Fall 2012, Vol. 10, No. 2

Qualitative & Multi-Method Research

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

Contents

Letter from the Section President

Gary Goertz
University of Notre Dame
ggoertz@nd.edu

With APSA being canceled *in extremis*, many of us were able to enjoy an unusually relaxed Labor Day weekend. However, it was very unfortunate that all the efforts that went into the preparation of panels, papers, and other events were not realized at the APSA meetings. I especially want to take this opportunity to thank several people for their leadership and hard work. David Waldner was the section's Division Chair, and organized a first-rate lineup of panels and roundtables. David Collier, Jay Seawright, Thad Dunning, Sherry Zaks, Diana Kapiszewski, and Naomi Levy organized and prepared to teach short courses that were cancelled at very short notice.

Absent the hurricane, we would of course have held our business meeting in New Orleans, and presented the section's various awards. I congratulate those who won awards this year, the citations for which can be found in the announcements at the end of this newsletter.

You received an email regarding section business and elections, which normally would also have been conducted at the business meeting, but which we have had to conduct electronically. I thank all those who have accepted to serve in various capacities for the section. Below we give the new and continuing set of officers, while members of the 2013 nominating and award committees are listed in the announcements at the end of this newsletter.

Over the years, while we have always made previous issues of the newsletter available online to everyone, we have reserved the two most recent issues for current section members. In the past, this meant a delay of about a year before the electronic versions of recent issues were posted. We are pleased to report that, thanks to APSA Connect, section members can now access contemporary issues by logging on to the Association's website.

Notwithstanding the difficulties surrounding the annual meeting, this is a promising time for qualitative and multi-method research. We now appear to be in the middle of a new and very substantial wave of writing. For example, the Cambridge University Press series "Strategies for Social Analysis" has four books in print and two more in press, and there are further volumes just out or forthcoming from Princeton, Michigan, Palgrave-Macmillan, and Routledge. These books (together with articles that have appeared in top journals) are improving and enriching how we teach and do qualitative research. Many of these books treat core topics such as multi-method research, case studies, and process tracing. Our section has been the venue for developing and road testing much of this work, with chapters and pieces of the volumes being presented at APSA panels, and seed ideas appearing in the newsletter. I want to express my deep appreciation for this new work, and I hope that it is an indication that there is much more to come!

Symposium: Peregrine Schwartz-Shea and Dvora Yanow's *Interpretive Research Design: Concepts and Processes* (Routledge Press, 2012)

Filling the (Interpretive Methods) Gap: Assessing the Contribution of Schwartz-Shea and Yanow for Graduate Students

Tanya B. Schwarz
University of California, Irvine
tschwarz@uci.edu

Peregrine Schwartz-Shea and Dvora Yanow wrote *Interpretive Research Design: Concepts and Processes* with three readers in mind: the graduate student, the experienced researcher, and the teacher of interpretive methods (p. 3). As a graduate student whose work falls under the interpretive umbrella, I am in the unique position to offer an "insider's" perspective on the accessibility and utility of this volume for the first group. To that end, my goal in the following is to articulate the strengths and weaknesses of this volume not only for graduate students with an interpretivist bent—including those who are either in the process of formulating a research project, applying for funding, or in a position in which they must defend their work to faculty and evaluators who may or may not be familiar with interpretive approaches—but also for the greater graduate student community.

As a graduate student first at the University of California, Riverside, and currently at the University of California, Irvine, and as a past participant of the Institute for Qualitative and Multi-Method Research, I have been exposed to top-notch training in a wide variety of methods and methodologies, including those that scholars often refer to as "interpretive." However, as evidenced in my discussions with graduate students from other programs, experiences like mine are uncommon. While most political science graduate programs in the United States offer (and often require) methods training, these courses, both in number and in focus, tend to privilege certain methods and methodologies over others, usually advocating positivism over its interpretive counterpart (or, alternatively, may claim to offer instruction in "mixed" methods, yet overemphasize positivist perspectives). Interpretive approaches are skimmed over or neglected altogether, thereby reducing their representation in the field. As such, Schwartz-Shea and Yanow's volume serves as a much-needed resource for those graduate students who lack readily available access to interpretive methods instruction at their home universities, as it provides these students with a clear and concise manual for addressing many issues that scholars engaged in interpretive research begin to encounter at the earliest stages of their careers.

In this book, Schwartz-Shea and Yanow present, in detail, the considerations that need to be addressed when formulating an interpretive research project. While they briefly revisit the step-by-step organization of a typical research design, this is not their primary focus. Instead, Schwartz-Shea and Yanow delve into the way conceptualizations of interpretive research designs differ from positivist models. They begin this enterprise with a discussion about the articulation of a research question, highlighting the importance of abductive reasoning and the hermeneutic circle for interpretive research (pp. 30–33)—crucial information for students who learn about the inductive/deductive distinction in their earliest methods seminars, but are not exposed to the possibility of a third option. Schwartz-Shea and Yanow then address differences in concept formation, theorizing, and notions of causality, noting that, unlike research originating from the school of positivist thought, in interpretive research, concepts are not formed *a priori*, but rather emerge organically through the course of the research process (p. 39). Moreover, theorizing is understood to occur based on knowledge arising out of specific historical and cultural settings (p. 47), and causality is perceived as constitutive (p. 52). Establishing these distinctions and elaborating on their onto-epistemological foundations is important for junior scholars, not only in the formation of their research, but also, and perhaps more crucially, during the justification phase of these projects, as concept formation, theory, and causality are often some of the most contentious points of discussion when preparing a prospectus or grant proposal. Students engaging in interpretive work can draw upon the language and arguments put forth by Schwartz-Shea and Yanow to articulate a strong defense of their own research approaches. Finally, although the authors assert that philosophy of science is not their primary focus, they do summarize and synthesize philosophy of science debates in concise and easy-to-understand language, thus making this book a much-needed reference for students—interpretivists and non-interpretivists alike—who have a difficult time understanding these debates or are unaware of their existence altogether.

Throughout the book, Schwartz-Shea and Yanow also elaborate on the importance of reflexivity and ethics. In my experience, the topic of ethics is not widely discussed in political science seminars outside of the usual concerns regarding the protection of human subjects in social science research (something usually addressed through institutional review boards). Here, Schwartz-Shea and Yanow broaden the sphere of discussion to highlight other ethical issues that a graduate student should be aware of, including the constitutive nature of research. Understanding the role of the researcher in data generation is a key lesson for any graduate student to learn early on, as it serves to make the researcher more reflexive during the course of formulating and conducting future research endeavors. Moreover, Schwartz-Shea and Yanow's elaboration on the problematic aspects of the notion of "objectivity" is useful for the interpretivist student who must justify interpretive research to scholars who rely on and expect objectivity in any political science project.

Schwartz-Shea and Yanow also provide detailed discussions outlining differences in terms often used in political science to describe processes and/or evaluate research projects. For instance, they distinguish between interpretive and positivist data-generating processes (i.e., exposure vs. sampling, intertextuality vs. triangulation, etc., pp. 86–88), analytic standards (i.e., hermeneutic sensibility vs. falsifiability, pp. 107–108), and evaluative standards (i.e., trustworthiness vs. validity, systematicity vs. rigor, and reflexivity and transparency vs. objectivity, pp. 91–103). Addressing these terms and highlighting their distinguishing features in a clear and concise way provides crucial and easily accessible information for formulating a research design.

For these and other reasons, Schwartz-Shea and Yanow's volume should be kept on the desk of every graduate student as a useful reference tool. For those who are uninterested in interpretive research, the book remains useful for articulating key methodological debates in easy-to-understand language. For those wanting to engage in interpretive research, this book provides important ways to conceptualize and shape future projects. Schwartz-Shea and Yanow fill a much-needed gap in the methods literature—providing an elucidating overview of interpretive approaches, their underlying presuppositions, and the related evaluative standards. That being said, while the issues addressed in the book will undoubtedly assist junior scholars—whether in the formulation of a research project or the articulation of a proper defense of one's work—significant hurdles remain in the broader political science community.

One challenge emerges with the writing of grant proposals. As Schwartz-Shea and Yanow point out, "communication may be stymied" when proposal writers and reviewers come from different epistemic communities with different onto-epistemological presuppositions (p. 136). Schwartz-Shea and Yanow acknowledge this problem and recommend that (1) proposal writers take care in communicating the goals of their research, and (2) proposal reviewers be aware of the differing evaluative standards required by interpretive approaches (p.137). Yet, while proposal writers, positivists and interpretivists alike, should be clear in articulating the foundations and goals of their research, this can be problematic if the reviewer is unfamiliar with the writer's methodological approach. How, then, can a proposal writer articulate her research in a way deemed acceptable by the reviewer in a proposal that is constrained by maximum length requirements, thereby preventing the applicant from fully fleshing out the assumptions and foundations of her work?

