: what if the state accepts women's movement goals that are non-feminist? Is this state feminism? (Non-feminist state feminism?) We bit the bullet and made the same distinction in this concept as we did in the women's movement/feminism concept. State feminism is the alliance between agencies and women's movement actors to achieve positive state response. State responses that incorporate feminist movement goals comprise a subset of state feminism. There are, thus, two types: Movement State Feminism and Transformative State Feminism.

All of this led to a reorganization of the capstone analysis. The propositions relating to Movement State Feminism and those pertaining to Transformative State Feminism are treated in separate chapters and the findings compared. We discovered, using a quantitative measure of feminist state response, that the postindustrial democracies have actually incorporated ideas and participants whose intent is to undermine the male-dominated underpinnings of the state itself, albeit at a much lower rate than the typically positive responses to women's movement demands more generally.

The main point here, however, is that our interests in expanding the KKV case-based comparative analysis toward quantitative analysis through a numerical data set of the cases required us to expand and deepen the conceptualization of key components of the research design. Often we read that qualitative researchers are more attentive than quantitative researchers to complex concepts. Our experience shows that the standards of cross-case reliability and validity essential to numerical measurement of case-descriptive information were the impetus for more rigorous and complex conceptualization.

Mixing Results of Three Methods to Build State Feminism Theory

We turn now to how we integrated three methods with different underlying analytic logics: statistical inference, crisp-set qualitative comparative analysis (csQCA), and tracing causal mechanisms in case studies. Initially our focus was on expanding the number of cases to allow statistical methods using the quantitative data set. The research goal was to assess the independent influence of women's policy agency activity (as allies) on state responses to women's movement demands while controlling for effects of the characteristics of the movement actors and the policy environment during the debates. This way we expected to cross the bridge from descriptive cases to quantitative territory.

The more we immersed ourselves in the details of the cases of policy debates and rich descriptions of the policy-making processes ranging from abortion battles in the 1970s to questions of state change in the late 1990s, however, the more we realized that newer methods of configurative comparative research showed promise for deepening our understanding and contributing to theory development. With QCA, for example, we learned we could retain some of the complexity of the cases and still come to conclusions about what factors produce positive state responses for women's movement actors and the role of agencies in those outcomes. Thus, we added a third method to our multi-method approach. And, as the following discussion shows, this decision made all the difference between reporting quite disparate findings from quantitative analysis and qualitative comparative methods and the integrated set of findings that resulted. To illustrate, we tell the tale of how we discovered the *Backup Theory of State Feminism*.

As indicated earlier, the central question of state feminism is how important women's policy agencies are to women's movement success with the state. One way to explore this question is in terms of a hypothesis: *Alliances with women's policy agencies are a significant cause of women's movement procedural and policy successes in postindustrial democracies*. This hypothesis lends itself to techniques of causal inference while controlling for the effects of other influences such as the degree of cohesion among movement actors, the priority of the issue to the movement, the openness of the policy subsystem to movement actors, and receptivity of policy makers to movement ideas and so on. We used ordinal regression (OR) techniques to test several models of state responses to women's movement demands, each including a measure of women's policy agency activity (degree of alliance). We found that agency activity is an independent influence on favorable state response: the probability of successful response increased significantly with more agency activity. However, other variables were significant as well—subsystem openness and issue priority to the movement—and these had higher odds for success.

The OR runs showed us we needed to understand more about agency activity in the context of the policy environment and movement resources, and so we considered another proposition: *Women's movements are more likely to be successful* *with the state when women's movement actors have more resources and a cohesion, they consider the matter a high priority, and the policy environment is open and compatible with movement goals*. *Activities of agencies complement these conditions in achieving women's movement success*. This question requires a method, like QCA, that allows discovery of which combinations of conditions are sufficient to achieve women's movement success with the state and when and if effective agency activity is one of those conditions. Our csQCA (Crisp Set Qualitative Comparative Analysis) analysis at first was disappointing: there was no evidence that it is ever necessary or even sufficient for women's policy agencies to be insiders for movements to gain success with the state. A closer look at the analysis of the job training debates, however, revealed something very interesting.