For interdisciplinary grants that are often reviewed by scholars based in the humanities or other interpretivist-friendly disciplines (anthropology, for example), the interpretivist political scientist may not need to adjust her language to be understood. In fact, in these cases, the interpretivist researcher may have an advantage over other positivist grants pursuers in that her approach would undoubtedly be familiar to the reviewers. However, a different situation emerges when the reviewers are primarily trained in and supportive of strictly positivist approaches. To ameliorate this issue, Schwartz-Shea and Yanow introduce alternative concepts and standards that may be used in place of their positivist counterpoints.

As noted above, Schwartz-Shea and Yanow discuss why certain terms and concepts often used in political science research are not suitable for interpretivist designs (at least with the current dominant meanings attached to these terms). Yet, due to the neglect of interpretive methods training in the political science community in the U.S., the alternative concepts and terminologies the authors introduce are not widely known. This presents a difficulty for grant applicants, as using unfamiliar terms is not unlike writing a proposal in a language unknown to the reviewer. Interpretivist graduate students are aware of this problem, and, consequently, may try to straddle a fine line in which the interpretive aspects of their projects are laid out, but features commonly expected from positivist scholars, including notions of "case studies," or even "hypotheses" are included. It could be argued that the inclusion of these terms actually weakens the proposal, as they are not properly suited to the project (p. 70). However, neglecting these common terms completely may leave the proposal writer in the dangerous position in which her research is perceived as not conforming to common standards of scientific study. This leaves the interpretivist scholar in a difficult position. Should she then focus only on those grants in which her methods "language" is more likely to be understood and valued? Or should she try to "translate" her project's assumptions and goals for the positivist reviewer, while running the risk of weakening her position? Neither option is ideal. While Schwartz-Shea and Yanow's book may help to alleviate this issue by educating non-interpretivist grant reviewers, how likely is it that non-interpretivists will pick up this volume to begin with?

Similar issues to the above may also arise when defending one's research to an academic advisor or dissertation committee. Graduate students are often in a precarious position with their faculty mentors, where power dynamics are key. As is the nature of the institution, students rely heavily on their advisors to support them in their endeavors, and, as such, often feel the need to bend to the will of their superiors, even if this is not always in the best interests of the graduate student. And, while many senior faculty may be more aware of the range of methodological approaches available to political scientists, junior faculty newly emerging out of graduate programs that focused largely (or solely) on positivist approaches may be perplexed by the proposed approach of the interpretivist graduate student. Schwartz-Shea and Yanow's book is important in that it provides an easily accessible and functional educational tool, which, if widely adopted, is capable of ameliorating the above issues. However, I fear that the general state of the political science discipline (especially within the U.S.), and, in particular, the trend to selectively educate students in positivist methods and methodologies to the exclusion of other approaches, may stymie Schwartz-Shea and Yanow's larger project.

It is with regard to this larger issue that Schwartz-Shea and Yanow may have missed an opportunity to make a larger impact in the methods/methodologies debates. The contribution of this latest volume is crucial for instruction in interpretive research. However, rather than openly renouncing the methodological hegemony of positivist methods (and the relationships and power structures undergirding this dominant discourse) Schwartz-Shea and Yanow try to maintain a neutral stance. Understandably, they are trying to educate, while still speaking within and across divides, in a way that is non-threatening. Yet, while tolerance and dialogue are important tools in any contentious discussion, I wonder if a stronger approach is needed in order to provide the necessary support to graduate students engaging in work that, while valuable and important, may not conform to mainstream evaluative standards (this may also include work that is not categorized as "interpretive" but still lies outside of the dominant approaches). After all, the argument can (and should) be made that these alternative methods and methodologies are just as essential to knowledge production and deserve an equal place at the table. And while attention to a broader range of instructive works, including *Interpretive Research Design: Concepts and Processes*, is important for moving in this direction, interpretive methods and methodologies will remain on the sidelines, and graduate students (and other scholars) who engage in interpretive research will continue to face difficulties, unless the broader discourse about "acceptable" methods and methodologies becomes more inclusive.

Speaking Across Epistemic Communities: The Promotion of the Interpretive Project

Corinne Heaven
University of Reading, United Kingdom
c.heaven@pgr.reading.ac.uk

Graduate students are one of the readerships that Schwartz-Shea and Yanow had in mind when writing *Interpretive Research Design: Concepts and Processes*. They "in particular need information about interpretive concepts and processes so that they can do empirical research that genuinely allows for an interpretive approach without having their confidence undermined at this stage of the game by uninformed critiques" (p. 2). Presenting this information in accessible, engaging, and jargon-free writing, Schwartz-Shea and Yanow identify the key differences between interpretive and positivist research designs. The authors make the convincing case for interpretive work that comes in addition to other ways of carrying out research in the social sciences. In doing so, Schwartz-Shea and Yanow draw out the reasoning behind interpretive research design and the "rhythm" that this logic of inquiry will take. The core of the book is devoted to the latter aspect, working out the possible avenues of generating and analyzing data as well as developing criteria to evaluate knowledge claims made by interpretive researchers. Drawing on various practical examples, the authors aptly demonstrate the strengths (and limitations) of interpretive work, making the book an essential reading for those interested in approaches that challenge the 'traditional' analysis and practice of social science.

Graduate students often find confusing the forest of terms for methodological approaches and ponder about the ways of designing their methodology, explaining their choice of methods as well as generating and analysing data. I, too, turned to *Interpretive Research Design* for guidance and inspiration. The book presents a fresh and innovative perspective on questions that unnerve every graduate student. Writing from this viewpoint, I make three core arguments in assessing the book.

First, reading *Interpretive Research Design* is in itself an exercise in reflexivity. The book helps to take a reflexive approach toward the choices we make when designing an interpretive research framework. It draws attention to the different ways of knowledge production and our own involvement as the creators of that knowledge. It is a practice of reflexivity because it reminds us of the reasons for choosing an interpretive approach, which neither follows automatically from the setting nor is a one-size-fits-all approach to every possible research problem. And importantly, this exercise helps to position ourselves within that epistemic community by identifying the audience we are addressing with our work. Second, the practical consequences for graduate students in designing their work are less straightforward. This is not a "how-to guide" (p. 9), and the level of generality that is applied when contrasting interpretive with positivist research methodology may lead to renewed confusion—a point I turn to in more detail below.

Put differently, for graduate students, the book is an excellent starting point to explore the world of interpretive research, but once one has set the scene for a research project, more specific debates are needed. Third, given this background I want to conclude by arguing that this contribution is more likely to have a strategic impact on the wider discipline by furthering the recognition for interpretive work across different epistemic communities as opposed to a practical guide for graduate students and scholars within the interpretive community.

A Practice in Reflexivity

If the process of writing a PhD thesis were a journey, the most helpful travel book I have read so far is *Interpretive Research Design*. Schwartz-Shea and Yanow are excellent tour guides. For example, many graduate students begin their project with the literature review, though it often seems a mystery as how to actually write one. How can one possibly decide whose work to include? Now, this no longer remains a mystery. Schwartz-Shea and Yanow suggest thinking about a literature review as a dinner party to which one invites the leading thinkers of different communities to join the conversation about the research question at hand. Who would be invited? What would they contribute to the question? Whose voice would be missing? It is here that the infamous "gap in the literature" reveals itself—or, to be precise, is created.

A second aspect then concerns the actual process of writing up the subsequent chapters. Whilst reading the book, it suddenly struck me that I was trying to comply with the "hourglass model of traditional research manuscripts" (p. 127), following a fixed model of writing one chapter after the other. Prior to reading the book, I was unaware that I was trying to apply a linear style of working to a research framework that lends itself to a recursive fashion. Schwartz-Shea and Yanow make it very clear that the process of interpretive research differs from the model of positivist research. This not only concerns the obvious: Yes, we do not speak of variables, operationalization, or sampling. But we also work in a circular fashion, starting from what the authors call a puzzle, moving back and forth between theoretical conceptions and empirical material. Interpretive research asks different questions, provides different answers, and, as a result, the entire research process looks and feels different.