Job training is an issue where women's movement actors have had great difficulty in many countries in penetrating the policy subsystems that control training and vocational education programs. When they have been successful, they have benefited from favorable policy environments—open policy subsystems and policy debates framed in terms that are compatible with movement goals. In those cases it did not matter whether or not there was an active, effective agency. There were some successes, however, where those favorable conditions were not present. For these, movements found success because there were women's policy agencies inside the policy debates and they brought about a successful outcome. Clearly, in job training debates, women's policy agencies play a backup role to help movements when otherwise favorable conditions are not present.

When we turned from *issue*-based analysis to a *country*-based analysis, we found more evidence of this backup role for women's policy agencies. Here, we classified the countries in terms of the proportion of women's movement successes among the debates across all issues. We looked for patterns that might explain the place of agencies as allies for the more successful movement outcomes in comparison with the less successful, again in relation to conditions of movement resources and the policy environment. This analysis included both csQCA, for countries where there were enough cases that met the conditions for this method, and case studies of causal mechanisms. Evidence supporting the backup role of women's policy agencies mounted. In Canada, for example, a high degree of fit between movement actor demands and the approach of policy actors at the beginning of debates is a sufficient condition for state response whether or not there is an active, effective agency. However, in one successful debate, those conditions were not present and a case study of causal mechanisms traced the cause to the activity of an agency. A similar pattern was found in Finland. The evidence became even stronger in looking at countries with few movement successes, such as France and the Netherlands where, in the debates in this study, only with the help of effective agencies were the movement actors successful.

To summarize, the backup theory of state feminism asserts that the greater the activity of policy agencies on behalf of women's movement actors, the greater the degree of positive response by the state to their demands. However, there are other variables that increase the odds of success more than agency activity. Such conditions, alone or in combination, are frequently sufficient for women's movements to gain positive responses from the state: priority of the issue, openness of subsystems, and compatibility of movement demands with policy actors' views. These movement actors achieve success regardless of any alliance with a women's policy agency. When these favorable conditions are not present, however, active, effective agencies are likely to help actors overcome the barriers and bring about success in policy and participation. These findings are made possible by and supported by the integration of all three methods—statistical inference, configurative comparative analysis, and causal mechanism case studies—which in turn validate each other and also allow for a more full and accurate picture of the place of women's policy agencies as allies of women's movement actors and as advocates for movement goals with the state.

Conclusion: Lessons Learned

The State Feminism Project clearly shows the benefits of integrating qualitative and quantitative approaches in a single study, countering fears of "epistemological incommensurability" (Ahmed and Sil 2009). The move from the qualitative to the quantitative phases of the project improved the precision and reliability of the operational definitions of key concepts, contrary to the conventional wisdom that qualitative analysis leads to better concepts. The dialog among ordinal regression, csQCA, and case studies in the capstone analysis helped us put together the components, like the backup theory, of a new theory of state feminism. By returning repeatedly to results from each method, it was possible, gradually, to make sense of the vast array of information about women's movements, agencies, and policy processes over time, across countries, and across issues. This concurrent integrated strategy produced mid-range theory, undermining sweeping generalizations from feminist and non-feminist scholarship about causes of social movement success, the effectiveness of women's policy agencies, and the receptiveness of states to women's movement activism. This study, therefore, undertakes "double bridging" (Checkel 2008), across both the qualitative/quantitative *and* the feminist/non-feminist divides.

To be sure, the time and resources invested in conducting the State Feminism project as well as its scope and magnitude make it unlikely that others will be able to assemble the number and range of original case studies necessary to reach the same level of mixed-methods integration. In addition, because of its origins in these cases, it could be argued that the RNGS numerical data set is not the typical-random sample data set with interval data producing linear regression; while it is quantitative, its nominal and ordinal measures contain the qualitative logic of its origins. Thus, the study is more likely to be compatible with research that uses more explicitly qualitative approaches than with scholarship that wholly embraces quantitative logics. And finally, the complexity and mid-range nature of the findings reflects the risk of undertaking mixed-method research: after all the work and time expended exploring the research questions, the answers lack the parsimony and elegance of macro theories. In the final analysis, however, finding complexity, bounded generalizations, and mid-range theory is likely to present the more accurate picture of the reality of politics; therefore, integrating the two cultures may lead to better science after all.

Notes

[1] See, for example, the recent symposium in this newsletter, "Cautionary Perspectives on Multi-Method Research." 7:2 (Fall 2008).

[2] For the project description, data set suite, and codebook, and other specifics on RNGS go to http://libarts.wsu.edu/polisci/rngs/.

[3] A second approach to feminist analysis is "standpoint feminism," where the scientific method is put into question for being "tainted" by patriarchy and exclusive of serious consideration of gender and women researchers. This approach employs interpretive and postmodern epistemologies, and researchers tend to reject standard social scientific protocols (Harding 1986).

[4] The 130 debates in the RNGS dataset come from 13 Western countries. Debates covering prostitution in Israel, political representation in Japan, prostitution in Australia, and job training in the EU were not included in the data set but appear in the issue books.

References

Ahmed, Amel and Rudra Sil. 2009. "Is Multi-Method Research Really 'Better'?" *Qualitative and Multi-Method Research* 7:2 (Fall), 2–6.

Ahram, Ariel I. 2009. "The Challenge of Conceptual Stretching in Multi-Method Research." *Qualitative and Multi-Method Research* 7:2 (Fall), 6–10.

Bennett, Andrew, ed. 2007. "Symposium: Multi-Method Work, Dispatches from the Front Lines." *Qualitative Methods* 5:1 (Spring), 9–28.

Checkel, Jeffrey T. 2008. "Introduction to Symposium on Bridging the Gap? Connecting Qualitative and Quantitative Methods in the Study of Civil War." *Qualitative Methods* 6:1 (Spring), 13–14.

Collier, David, Henry E. Brady, and Jason Seawright. 2004. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Henry E. Brady and David Collier, eds. (Lanham MD: Rowman and Littlefield), 195–228.

Coppedge, Michael. 2009. "Speedbumps on the Road to Multi-Method Consensus in Comparative Politics." *Qualitative and Multi-Method Research* 7:2 (Fall), 15–17.

Creswell, John W. 2003. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd ed. Thousand Oaks, CA: Sage.

Dunning, Thad. 2007. "The Role of Iteration in Multi-Methods Research." *Qualitative Methods* 5:1 (Spring), 22–24.

Harding, Sandra. 1986. *The Science Question in Feminism*. Ithaca, NY: Cornell University Press.

Haussman, Melissa and Birgit Sauer, eds. 2007. *Gendering the State in the Age of Globalization. Women's Movements and State Feminism in Postindustrial Democracies*. Lanham, MD: Rowman and Littlefield.

Lieberman, Evan. 2005. "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99:3 (August), 435–452.

Lieberman, Evan. 2010. "Bridging the Qualitative-Quantitative Divide: Best Practices in the Development of Historically Oriented Replication Databases." *Annual Review of Political Science* 13, 37–59.

Lovenduski, Joni, Claudie Baudino, Marila Guadagnini, Petra Meier, and Diane Sainsbury, eds. 2005 *State Feminism and the Political Representation*. Cambridge: Cambridge University Press.

McBride, Dorothy E. and Amy G. Mazur. 2008. "Women's Movements, Feminism and Feminist Movements." In *Politics, Gender and Concepts: Theory and Methodology.* Gary Goertz and Amy G. Mazur, eds. (Cambridge: Cambridge University Press), 219–243.

McBride, Dorothy E. and Amy G. Mazur. 2010. *The Politics of State Feminism: Innovation in Comparative Research*. Philadelphia: Temple University Press.

McBride Stetson, Dorothy and Amy Mazur, eds. 1995. *Comparative State Feminism*. Thousand Oaks, CA: Sage Publications.

McBride Stetson, Dorothy, ed. 2001. *Abortion Politics, Women's Movements and the Democratic State: A Comparative Study of State Feminism*. Oxford: Oxford University Press.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Quantitative and Qualitative Research." *Political Analysis* 14:3 (Summer), 227–249.

Mazur, Amy G., ed. 2001. *State Feminism, Women's Movements, and Job Training: Making Democracies Work in the Global Economy*. New York and London: Routledge.