Third, Schwartz-Shea and Yanow emphasize that the criteria for evaluating positivist research design simply do not apply to interpretive logics of inquiry. This is a crucial point, especially for those graduate students who are socialized in institutions that are characterized by a predominantly positivist research community. Standards of generalizability will not apply to an interpretive project, for example. Quite the contrary; it should be evaluated whether it is context-specific enough. An interpretive project will not test the accuracy of concepts, but will explore how concepts are used in the field. In other words, the character of interpretive research is distinctively different. For graduate students in particular, it is important to think about the choice they make when joining a specific epistemic community, since the research project will signal where one is planning to position her- or himself. Why, for example, does one choose to carry out a narrative analysis as opposed to a causal explanation? This is not always self-explanatory, and the more one reflects upon these questions, I believe, the more one can further the interpretive research agenda as a whole.

In sum, the book offers orientation and helps not only to position ourselves, but the interpretive community in the wider discipline. It is therefore an invaluable contribution that will be of interest for both interpretive and positivist researchers.

Limitations of *Interpretive Research Design*

Unlike other contributions (see for example Wagenaar 2011, Alvesson and Sköldberg 2009), the authors dedicate large parts of the book to positivist methodology, contrasting it with the logic of interpretive inquiry. A column introduced by the authors summarizes this comparison by looking at the following criteria: research orientation (meaning-making vs. measurement), design attitude (abductive logic of inquiry vs. deductive and inductive logics of inquiry), the process of inquiry (hermeneutical circle-spiral vs. hypothesis testing; bottom-up, *in situ* concept development vs. sampling), the analysis of evidence (hermeneutic sensibility vs. falsifiability) and evaluative standards (trustworthiness vs. validity, reliability, replicability) (p. 113). It is this aspect of the book that is perhaps the most striking one. Why would one go into the details of the positivist logic of inquiry when writing a book about interpretive research design? On a practical level, one might expect graduate students to use the column as a manual to respond to potential reservations raised by "uninformed critics" by pointing out the different standards of positivist and interpretive research. Or, one might argue that this, too, is a practice of reflexivity by the authors.

But of course, the clear-cut comparison between positivist and interpretive research presupposes a level of generality that might cause confusion for the researcher who has already established ideas about the design of a project. I would have liked the authors to be clearer on what is to be understood by interpretive research. That is not to say that I would have liked Schwartz-Shea and Yanow to provide an authoritative definition of the interpretive community (as this would, from a philosophical standpoint, run against their very own project). The approach they chose is clearly very inclusive and dynamic. However, while they are very strong in showing the differences to positivist research design, their own philosophical roots and ontological and epistemological commitments are somewhat lurking beneath the surface (they briefly point out that their work is informed by science studies or the sociology of knowledge perspective). For Schwartz-Shea and Yanow then, the common aim of interpretive research design is "a focus on meaning-making and production of contextualized knowledge." While most "interpretive" graduate students and scholars would perhaps agree with that claim, it nevertheless remains abstract and general. How does one achieve this goal? There are, of course, numerous ways of doing so and the book does not (and—due to its length—cannot) address these in a detailed manner, if nothing else because designing a research framework is something that has to be practiced, not theorized.

On a day-to-day basis, this level of generality causes some limitations concerning the practicality of the book. The authors make very clear the purpose of interpretive research, but not so much *how* one should go about conducting this style of research. Put differently, the book shows what kind of questions interpretive scholars can ask and what kind of answers one can expect, but not so much *how* one will form these questions and design the answers. Let me explain this: Reading the book from a Foucauldian perspective, I sometimes found myself at odds with certain concepts, agency and subjectivity being amongst them. "Foucauldians" view the latter as both an object and subject of knowledge and power, while in *Interpretive Research Design* the subject seems to resemble a foundational figure. Without going into details, this exemplifies that conceptions around the autonomy or consciousness of subjects differ among interpretive researchers. The interpretive community is less homogenous than portrayed in the book, and it might have been worthwhile to address this pluralism in more detail. This is relevant not only for practical considerations, but also in order for researchers with different backgrounds to identify themselves with the interpretive project. Some scholars will want to identify themselves with a community and will find the ambiguity of not knowing where one belongs unsettling. Our terminology, our methodology, our research practices all signal a specific outlook and show others on which ontological and epistemological levels we operate and to whom we want to communicate our research.

Yet, and this goes without saying, the book is not directed at Foucauldian scholars in particular, a community that itself might be described as multi- rather than interdisciplinary (Walters 2012). Nor is it directed to any other "sub-community." Rather, the book addresses the interpretive community in general. This making of a somewhat uniform interpretive community, is, in fact, a powerful move. It might be here that the success of the book lies, as the way Schwartz-Shea and Yanow define interpretive vis-à-vis positivist research may have significant consequences for future practice.

Conclusion: Positioning Interpretive Research Design

The approach Schwarz-Shea and Yanow take by creating a clear dichotomy between an interpretive and a positivist logic of inquiry is, as noted above, not without problems. For a start, there is considerable variation within the interpretive research community, and perhaps less interaction than we would often wish for. But choosing the label "interpretive" for a wide range of approaches might be an effective strategic choice. As the authors note, "Our purpose is to help those from diverse research communities *recognize* interpretive research as a distinctive logic of inquiry and to *develop* what this means at the design stage" (p. 13). Well, one might ask whether interpretive research requires this kind of promotion. Do we still need to pave the way for interpretive work in the social sciences? Does this run the danger of undermining a community that is already well established and growing? Now, depending on one's own position, the answer to this question will obviously differ. The interpretive project may be more or less recognized within different epistemic communities, different institutions, or different regions. As someone writing with a background influenced by the "Franco-German critical or interpretive tradition," I would have appreciated drawing more attention to the variety in the community. Someone from a different reception context might argue that the emphasis should lie on the interpretive project as a whole vis-à-vis the positivist research agenda, despite its heterogeneity.

So, if interpretive work requires promotion and positioning in the discipline, then *Interpretive Research Design* serves this purpose well. After all, the authors invite the reader to explore this community more closely after reading the book, hoping that their contribution will encourage efforts to gain "a greater understanding from the inside" (p. 139) of interpretive research. I wonder whether the authors therefore primarily have in mind a reader who is looking from the "outside." And indeed, anyone who is reading from the outside will recognise interpretive research as a distinct scientific practice that contributes to important plurality in the social sciences.

In sum, Schwartz-Shea and Yanow's book is an invaluable contribution to the discipline. In a clear and concise manner, the authors navigate the deep and sometimes stormy waters of speaking *across* epistemic communities. And most importantly, they set the course for continued debate about the range of logics of inquiry in the social sciences.

References

Alvesson, Mats and Kaj Sköldberg. 2009. *Reflexive Methodology: New Vistas for Qualitative Research*. London: Sage Publications.

Wagenaar, Hendrik. 2011. *Meaning in Action: Interpretation and Dialogue in Policy Analysis*. Armonk, NY: M.E. Sharpe.

Walters, William. 2012. *Governmentality: Critical Encounters*. London: Routledge.

The Need to Teach Rather than Translate

Katherine Cramer Walsh
University of Wisconsin-Madison
kwalsh2@wisc.edu

Reading and digesting *Interpretive Research Design* has been an enormously valuable experience in my professional development. I found it reassuring, encouraging, and also challenging. Reading it was like discovering the lost tribe of which I am a part. It gave voice to many of the struggles I have experienced as a political scientist who tends toward interpretive methods. And yet I now find myself torn over whether to fully embrace the language it suggests, as opposed to continuing to use the terms commonly used by the broader discipline with which I try to communicate.

I study American public opinion, and my main method of doing so is not through positivist analysis of opinion poll data but interpretive analysis of observations of conversations. My methods and my methodology (see Schwartz-Shea and Yanow 2012: 4) are unusual in my field. I view that as an asset, but it is a challenge, as this book so clearly lays out.

The main challenge is primarily one of communication.

Any good communicator knows that she has to put her message in terms that her audience will understand. To be persuasive, not only does the message need to be understandable; it needs to resonate. When I communicate my research to the political science community I often expect that I have to translate my words into a learned language, rather than the language that is most natural to me. I have made myself use the terms that my colleagues expect to see: dependent and independent variables, sampling, bias, and generalizability, for example. These are not always the terms that best fit my research methodology, as Schwartz-Shea and Yanow explain in detail.

In most of my work, I attempt to understand how people make sense of politics. I examine what perspectives particular people bring to bear as they try to interpret the political world, and how they do so. When pressed, I have stated that my dependent variable in such a study is the understandings people develop for themselves. But I have come to terms over the years with the fact that that terminology is not quite right. I am usually not so interested in what independent variables predict those understandings, but rather in how people put different considerations together to arrive at an understanding. The language of variables does not always allow me to convey an accurate meaning of my work. My focus is not on how varying one factor produces a different understanding, but on how people collectively create understandings in particular contexts.