Mazur, Amy G. and Dorothy E. McBride. 2008. "State Feminism" In *Politics, Gender and Concepts: Theory and Methodology*. Gary Goertz and Amy G. Mazur, eds. (Cambridge: Cambridge University Press), 244–269.

Mazur, Amy and Janine Parry 1998. "Choosing not to Choose in Comparative Policy Research Design: The Case of the Research Network on Gender, Politics and the State." *Journal of Policy Studies* 26:3 (September), 384–395.

Outshoorn, Joyce, ed. 2004. *The Politics of Prostitution: Women's Movements, Democratic States, and the Globalization of Sex Com-*

Announcements

APSA Short Courses Created (or Co-Organized) by Division 46: Qualitative and Multi-Method Research Wednesday, September 1, 2010, Washington, DC

Short Course 1: Multi-Method Research

Time: 9:00am–1:00pm
Lead Instructor: David Collier, University of California, Berkeley

Attention has increasingly focused on how qualitative methods can be linked to other analytic tools, including large-N quantitative analysis and formal modeling. To this end, methodologists have urged scholars to "nest" their case studies within small- to medium-N comparisons, and/or within large-N quantitative analysis.

Given that many political scientists are now convinced that good research necessarily employs multiple methodologies, how can different approaches be combined to maximize analytic leverage? How useful are the multi-method techniques under discussion here? Is it sometimes better to stick with one method, and to focus on using it with great skill? This short course explores alternative examples and strategies of multi-method research, with the goal of addressing these questions.

Short Course 2: Designing and Conducting Field Research

Time: 2:00pm–7:00pm
Lead Instructor: Diana Kapiszewski, Univ. of California, Irvine

This short course aims to help scholars hone their empirical research skills. Two foundational premises of the course are that fieldwork begins long before one enters the field, and that the way fieldwork is carried out—that is, what data are collected and how they are collected—has important ramifications for analysis and theory-generation. Although field methods are usually associated with "studying politics abroad," we discuss techniques that may be applied inside and outside the U.S.

We will begin by examining different "varieties" of fieldwork and exploring how research design and fieldwork intersect—since planning for the effective use of field methods and the efficient collection of data are crucial aspects of overall research design. The bulk of the course will focus on logistical and intellectual aspects of conducting field research. We will discuss the usefulness and practice of interactive and non-interactive forms of data-collection (including interviewing, oral histories, focus groups, using archival sources, collecting documents and ephemera, and ethnographic study). The last section of the course explores the challenges involved in managing, analyzing, and evaluating data both in and out of the field. The course includes several hands-on activities.

Scholars typically initiate their projects by mapping out their analytic questions, thinking about how they will measure their variables and what evidence they will need to support their claims, and beginning to identify potential sources for the data they hope to collect. Yet even if the research is well planned and adequately funded, obstacles can arise. Key respondents may be unhelpful or unavailable. Valuable archives and other collections of primary materials may be accessible only on a limited basis or may be poorly organized. Data necessary for constructing sampling frames for formal or informal interviewing may simply not exist. Time or money may run out before essential data have been collected.

The course will help analysts to anticipate and address many of the challenges involved in designing and conducting field research. We discuss strategies that will allow scholars to: (1) convert their research design into a "to get" list; (2) identify and begin to investigate data sources before leaving their home institution; (3) make optimal use of relevant technologies (e-mail, web, cell phones, portable photocopying equipment, scanners, digital cameras, and voice and video recorders); (4) respond to the availability of data not anticipated in the original research design, and to the inaccessibility of data that were originally to be collected; (5) organize and manage information gathered; (6) establish key contacts and interact constructively with actors of all types in the host community; (7) cope with professionally, politically, and personally uncomfortable situations; and (8) make the transition from data collection to data analysis and writing in a timely manner.

Following the end of the formal class at 6:00pm, the instructors will hold a "workshop" in which short-course participants will have the opportunity to discuss their own research and the design and conduct of their own fieldwork in a smaller-group setting. We encourage students to stay for this more informal conversation, and to bring along questions about their work.

Participants will be provided with document templates that may be useful when carrying out field research, including sample correspondence. The course is valuable for students planning dissertation projects, for scholars who would like to develop or improve their data collection skills, and for those who teach classes on research methods.