Interpretive Research Design calls into question whether I still too often shove explanations of my methodology into words that do not quite fit. Take for example, sampling. Schwartz-Shea and Yanow argue that in an interpretivist project, using "sampling" to describe our selection of cases is not as accurate as "exposure." They encourage researchers doing interpretivist work to select situations, experiences, people, etc., "to maximize research-relevant variety in the researcher's exposure to different understandings of what is being studied" (p. 85). I normally label this sampling, but this book suggests doing so is not consistent with the methodological logic of an interpretivist approach. Sampling "signals researcher control over the selection process, an implication that often does not hold for interpretive research settings" (p. 87).

I am torn. I want my work to be read and understood by others in my field, the vast majority of whom do not study public opinion in an interpretivist manner. And in my current work, I do exercise control over which conversations I study, even though I do not recruit the participants. I "sample" the communities, and then seek advice on which local venues might have a group of regulars whom I can gain access to. So I tend toward the language of sampling.

This book acknowledges that of course researchers want to be taken seriously and be respected and thus strive to communicate their work in widely understood ways. But it argues that forcing an approach into terms that do not fit means the work will be evaluated according to criteria that do not fit either. This, in the end, disadvantages the work: It reduces the chances that it will be funded, and that it will ever be conducted, much less communicated to other scholars. I plan to continue to use the language of sampling, but on the advice of this book, I will be more explicit about what that means for an interpretivist study.

The book presents an even stronger challenge to my use of the term "generalizability" in the context of interpretivist research. The normative understanding of social science is that its purpose is to generalize to a broader population. Following the logic of statistics, we study a part in order to say something about the broader whole. If we can't say something about the broader world, then what is the point?

I have addressed this question in the past by arguing that the purpose of my interpretive work is not to generalize to a population but to generalize to theory—that I hope to contribute to the literature by helping to develop the theories that are then subsequently tested via positivist approaches. This book challenges that translation as well. On page 47, the authors write that in contrast to the question "'Are the results generalizable?'…members of interpretive research communities ask: 'Is the research sufficiently contextualized so that the interpretations are embedded in, rather than abstracted from, the settings of the actors studied?'" This answer is not really about generalizability at all.

I anticipate that if I claim that as the criterion by which my work should be judged, colleagues will ask, "Where does that get us? What good does it do to understand the way a particular set of people in a particular place understand politics?" My answer is twofold: First, most if not all of our research actually studies something particular—a particular moment in time, if not a particular place or group of people. Second, if the overarching goal of the discipline is to generalize to broad populations, how can we reasonably do so if we do not know much about the microprocesses that go on within them? That latter answer is what brings me back to the language of generalizability. I see my work as contributing to the project of generalizability even if my claims are not about the broader population.

I am mindful that presenting my work this way may subjugate it to "helper," not "leader," status in the discipline in some eyes. But I disagree. I do not think it is possible to make accurate generalizable claims if the understandings do not accurately reflect what goes on in particular lives.

Interpretive Research Design has also made me question my use of the term "bias." One of the most common questions I get from reviewers or from audience members when I present my work is whether my presence influenced the conversations I observed. I tend to answer these questions in terms of the positivist language of "bias," since that is the language through which I am typically posed the question. I usually respond by talking about how I make attempts to gauge how my presence alters the conversation (such as listening to a conversation behind me or paying attention to what my recorder picks up while I am briefly out of the room). But here again, the Schwartz-Shea and Yanow book is instructive: Why convey the role of my presence in the language of bias when the relevant language is really one of reflexivity (ch. 6)? Why engage this as a question of *whether* my presence biases what I observe, when it is really a question of *how* my presence alters the context and therefore what takes place within it? Of course my presence alters the conversations. That is a given, and oftentimes a gift (e.g., Schwartz-Shea and Yanow 2012: 110).

Perhaps the greatest challenge that this book presented to me is the caution against striving for mixed methods as the sign of the ultimate research design. In recent years, I have been deeply grateful to have many scholars express support for my approach to the study of public opinion. They have also repeatedly encouraged me, however, to more closely combine my interpretive work with positivist analysis of survey data. I have agreed with that advice completely—until reading this book. The final chapter raises the possibility that it might be quite illogical to combine quantitative analysis of poll results with interpretive analysis of observation of conversations. If the two approaches are based on two very different logics, then does one really cover for the other's weaknesses?

The answer I am sticking with for now comes on page 134. The authors state that combining approaches makes sense if "mix[ing] methodologies within a single research topic, …[as opposed to] mix[ing] methodologies within a single research question." If my main focus is on the way people in particular communities interpret public affairs, then I use an interpretive analysis of conversations in those communities. When I notice the use of a perspective across a variety of communities, I become curious about whether or not that perspective is as widespread as it seems from the set of cases I have sampled (or exposed myself to). So then I turn to poll data to test this hypothesis. But at that point, the question I am pursuing is slightly different from the one that generated the research.

This book acknowledges all of these dilemmas and challenges, and indeed seems to be motivated in part by them. The authors seem to be saying, "Buck up. It may not be fair that you have to do this, but in order to have your work evaluated fairly, you need to expend some energy explaining and justifying your methodology and providing the criteria—the language—by which it should be judged, rather than forcing your work into a language that doesn't fit."

In other words, the main lesson I take away from this book is that writing a proposal for interpretivist work is not so much about translation as it is about teaching. Schwartz-Shea and Yanow urge us to teach to others the value of this kind of work and how it ought to be judged. The book conveys a desire for interpretivism to not take on the language of another methodology, but for all methodologies to jointly contribute to the underlying goal of social science: a better understanding of the world in which we live through a "systematic" approach, "conducted with an attitude of doubt" (p. 17).

Interpretive and Experimental Thinking in the Social Sciences

Bernhard Kittel
University of Vienna, Austria
bernhard.kittel@univie.ac.at

In their introduction to interpretive research design, which programmatically opens a new series on interpretive methods at Routledge, Peregrine Schwartz-Shea and Dvora Yanow (henceforth SSY) have laid out an impressive landscape of the fundamental ideas guiding research inspired by the interpretive paradigm. Both authors are well-established scholars working from that perspective, and have contributed much to the advancement of the research community's understanding of interpretive work, not least through their previous joint methodological project, the co-edited volume *Interpretation and Method: Empirical Research Methods and the Interpretive Turn* (2006). Their new textbook builds on the previous volume, but also demonstrates a remarkable development in both interpretive methodology and the debate between representatives of interpretivism and positivism in recent years. The tone of the exposition is both self confident and critically reflective, clearly marking the distinctive features of interpretive research. But in contrast to past times, the references to positivist work are characterized by a more conciliatory mood and the description of contrasts is based on a clearly manifest interest in understanding the "other way of knowing" (Moses and Knutsen 2012).[1]

The methodology debate in political science has matured much since King, Keohane, and Verba's (1994) attempt to understand all empirical social science research through the lenses of a quantitative template. While their argument was phrased in terms alien to interpretivists, it triggered a response by qualitative scholars adhering to the positivist paradigm (Brady and Collier 2004; Gerring 2012), which, in turn, made clear that the divide between qualitative and quantitative methods does not overlap with positivist and interpretive methodologies. A consequence of this response, however, was the marginalization of interpretive research in the United States, though much less so in Europe.

With this textbook, interpretivists do not need external interpreters anymore to make themselves heard. Most clearly, this position of self confidence and conciliation can be observed in the chapters on quality standards. For example, SSY explain the problems that traditional quality criteria pose to the interpretive perspective, and why these problems have deeper implications than other perspectives seem to recognize. However, they accept the salience of the intention underlying these criteria, basically ascertaining the trustworthiness of claims made by the researcher and extensively discuss the relevant criteria in interpretive work.

In the present contribution, I will take some references to experimental work in SSY as a starting point to elaborate on some links and differences in the views and actual practices of researchers working from an interpretive and an experimental perspective. My claim, in short, is that these two perspectives, despite all fundamental differences, have more common ground than the methodological juxtaposition found in textbooks, including SSY, might suggest, and that both may have much to gain from cooperation.