APSA Panels/Roundtables Created (or Co-Organized) by Division 46: Qualitative and Multi-Method Research September 2–September 5, 2010, Washington, DC

QMMR Meets IR: Rethinking Classic and Contemporary Methodological Issues in International Relations

Chair: Robert Kaufman Adcock, George Washington University
Participants:
Robert Kaufman Adcock and Tristan Volpe, George Washington University: "Integrating Economic and Sociological Approaches to International Relations? Multi-Methods and a Classic Methodological Puzzle."
Eric Grynaviski, Ohio State University: "Interpretivism, Game Theory, and Counterfactual Analysis: Bridging a Divide in the Study of International Politics."
Patrick Thaddeus Jackson, American University: "The Road Not Taken: Analyticism and Configurational Analysis in IR."
Kevin A. Clarke and Bear Braumoeller, University of Rochester: "Causation, Description, and Inference in International Relations."
Discussant: James D. Johnson, University of Rochester

New Approaches to Qualitative Methods

Chair: Darren G. Hawkins, Brigham Young University
Participants:
Jason Seawright, Northwestern University: "Means and Ends: Evaluating Case-Selection Strategies in Light of Common Case-Study Objectives."
Adam Glynn and Kevin Quinn, Harvard University: "Combining Case Studies and Regression (or Other Large-N Techniques) for Population Causal Inference."
Richard Nielsen, Harvard University: "Case Selection via Matching."
Matthew A. Kocher, Yale University: "Comparative Politics as Comparative History."
Discussant: Darren G. Hawkins, Brigham Young University

Diffusion Studies: Methodological and Conceptual Issues

Chair: Kurt Weyland, University of Texas, Austin
Participants:
Fabrizio Gilardi and Fabio Wasserfallen, University of Zurich: "Research Design and the Study of Policy Diffusion."
Katerina Linos, Harvard University: "International Influences on National Health Reforms: A Mixed-Methods Approach."
Zachary Elkins, University of Texas, Austin: "From Spanish Revival to Beaux Arts: Spatio-Temporal Trends in the Architecture of National Constitutions."
Kurt Weyland, University of Texas, Austin: "The Emulation, Preemption, and Prevention of External Impulses—Multiple Facets of Diffusion?"
Discussants: Beth A. Simmons, Harvard University; Charles R. Shipan, University of Michigan

Concept Analysis: Advances and Applications

Chair: Cas Mudde, University of Antwerp
Participants:
Robert M. Mauro, State University of New York, Albany: "The Meaning of Ideology and Social Action."
Fred Eidlin, University of Guelph: "Concepts as Depth Probes, Concepts as Inquiry Blockers."
Lise Morjé Howard, Georgetown University: "Ethnocracy as Regime Type."

Sarah Chartock, College of New Jersey: "'Corporatism with Adjectives?' Conceptualizing Civil Society Incorporation and Indigenous Participation in Latin America."
Jurg Martin Gabriel, Swiss Federal Institute of Technology: "Concepts Matter—Reflections on Political Science Methodology."
Discussants: Cas Mudde, University of Antwerp; Adrienne LeBas, American University

Lost in Translation? Structural Challenges to Cross-National Political Measurement

Chair: Cas Mudde, University of Antwerp
Participants:
Andreas Schedler, Centro de Investigacion y Docencia Economicas: "Academic Market Failure: Data Availability and Quality in Comparative Politics."
Tom Ginsburg and James Melton, University of Illinois: "On the Interpretability of Law: Adventures in the Decoding of National Constitutions."
Nicholas Charron, Goteburg University: "The Quality of Government Data."
Jose Antonio Cheibub, University of Illinois, Urbana-Champaign: "Towards the Public Provision of Cross-National Political Data."
Kristen Renwick Monroe, University of California, Irvine: "Measuring Dependent Variables, Testing for the Same Phenomenon in Cross-Cultural Settings, and Using Qualitative Work to Develop Quantitative Tools."
Discussant: Philip Keefer, The World Bank

Designing and Conducting Field Research in the Social Sciences

Chair: William Reno, Northwestern University
Participants:
Diana Kapiszewski, University of California, Irvine: "Conceptualizing and Preparing for Fieldwork."
Lauren M. Morris MacLean, Indiana University: "Qualitative Interviewing: In-Depth Interviews, Focus Groups and Oral Histories."
Benjamin L. Read, University of California, Santa Cruz: "Site-Intensive Methods: Participant Observation and Ethnography."
Anu Chakravarty, University of South Carolina: "Field Methods in a Post-Conflict Setting: A Close Range Analysis of 'Partially Trusting' Field Relationships."
Discussants: Jeremy Pressman, University of Connecticut; Craig A. Parsons, University of Oregon

Roundtable: Analytic Electicism in the Study of World Politics

Chair: Peter M. Haas, University of Massachusetts, Amherst
Participants: Peter M. Haas, University of Massachusetts, Amherst; Christian Reus-Smit, Australian National University; T.V. Paul, McGill University; Alice D. Ba, University of Delaware; Andrew Bennett, Georgetown University
Discussants: Rudra Sil, University of Pennsylvania; Peter J. Katzenstein, Cornell University

Roundtable: Standards for Qualitative Research

Participants: Colin Elman, Syracuse University; Andrew Bennett, Georgetown University; James Mahoney, Northwestern University
Discussants: Ronald L. Rogowski, University of California, Los Angeles; Brian D. Humes, National Science Foundation

Field Research under Authoritarianism: Where Methods Meet Regime Type

Chair: Marc Morjé Howard, Georgetown University
Participants:
Ariel Ahram, University of Oklahoma: "A Scanner Darkly: Seeing Middle East Cultures through the Lens of Authoritarianism."
Paul Goode, University of Oklahoma: "What You See is What You Get: Democratic Reversals and Scientific Closure in Eurasia."
Andrew Mertha, Cornell University: "Doctor Without Borders: Going Rogue in China and Cambodia."
Ryan M. Sheely and Janet Lewis, Yale University: "Regimes, Riots, and Randomization: The Politics of Fieldwork in Contemporary East Africa."
Discussant: Mark Beissinger, Princeton University

Meet the Author: Margaret Somers, *Genealogies of Citizenship: Markets, Statelessness, and the Right to Have Rights*

Chair: Eileen McDonagh, Northeastern University
Participants: Frances Fox Piven, City University of New York, Graduate Center; Gary Herrigel, University of Chicago; Jeff Manza; Northwestern University; Michael C. Tolley, Northeastern University; Margaret R. Somers, University of Michigan

Interpreting Situated Discourse

Chair: Charles L. Mitchell, Grambling State University
Participants:
Caroline Shenaz Hossein, University of Toronto: "Going Local and Getting Access to the Garrisons in Kingston, Jamaica."
Adam Avrushin, University of Chicago: "Policies, Money, and Child Welfare Caseworkers: What's Guiding Best Interest Decision-Making?"
Emily D. Shaw, University of California, Berkeley: "Depth Frame Analysis."
Cyrus Ernesto Zirakzadeh, University of Connecticut: "Uncovering Meta-Narratives in Non-Narrative Exposition: *New York Times* Coverage of Post-Election Protests in Iran."
Daniel E. Esser and Ben Williams, American University: "How Poverty Trumps Inequality: A Tracer Study of Development Discourse."
Discussants: Charles L. Mitchell, Grambling State University; Kevin Costa, University of Massachusetts, Amherst

Mobilizing Across Borders: Transnational Dynamics of Civil War

Chair: Jeffrey T. Checkel, Simon Fraser University
Participants:
Kristin Marie Bakke, University College London: "Transnational Insurgency and Domestic Insurgent Mobilization: A Case Study of the Chechen War."
Stephan Hamberg, University of Washington: "Transnational Advocacy and Demobilization of Child Soldiers."
Fiona B. Adamson, University of London, SOAS: "Mechanisms of Diaspora Mobilization and the Transnationalization of Civil War."
Jeffrey T. Checkel, Simon Fraser University: "Transnational Dynamics of Civil War."
Discussants: Scott Gates, International Peace Research Institute; Elisabeth Jean Wood, Yale University