Let us first examine the concept of human being. According to SSY, "(i)n interpretive research, human beings are understood not as objects, but as agents" (p. 46), who "actively and collaboratively" construct and populate their social world. The interpretive researcher's aim is to understand the subjective meanings that human beings give to their actions. Clearly, this claim is meant to differentiate their approach from the positivist and quantitative template and implicitly criticizes that approach for "decomposing" individuals into bundles of variables. While this may, at least to some extent, adequately describe the perspective of survey-based research, it does not capture the way in which individuals are treated in laboratory experimental research. Experimentalists are interested in the way in which human beings respond to certain stimuli and whether there are any regularities in these responses. In this sense, they construct a world in the laboratory which is as real as a decision condition to the subjects taking part in the experiment as is the "natural" social world to subjects in their everyday world. For example, experimental economists may be interested in the way human beings respond to price changes, thereby generating a new market equilibrium, or political scientists may be interested in the way people change their voting strategy as a function of information on the relative shares of parties in a poll, thereby contributing to the electoral victory of a particular party (for state-of-the-art overviews, see Plott and Smith 2008; Druckman, Green, and Kuklinski 2011; see also Kittel, Luhan, and Morton 2012). But they maintain the holistic concept of the subject in possession of agency.[2] They are less interested in whether particular socio-economic traits or specific attitudes increase or decrease the likelihood of buying at certain prices or of voting strategically. They maintain the holistic concept of the human being, however, for a different reason than interpretivists. Experimental research is not interested in the specific condition of a particular individual and the uniqueness of a specific situation, but averages out such conditions through random assignment to the group receiving and not receiving the stimulus of interest. It is the subject that acts in response to either the treatment or the control condition, but it is only the decision to act that is of interest to the experimentalist, not the subject as such (Morton and Williams 2010).

This leads to another claim by interpretivists, that individual agency cannot be understood in a decontextualized manner (SSY, p. 48). Implicitly, hereby, positivists are accused of disregarding context. What, however, is the context? It is infinite, and thus a researcher has to make myriad decisions about which elements of the context are to be considered relevant to the subject's actions of interest. Interpretivists deal with this question by exploring the meaning of actions, objects, and relations to the subjects by observing and interviewing them, or by reading their artifacts. They thus emphasize the meaning of context from the subject's perspective. They focus on the subject's "reading" of the situation, hence on the subjective cognition of relevant context, and the scope of the context is defined by the subject's range of attention, to the extent that the researcher is able to "look into" the subject's mind. Experimental economists and those experimentalists in political science and sociology whose research is inspired by that logic, in contrast, attempt to induce in subjects one specific understanding of the situation through monetary incentives and then (more or less) systematically vary single elements of the context one by one in order to test whether it is relevant to the subjects' actions. Interpretivists thus aim at understanding the relationship between context and action from a "within" logic, assuming that it is the representation of the context in the individual that causes the action. Experimentalists attempt to hold constant the "within" logic in order to understand that relationship through variation of the context.

There is a second twist to the problem of contextuality. Knowing the context tells only part of the story. It allows the researcher to identify social conditions, motives, or reasons of an individual that help in understanding why she responds in the observed way to a stimulus in order to make a claim about the meaning of that action to the individual in the relevant context—or, put otherwise, to *understand* the action. But interpretivists have to assume some "mechanism" linking context to action. Although the term "mechanism" may sound suspicious to the interpretive researcher, it does not mean that there is anything mechanical in subjects' reactions to context conditions. You do not push a button and a subject acts according to a script.[3] A mechanism is a regularly occuring logic of action which does not need further differentiation in order to be accepted as a statement about the relationship between two observations. Interpretivists can reconstruct a mechanism from tales told by the agent or by other observers of the agent's actions, but the judgement about the trustworthiness of such claims ultimately rests on plausibility probes, hence whether the action is either considered typical because of prior knowledge about regularities of action in certain circumstances or considered reasonable in its deviation from expected behaviour because of a compelling story that motivates the deviation. Interpretivists will claim that their knowledge of the context gained through extended exposure to the field will guide them. But in my reading, this is a disguised statement about regularities. Understanding through extended exposure is nothing else than a summary of observations collected in the field, no different in principle from the understanding, or tacit knowledge, in social and other situations that we gain through repeated experiences in our social world.

It is this missing link of all observational research, hence also of interpretive work, which the experimental approach can help illuminate. By systematic variation of context conditions of individual actions, experiments can provide knowledge about what we can expect under specific circumstances. They help us differentiate between expected and surprising behavior and thus provide a foundation on which claims about the trustworthiness of a specific interpretation of behavior can build. In their focus on the regular, the replicable, the general, the law-like, experimentalists isolate the stimulus-effect conditions as much as possible from interfering factors. They thus decontextualize social interactions. By inducing preferences through monetary payoffs, the context is expected to trigger a particular mechanism which is scrutinized by comparing realized to expected outcomes. They cut away all stuff that is considered of crucial interest by the interpretive researcher because it is the source of making sense of experimental results.

At the same time, this emphasis on isolation by experimentalists implies that interpretivists have something to offer to them. Experimental findings suffer from the opposite problem of interpretive work: In order to become relevant for understanding natural situations, they need to be amended by the specific context. To the extent that individual traits are part of the context of decisions and systematically affect behavior, survey data can be supportive, but it is the extensive knowledge of context and subjective meaning-making of the context that interpretive research can supply, which helps us in understanding why subjects in a certain situation, which must be assumed to be unique in the exact contextual constellation, follow one particular of several potential logics of action that experimental findings suggest. The subject's mind—her motives, reasonings, and ways of making sense of the context—remains a black box to the experimental researcher.

Another layer of misunderstanding of experimental work by interpretivists can be observed with regard to the logic of reasoning. Typically, experimental research is depicted as the prototype of positivist methodology adhering to a deductive logic of reasoning (SSY, pp. 26–7, 113).[4] A mathematical, for example, game-theoretic model will indeed be derived in a deductive manner from first principles, such as rationality assumptions and specific axioms of behavior. However, experimentalists are much more cautious in their claims when it comes to testing such deductive results: "Predictive hypotheses, as expectations about observations, are sometimes inspired initially and directly from a theoretical model. I choose the word 'inspired' rather than the word 'derived' advisedly; this is because of the large reservoir of personal experiential knowledge that we must draw on in order to make operational a test of any theoretical construct with empirical observations. That knowledge is always subject to revision in the light of new direct examinations" (Smith 2010: 4).

Empirical, and thus also experimental, research always entails an inference from the observed to the unobserved, and thus is inherently inductive in nature. In this vein, Bardsley et al. (2010: 145) have recently argued that "a method of systematic inductive inquiry is now evolving in experimental economics." Many experiments are not designed to "test" a formal model stated as a set of mathematical equations, but try to make sense of some unexpected observations by exploring different possible explanations. Hence, the analyst attempts to understand the meaning that subjects attach to their actions in the laboratory. One of the most fascinating and provocative unexpected findings from experimental work, social preferences, has spurred a whole industry of exploratory experiments. While a small number of deviations from behavioral expectations that are deductively derived from the game theoretical axioms might be considered as erroneous decisions, the consistent deviations from egoistic utility maximization observed in a variety of contexts needed to be accepted as a challenge. Extensive experimental work, confirmed by many replications, has consistently shown that other-regarding considerations such as reciprocity, envy, and altruism are important elements of subjects' action repertoires alongside the traditional modeling assumption of selfish utility maximization (Fehr and Gintis 2007).[5] The difference to interpretive work is that the experimentalist embarks on an enterprise of probing different potential explanations of the unexpected finding by designing different experimental conditions.

The example of social preferences also entails an element that links back to the discussion of subjectivity above. The departure from induced preferences under the individual egoistic utility-maximizing rationality assumption is not considered irrational behavior anymore, but behavior that makes sense to the subjects. Thus, the acceptance of social preferences reinserts the subjectivity that may have been lost in early formal models of behavior, without, however, giving up the claim to search for regularities.

But how does one obtain interesting potential explanations of such a puzzling finding? According to SSY, interpretivists have an answer by proposing a third, alternative, "logic of reasoning" next to deduction and induction. Abduction is described as the "search for possible explanations that would render the surprise less surprising" (SSY, p. 28). More precisely, "(i)n this puzzling-out process, the researcher tacks continually, constantly, back and forth in an iterative-recursive fashion between what is puzzling and possible explanations for it, whether in other field situations...or in research-relevant literature. The back and forth takes place less as a series of discrete steps than it does in the same moment: in some sense, the researcher is simultaneously puzzling over empirical materials and theoretical literatures" (SSY, p. 27). On the pages following these descriptions, SSY reiterate the idea of reiteration and recursion, and elaborate on examples of abductive reasoning, but they do not offer any logic of abduction in the sense of the logics underlying deduction and induction. The latter two are truth claims that state relationships between something known and something unknown. The acceptability of both claims rests on some criterion, such as a proof of the lack of contradiction of statements or a statistical significance test. Abduction does not entail a truth claim because it seems to refer to a relationship between a puzzle and a potential explanation. There is thus no anchor in something commonly accepted from which research reaches out. SSY do not offer any criterion according to which abductive reasoning might be proven wrong. Experimentalists would, in Popper's (1959) tradition, suggest intellectual curiosity and creativity as the basic conditions driving the process of identifying possible explanations to a puzzling finding. They would, however, deny to this process any "logic." How do we decide whether an abductive line of reasoning is a contribution to our knowledge about the world? Even more, submitting this process to any systematizing logic would mean that certain possible lines of thinking may be blinded out by the logic itself. From this perspective there is a puzzling contradiction in SSY's argument between the emphasis on openness to any new experience in the field and the explicit rejection of systematic criteria for the collection of empirical material on the one hand, and the ambition to subordinate the processes of theorizing and finding new ideas to some logic of reasoning on the other.

As long as formal theories have been based on the classical self-interested conception of rationality, the close relationship between rational choice and quantitative methodology posited by Goldthorpe (1996) seemed quite reasonable. However, the repeated finding that experimental results were at odds with the theoretical predictions has spurred interest in methods able to shed light on motivation. While one route has been to dig deeper into the neural processes in human brains (Glimcher et al. 2008), another route has been to search for other means of eliciting statements from which inferences about motives can be made. Because interviewing is always confronted with potential misrepresentation bias, it is observation of communicative behavior which, in my view, offers a promising route. Field observations can certainly help to uncover motives but at the cost of limited possibility to identify context conditions triggering the one or the other motive. A recent development in experimental social science has been the provision for unconstrained communication through computerized chat facilities, which allows researchers to observe communicative action in social interaction (Frohlich and Oppenheimer 1993; Karpowitz and Mendelberg 2011; Dickson, Hafer, and Landa 2008; Kalwitzki, Luhan, and Kittel 2012). One lesson learnt from a comparison of this research with experiments restricting communication content (Goeree and Yariv 2011) is that the fewer constraints are imposed on communication, the lower the predictive capacity of current formal models. Groups as contexts strongly influence the meaning which subjects give to their actions. Even if the ultimate goal of these experiments is to test theories by comparing their predictions to empirical regularities, the coding of chats involves a contextualized interpretation of statements involving research strategies inspired by the interpretive paradigm. Perhaps someday creative researchers might find coding algorithms that substitute computer power for the capacity of intellectual understanding of human brains. For the time being, however, experimentalists can learn much from their interpretive peers for studying experimental results involving unconstrained communication. The textbook by SSY is definitely an excellent point of departure for such work.

It is a truism that interpretivists and experimentalists have different views of the world. But because they also ask different questions, neither covers the full range of possible questions and answers; even more, both systematically blind out those questions that are relevant to the other. Thus, in order to make sense of our social world, interpretivists and experimentalists generate complementary knowledge, and only the combination of both elements can provide a fuller account of human interaction. In this sense, the two perspectives provide different elements of the Hempel-Oppenheim (1948) logic: Experimentalists provide the regularity, interpretivists provide the unique context of a specific situation, and together they construct an explanation that allows us to observe the general in the unique and the unique in the general. Hence, interpretivists and experimentalists are like two sisters, each owning only one of two keys needed to open the safe inherited from their grandparents.

Notes

[1] I should add that the hostility dominating this debate in the past has been fuelled by both sides.

[2] Interestingly, if language, as interpretivists (SSY, p. 46) maintain, confers meaning, it must be noted that experimentalists speak of "subjects" when they refer to participants in laboratory experiments.

[3] Although all too often we observe human interactions to apparently follow a script. Most escalations of conflict seem to follow an automated logic of retorsion.

[4] SSY (p. 52) seem to view statistical and mathematical ("formal") models as equivalent. However, a mathematical model is a deductive theoretical statement about assumed relationships derived from first principles, while a statistical model is an empirical representation of such a statement in an inductive inferential logic.

[5] The conclusions to be drawn from these experimental findings are still under debate. See Binmore and Shaked (2010a, 2010b), Fehr and Schmidt (2010), and Eckel and Gintis (2010).

References

Bardsley, Nicholas, Robin Cubitt, Graham Loomes, Peter Moffatt, Chris Starmer, and Robert Sugden. 2010. *Experimental Economics: Rethinking the Rules*. Princeton: Princeton University Press.

Binmore, Ken and Avner Shaked. 2010a. "Experimental Economics: Where Next?" *Journal of Economic Behavior and Organization* 73, 87–100.

Binmore, Ken and Avner Shaked. 2010b. "Experimental Economics: Where Next? Rejoinder." *Journal of Economic Behavior and Organization* 73, 120–121.

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

Dickson, Eric S., Catherine Hafer, and Dimitri Landa. 2008. "Cognition and Strategy: A Deliberation Experiment." *Journal of Politics* 70:4, 974–989.

Druckman, James N., Donald P. Green, and James H. Kuklinski. eds. 2011. *Cambridge Handbook of Experimental Political Science*. Cambridge: Cambridge University Press.

Eckel, Catherine and Herbert Gintis. 2010. "Blaming the Messenger: Notes on the Current State of Experimental Economics." *Journal of Economic Behavior and Organization* 73, 109–119.

Fehr, Ernst and Herbert Gintis. 2007. "Human Motivation and Social Cooperation: Experimental and Analytical Foundations." *Annual Review of Sociology* 33, 43–64.

Fehr, Ernst, and Klaus M. Schmidt. 2010. "On Inequity Aversion: A Reply to Binmore and Shaked." *Journal of Economic Behavior and Organization* 73, 101–108.

Frohlich, Norman and Joe Oppenheimer. 1993. *Choosing Justice: An Experimental Approach to Ethical Theory*. Berkeley: University of California Press.

Gerring, John. 2012. *Social Science Methodology. A Criterial Framework*, 2nd ed. Cambridge: Canbridge University Press.

Glimcher, Paul W., Colin Camerer, Russell A. Poldrack, and Ernst Fehr. 2008. *Neuroeconomics: Decision Making and the Brain*. Amsterdam: Academic Press.

Goeree, Jacob K. and Leeat Yariv. 2011. "An Experimental Study of

Collective Deliberation." *Econometrica* 79, 893–921.
Goldthorpe, John. 1996. "The Quantitative Analysis of Large-Scale Data-Sets and Rational Action Theory: For a Sociological Alliance." *European Sociological Review* 12:2, 109–126.
Hempel, Carl G. and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Philosophy of Science* 15, 135–175.
Kalwitzki, Thomas, Wolfgang J. Luhan, and Bernhard Kittel. 2012. "Experimental Chats: Opening the Black Box of Group Experiments." In *Experimental Political Science: Principles and Practices*. Bernhard Kittel, Wolfgang J. Luhan and Rebecca B. Morton, eds. (London: Palgrave-Macmillan).
Karpowitz, Christopher F. and Tali Mendelberg. 2011. "An Experimental Approach to Citizen Deliberation." In *Cambridge Handbook of Experimental Political Science*. James N. Druckman, Donald P. Green, James H. Kuklinski, and Arthur Lupia, eds. (Cambridge: Cambridge University Press), 258–272.
King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry. Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
Kittel, Bernhard, Wolfgang J. Luhan, and Rebecca B. Morton, eds. 2012. *Experimental Political Science: Principles and Practices*. Basingstoke: Palgrave-Macmillan.
Morton, Rebecca B. and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality. From Nature to the Lab*. Cambridge: Cambridge University Press.
Moses, Jonathon and Torbjørn Knutsen. 2012. *Ways of Knowing. Competing Methodologies in Social and Political Research*, 2nd ed. Basingstoke: Palgrave-Macmillan.
Plott, Charles R. and Vernon L. Smith. eds. 2008. *Handbook of Experimental Economics Results*, Vol. 1. Amsterdam: North-Holland.
Popper, Karl. 1959. *The Logic of Scientific Discovery*. London: Routledge.
Smith, Vernon L. 2010. "Theory and Experiment: What Are the Questions?" *Journal of Economic Behavior and Organization* 73:1, 3–15.
Yanow, Dvora and Peregrine Schwartz-Shea. 2006. *Interpretation and Method. Empirical Research Methods and the Interpretive Turn*. Armonk, NY: M.E. Sharpe.