Ideas, Power, and Public Policy

Chair: Robert H. Cox, University of Oklahoma
Participants:
Vivien A. Schmidt, Boston University: "Analyzing Ideas and Tracing Discursive Interactions in Institutional Change: From Historical Institutionalism to Discursive Institutionalism."
Craig A. Parsons, University of Oregon: "Ideas, Position, and Supranationality.
Daniel Beland, University of Saskatchewan and Robert Cox, University of Regina: "Ideas, Power and Politics."
Robert C. Lieberman, Columbia University: "Ideas and Institutions in Race Politics."
Discussant: Kathleen R. McNamara, Georgetown University

Advances in QCA/fs Analysis

Chair: Andrew Yeo, Catholic University of America
Participants:
Dexter Boniface, Rollins College: "Rethinking Presidential Falls: A Qualitative Comparative Analysis."
Martin Binder, Social Science Research Center, Berlin: "The Selectivity of UN Humanitarian Intervention: A Fuzzy-Set Analysis."
Steven Samford, University of New Mexico: "Rethinking Trade Liberalization in Latin America: An Application of Fuzzy-Set QCA."
Carsten Q. Schneider and Ingo Rohlfing, Kozep-Europai Egyetem: "It's Complex! Combining Qualitative Comparative Analysis (QCA) and Case Studies in Multi-Method Research."
Discussant: Andrew Yeo, Catholic University of America

History, Narrative, and Process Tracing

Chair: Jens Borchert, University of Frankfurt
Participants:
Derek Beach and Rasmus Pedersen, University of Aarhus: "Observing a Causal Mehanism with Process-Tracing Methods: A Mechanistic Approach."
Daniel HoSang, University of Oregon: "The Interdependence of Ideas and Institutions: Tracing the Emergence of Proposition 187 and the Immigration Debate, 1965–1994."
Peter Starke, Universitaet Bremen: "Anecdotal Evidence: The Role of Narrative in Process Tracing."
Cory McCruden, Yale University: "The Use and Misuse of Oral History in Political Science."
Pascal Vennesson, European University Institute: "Process Tracing in Action: Bridging the Positivist-Interpretivist Divide?"
Discussant: Hillel David Soifer, Temple University

Theorizing Institutional Change

Chair: Kathleen Thelen, Massachusetts Institute of Technology
Participants:
Jacob S. Hacker and Paul Pierson, Yale University: "Drift as a Mechanism of Institutional Change."
Peter A. Hall, Harvard University: "The Political Origins of our Economic Discontents: From Keynesianism to Neo-Liberalism and Beyond."
James Mahoney and Kathleen Thelen, Northwestern University: "Explaining Gradual Institutional Change."
Discussant: Julia Lynch, University of Pennsylvania

Politics, Culture, and Change

Chair: Dara Z. Strolovitch, University of Minnesota, Twin Cities
Participants:
Kimberley S. Johnson, Barnard College: "The League of Women Voters, Southern Politics, and the Civil Rights Movement."
Julie L. Novkov, State University of New York, Albany: "Sacrifice and Civic Membership: The War on Terror."
Patricia Strach and Kathleen Sullivan, Harvard University: "Public-Private Partnerships in Institution Building: Municipal Garbage Collection."
Priscilla Yamin, University of Oregon: "Sixties Culture Wars, Loving and the Politics of Marriage, Race and Gender."
Carol Nackenoff and Kathleen Sullivan, Swarthmore College: "Municipal and County Courts as Sites of Political Development."
Discussant: Kristi Andersen, Syracuse University

Comparative Approaches to Political Change: Ethnography Unveiled?

Chair: Atul Kohli, Princeton University
Participants:
Deborah L. Wheeler, United States Naval Academy: "Information (without) Revolution? Ethnography and the Study of New Media Enabled Change in the Middle East."
Carrie Rosefsky Wickham, Emory University: "The Dynamics of Islamist Movement Change: Revisiting the Participation-Moderation Thesis."
Lloyd I. Rudolph, University of Chicago: "Caste and Politics in India: The Traditional Roots of Associational Life."
Susanne Hoeber Rudolph, University of Chicago: "Gandhi's Leadership: The Traditional Roots of Charisma."
Discussant: Rebecca Bill Chavez, United States Naval Academy