Art, Methodology, and the Practice of Designing (Interpretive) Research

Christian Bueger
Cardiff University, Wales
buegercm@cf.ac.uk

Methodology can be understood as the art of building references that last. As an art it is concerned about capturing the world in language, translating it into the discourse of academic communities, and creating chains of circulating references that resist controversy and the critique of our peers. Texts about art come in different genres, as do methodology books. A visit to the art section of a local bookstore reveals several of such genres. First, there are the "how-to" books. Picking up the *Joy of Painting with Bob Ross* (Ross 1993) tells us how to hold the brush, how to organize our canvas and color palette, and how to paint in 20 easy steps a picture that is supposed to resembles a landscape or similar. Second, there are the broad, encompassing histories of the development of art. *A World History of Art* (Honour and Flemming 2009) is a book we hardly would read from page 1 to 996. Yet, it comes in handy if we want to situate an artist or a piece of art in an epoch or certain style. Then there are, third, the published lectures which are more difficult to digest and present systematic and often philosophical overviews of a distinct tradition or style. A book such as Kirk Varnedoe's (2006) brilliant *Pictures of Nothing* allows us to follow the linkages between artists and styles and gives us an understanding of why a black-painted canvas has been considered to be a revolutionary piece of art at some point of time. Fourth, there are the exhibition catalogues which focus on the retrospective of a distinct artist or the theme of an event. They allow us to appreciate an exhibit or to keep it in good memory. Finally, there are also art manifestos such as the *Dada Manifesto* or the *Situationist Manifesto*. Manifestos want to bring about a revolution in thinking about and doing art. They are tools for making sense of one's own practices and allow us by subscribing to them to become part of a gang, movement, community or network.

Peregrine Schwartz-Shea and Dvora Yanow's new book on the practice of designing interpretive research sits comfortably between these genres. If my local bookseller wouldn't sort his art section alphabetically, he would have trouble sorting it in. The little white book with grey stones on its cover tells you how to organize your canvas, but not how to hold your brush. It resists a "Bob Ross" style of outlining the 20 steps you have to follow to produce a good piece of interpretive research. In the words of the authors, it is a book that "engages some very practical issues [but …] it is not a how-to volume" (p. 9). Relying on the idea that research is driven by tacit knowledge and practical experience, the book provides a guide for the practical work and the many choices that go into the practice of designing research. It is successful in letting the reader experience these practical choices, for instance, in drawing on the problem of how to interpret a photograph of an artifact, or in using brief illustrations from the experience of researchers throughout the book. The lively metaphorical style further contributes to this experience. The authors compare literature reviews with dinner parties (p. 35). They describe the interpretive researcher as a captain of a ship "attuned to weather conditions and riding the resulting waves" rather than strictly following an initial course (p. 36). The book is also a *tour de horizon* through different ideas and authors which belong to the family of interpretive research. It helps to situate ideas, such as abduction or the hermeneutic circle, without turning encyclopedic. It tells you how to draw linkages and appreciate commonalities of interpretive researchers, such as the importance of context and process. But it isn't exclusively concerned with these linkages or explaining which piece of design should be considered as path breaking for this and that reason. Vividly, it is also a book that tells us how to appreciate interpretive work, whether it is the interpretive study we recently have read, the student dissertation on our desk for assessment, or reflections on our own experience of probing out interpretive ideas.

Finally, Schwartz-Shea and Yanow also have written an intervention. Notably in the later chapters they address audiences such as reviewers and grant authorities to make a case that interpretive research requires different standards of evaluation and faces different challenges than the natural sciences or the traditionalist positivists do. It is also clearly a book written to serve as a source of identification; it wants to organize and facilitate exchanges of scholars, win the hearts and minds of skeptics, and open a pathway for newcomers who want to engage in interpretive research. Indeed, it ends with such a call for engagement.

Hence, Schwartz-Shea and Yanow's short book gives you many things in one. Of course, that the book plays with different genres and wants to serve multiple purposes can also be turned against it. The reader who wanted his how-to book, the philosophical overview, or the detailed discussion of different interpretive positions will have to look elsewhere. Yet he would miss out on a little gem that opens space for reflexive thinking (whether one wants to subscribe to the club of interpretive research or not). It triggers reflexivity about what needs to be assembled, which resources need to be considered, who needs to be convinced, what is there to plan, and which choices have to be made to perform the practice of planning and designing of research.

Talking About Practices

The reason why the book succeeds in opening such reflexive space is that it has a clear argument: it argues that designing and planning research is a "practice." As the social studies of sciences have reasoned for decades, science, research, and methodology are fields or bundles of practices (Rouse 1996, Schatzki 2001). Together with the exclaimed "practice turn" and engagement with contemporary classics such as Bruno Latour's (1987) *Science in Action*, or Karin Knorr Cetina's (1999) *Epistemic Cultures*, also political science is slowly learning this message and the implications it has for how we do research, how we write it up, how we teach it, and how we evaluate it (Bueger 2012, Kessler and Guillaume 2012). Throughout the book Schwartz-Shea and Yanow develop an understanding of designing research as practice in this sense. Practices are organized social activities; they are materially situated and involve things and their use, bodies, mental states, and emotions, and they involve all sorts of practical understandings and rules of thumb.

Rules, such as Bob Ross' rules for painting a landscape or classical positivist rule sets, are from such a perspective relevant in so far as they help organize a social community and provide points of orientation of how to act under situations of uncertainty (such as those newcomers to a practice face). To perform a practice and to follow rules in action, however, much more and centrally, practical knowledge is necessary. Hence, Schwartz-Shea and Yanow's hesitation to outline rules. Moreover, we appear to be in a situation in which the problem is not the lack of rules, but of attempts to explicate practical knowledge. Rightfully the authors foreground that planning is about making choices and that rules are only helpful if they assist us in making informed choices and coping with practical challenges such as how to convince an ethics committee or how to fill out a data management sheet.

Schwartz-Shea and Yanow outline a thorough understanding of research design as an inherently social activity. Research design is about engaging with scholarly communities, developing research questions, and writing literature reviews which are "intelligible" to communities. It is coping with the demands of grant authorities and universities, and is ultimately about persuasion—that is, convincing others about the quality and benefits of one's research (p. 3). They understand research as well as the practice of designing it as a process, that it is a constant movement of coping, negotiating, and translating between theory and data, between prior knowledge and field experience. Research is contingent, it is about surprises and drawing up creative solutions which neither symbolic structures nor the situation at hand determine. As a practice it is ethical, but also material. It concerns body movements, emotions, and sexuality. And it is about things and equipment, that is, inscription devices, tape recorders, or social networking platforms. Research does not stop when data gathering and analysis over. Planning has to acknowledge writing and dissemination, convincing peer reviewers or publishing houses.

The authors' sketch of the practice of designing research is convincing. Yet, read as an intervention in the intellectual discourse on methodology, one would have wished that the book would have been much more explicit and systematic about outlining the theoretical understanding of science as practice that underlies it and its consequences for how we learn and evaluate research (whether interpretive or otherwise). Writing a book about practice, explicating practical knowledge is difficult, and that is even more the case if one conducts the experiment of writing a multi-purpose book as Schartz Shea and Yanow have. Yet for my personal taste for a book about practice, there are too few empirical stories about practice. I am tempted to suggest that the book should come with complementary reading such as Latour's (1999) essay "Circulating Reference" in which he describes in painstaking detail all the little steps, all the choices, the hard work of constructing references that were necessary to translate the Amazon Forest into an academic article. As he documents, building chains of references, relating things to words and references to signifiers is not only hard work but involves taking risks. Any little step might cause the chain to break down and it has to be fixed again. Drawing on such empirical stories of decision making would have notably increased the value of the book in teaching, but it would also have induced to document better the high level of uncertainty researchers face.

Performativity and Engaging with Exoteric Communities

One of the virtues of the book is, as discussed above, that it extends our understanding of what is part of the practice of designing research, including convincing colleagues and planning for writing, but centrally also more recent challenges such as coping with an ethics committee or drafting a "data management" plan. While certainly these contemporary challenges could have been given lengthier treatment, it points us to a major shortcoming of the book: The extension does not go far enough. Contemporary research designs (certainly in the UK) have to be increasingly justified by demonstrating their relevance for non-academic communities. "Impact," the resonance of knowledge in non-academic communities, has become a core criterion for evaluation. Rightfully, the authors suggest that writing up research is a "scholarly, political act" (p. 39), it is an act of "world making" (p. 113). Rightfully, they stress the importance of asking "Knowledge for what purposes? Knowledge for whom?" (p. 46). Yet throughout the book—to use Ludwik Fleck's vocabulary—they relate these questions only to the "esoteric communities" of science and not to its exoteric ones. Research not only is relevant for esoteric communities, and the consequences it might have extends beyond the communities studied. In International Relations and especially security studies there is a rich body of literature (Huysmans 2002, Bueger and Villumsen 2007, Ish-Shalom 2009) that shows how research results travel, catch the attention of decision makers, and are used to make sometimes good and sometimes very bad judgments. Extending research design to consider questions of impact, relevance, and consequences is not only a necessary step because the funding bureaucracy forces us to ponder them. It might be exactly here that interpretive research has very good answers and a comparative advantage to other isms. Interpretive research entails a different form of engagement with the world of practitioners and indeed has always been reflexive about consequences. Extending our understanding of designing research in such a way leads us to the performativity of methodology (Mol 1999, Law 2004). If methodology is about making worlds, than methodological decisions are not only about convincing our colleagues but about the question of which worlds we want to enact and how in our research practices.