The Return of Cross-Regional Qualitative Comparative Analysis (QCA)

Chair: Sidney Tarrow, Cornell University
Participants:
Rudra Sil, University of Pennsylvania: "Between Area Studies and Universal Models: Cross-Regional Comparative Analysis (CRCA) and Middle-Range Theory-Building."
Amel F. Ahmed, University of Massachusetts, Amherst: "Beyond Exceptionalism: Cross-Regional Analysis of American and European Democratization."
Cheng Chen, SUNY, Albany: "'Authoritarian Capitalism' in Post-Communist Russia and China: Are We Comparing Apples and Oranges?"
Marcus Kreuzer, Villanova University: "Old Theories, New Contexts: Cross-Regional Analysis and Theory Development."
Jefferey M. Sellers, University of Southern California: "Beyond Subnational Comparison: The Transnational Comparative Method."
Discussant: Sidney Tarrow, Cornell University

The Methods Café

Chairs: Adam Avrushin, University of Chicago; Debra Thomas, University of Toronto
Participants:
Rudra Sil, University of Pennsylvania: "Analytic Electicism: Beyond the Micro-Meso-Macro and Ideal-Material Divides."
Joe Lowndes, University of Oregon: "APD: Methodological Issues."
Emily Hauptmann, Western Michigan University: "Archival research."

Robert Adcock, George Washington University and Douglas C. Dow, University of Texas, Dallas: "Concepts: Formation, History, and Analysis."

Raymond Duvall, University of Minnesota and Lisa Wedeen, University of Chicago: "Critical Constructivist and Discourse Analyses."

Cecelia Lynch, University of California, Irvine: "Critical Studies of Ethics and Agents."

Mary Hawkesworth, Rutgers University: "Feminist Methods."

Timothy Pachirat, The New School for Social Research and Katherine Cramer Walsh, University of Wisconsin, Madison: "Field Research: US (Participant Observation, Political Ethnography, etc.)"

Ron Schmidt, California State University, Long Beach and Dvora Yanow, Vrije Universiteit, Amsterdam: "Interpretive Policy Analysis: Value-Critical, Policy Discourse, Policy Spaces."

Nancy Wadsworth, University of Denver: "Intersectionality Research: Race, Ethnicity, Gender, Class, Sexuality, Religion."

Lee Ann Fujii, George Washington University and Frederic Charles Schaffer, University of Massachusetts, Amherst: "Interviewing: Ordinary Language Interviewing and Life History Narratives."

Samer Shehata, Georgetown University and Jan Kubik, Rutgers University: "Political Ethnography: Field Research 'Overseas.'"

Kevin Bruyneel, Babson College and Edmund Fong, University of Utah: "Post-Colonial Analysis."

Pamela Brandwein, University of Michigan and Julie Novkov, State University of New York, Albany: "Public Law."

Patrick Jackson, American University: "Scientific Methods: The Science Question in Political Science Research Design."

Peregrine Schwartz-Shea, University of Utah and Elisabeth Wood, Yale University: "Teaching Qualitative-Interpretive Methods."

Lloyd Rudolph, University of Chicago: "Validating Subjective Knowledge: Personal Documents in Political Analysis."

Qualitative and Multi-Method Research
School of Government and Public Policy
University of Arizona
315 Social Sciences Building
P.O. Box 210027
Tucson, AZ 85721-0027
USA

Nonprofit Org.
U.S. Postage
PAID
TUCSON, AZ
Permit No. 271

Qualitative and Multi-Method Research (ISSN 2153-6767) is edited by Gary Goertz (tel: 520-621-1346, fax: 520-621-5051, email: ggoertz@u.arizona.edu). The assistant editor is Joshua C. Yesnowitz (email: jcyesnow@bu.edu). Published with financial assistance from the Consortium for Qualitative Research Methods (CQRM). Opinions do not represent the official position of CQRM. After a one-year lag, past issues will be available to the general public online, free of charge, at http://www.maxwell.syr.edu/moynihan/programs/cqrm/section.html. Annual section dues are $8.00. You may join the section online (http://www.apsanet.org) or by phone (202-483-2512). Changes of address take place automatically when members change their addresses with APSA. Please do not send change-of-address information to the newsletter.