Leaving these caveats aside, Schwartz-Shea and Yanow's book is a successful experiment for writing a book about a practice for its practitioners. It extends our understanding of the practical work and experience required to perform the practice of designing research, and as an intellectual intervention it opens up reflexive space for further praxiographic studies on methodology and research practice.

References

Bueger, Christian. 2012. "From Epistemology to Practice: A Sociology of Science for International Relations." *Journal of International Relations and Development* 15:1, 97–109.

Bueger, Christian and Trine Villumsen. 2007. "Beyond the Gap: Relevance, Fields of Practice, and the Securitizing Consequences of (Democratic Peace) Research." *Journal of International Relations and Development* 10, 417–448.

Honour, Hugh and John Flemming. 2009. *A World History of Art*, 7th ed. London: Laurence King.

Huysmans, Jef. 2002. "Defining Social Constructivism in Security Studies: The Normative Dilemma of Writing Security." *Alternatives: Global, Local, Political* 27:1, 41–62.

Ish-Shalom, Piki. 2009. "Theorizing Politics, Politicizing Theory, and the Responsibility That Runs Between." *Perspectives on Politics* 7:2, 303–316.

Kessler, Oliver and Xavier Guillaume. 2012. "Everyday Practices of International Relations: People in Organizations." *Journal of International Relations and Development* 15:1, 110–120.

Knorr Cetina, Karin. 1999. *Epistemic Cultures. How the Sciences Make Knowledge.* Cambridge: Harvard University Press.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge: Harvard University Press.

Latour, Bruno. 1999. "Circulating Reference. Sampling the Soil in the Amazon Forest." In *Pandora's Hope: Essays on the Reality of Science Studies*. Bruno Latour, ed. (Cambridge: Harvard University Press), 24–79.

Law, John. 2004. *After Method. Mess in Social Science Research*. London/New York: Routledge.

Mol, Annemarie. 1999. "Ontological Politics. A Word and Some Questions." In *Actor Network Theory and After*. John Law and John Hassard, eds. (Oxford: Blackwell Publishers).

Ross, Robert N. 1993. *The Joy of Painting with Bob Ross*. Sterling, VA: Bob Ross Incorporated.

Rouse, Joseph. 1996. *Engaging Science. How to Understand its Practices Philosophically*. Ithaca: Cornell University Press.

Schatzki, Theodore R. 2001. "Practice Theory." In *The Practice Turn in Contemporary Theory*. Theodore R Schatzki, Karin Knorr Cetina, and Eike von Savigny, eds. (London/New York: Routledge), 1–14.

Varnedoe, Kirk. 2006. *Pictures of Nothing: Abstract Art since Pollock* (The A.W. Mellon Lectures in the Fine Arts). Princeton: Princeton University Press.

Reading Our Readers

Peregrine Schwartz-Shea
University of Utah
psshea@poli-sci.utah.edu

Dvora Yanow
Wageningen University, The Netherlands
dvora.yanow@wur.nl

Interpretive Research Design: Concepts and Processes is less a "how-to" book than a "how to think about such matters" book, to paraphrase Atkinson (1992: 8), as these five comment authors have understood. Although we initially thought we would be writing a "how-to" book, the material talked back (cf. Schön 1987), refusing to go in that direction, moving us more and more toward the manuscript now in print. We did not want to write a manifesto in tone—a matter both Tanya Schwarz and Christian Bueger raise—because of our desire to reach across epistemic communities and to speak in an educating spirit (described further below). Nor did we want to write a history of interpretive research contributions or attempt a catalogue of particular thinkers—points made by Corinne Heaven and Bueger—as both have been done elsewhere and were inconsistent with our purpose of focusing on issues arising in designing research. Moreover, that purpose was precisely to intervene in contemporary methods discourses: The language of research design has become so ossified, so closed, that an interpretive logic of inquiry cannot fit within its confines. Moreover, as Katherine Cramer Walsh and Bernhard Kittel have understood, we wrote not in the spirit of engaging "methods wars," but rather more in a "Can't we all just get along?" vein, asking in a more intentionally educational spirit what researchers of all stripes need to know in order to talk with, rather than past, one another.

We thank all five commentators for reading the book "from the inside," as it were, with an ear attuned to what we were trying to achieve, and to *QMMR* editor Robert Adcock for organizing the exchange. As Tanya Schwarz notes in her opening sentence, we had three classes of reader in mind in writing the book: graduate students, experienced researchers, and teachers of methods. We see that, indeed, our five readers—sequenced here in some sense in keeping with that order—have taken different things from the book, in part reflecting the needs and concerns of these three different kinds of reading.

We are very pleased that both Tanya Schwarz and Corinne Heaven think the book (which we refer to, below, as SSY) will be useful to students (and not only those doing interpretive research) navigating contemporary methods debates as they wend their way through dissertation committee meetings, proposal and grant submissions, and the subsequent steps that face newer members of the academic community—especially in those areas of research that remain dominated by positivist-influenced approaches. We are tremendously concerned, as is Schwarz, with graduate students lacking access to interpretive ways of thinking in their home institutions, and we are pleased to think that this book would be one resource to potentially fill that gap.

Schwarz points to some key challenges for researchers seeking to pursue interpretive empirical work. With respect to writing grants for departments, agencies, and foundations that remain positivist, we agree with her that this remains a fundamental challenge in today's research world. One answer to the question she poses—how can proposal writers explicate their research in acceptable fashion when length limits preclude "fully fleshing out the assumptions and foundations of their work"?—is that graduate students (and others) need not reinvent the wheel. Over a century's worth of literature has laid the groundwork for such undertakings, going back to Schütz, Husserl, and others; secondary literature explicates these further; and current methodological discussions, such as in our book, ground these ideas in research processes. Whether due to the proliferation of methods courses and textbooks or because of the time and "productivity" pressures that can thwart exploratory impulses, students seem not to pursue these sorts of discovery processes on their own, following the breadcrumb trail as one source leads to another and using these to advance their desired methodological arguments. The hurdle is to learn that these approaches and works exist[1] and then, discovering what they say, to use them expeditiously to support one's arguments.

Concerning interacting with a dissertation advisor or committee members who are not conversant with interpretive methods, we hope to have provided translation tools that would enable a "third way" between engaging only interpretive scholars (and granting agencies) and rewriting the research in a positivist mode. In an "educating" spirit, an interpretivist can write or say, explicitly, "Were I doing this research project in a positivist mode, I would at this point in the proposal (or manuscript) discuss X (reliability, validity, or whatever the relevant term); but in interpretive research, this concept is called Y (citations here)" and so on. Depending on the character of the adviser and of the relationship, the student might gift her with a copy of this book and a plea to read certain passages for subsequent discussion, in the hope of talking their way to mutual understanding.

With respect to influencing the broader world of positivist research, in our view this is not a battle that political science graduate students and junior faculty can fight on their own, whatever resources might be available to them. We continue to work with other senior scholars in qualitative-interpretive arenas across the social sciences to articulate and advance the utility, indeed the necessity, of interpretive-qualitative methods for the understanding of social life in its various domains (see, e.g., Becker's 2009 critique of U.S. National Science Foundation approaches). With the success of the Sage *Handbooks* (now in a fourth edition; Denzin and Lincoln 2011), new journals, and other mainstream interpretive-qualitative publications, there is now a critical mass of contemporary scholarship that can be used to educate grant-making and other agencies about the changes in the understanding of natural, physical, and social science that have been developing over the last half century (whether in philosophy, feminist theory, science studies, or the various social science disciplines and fields of inquiry). Whether such efforts will ultimately be successful is not clear; government funding of social science was controversial when these public funding efforts began (King 1998; Larsen 1991), and Summer 2012 efforts by Congress to defund the political science program at NSF (Voeten 2012) give us pause. Still, in 2009 NSF funded our Workshop on Interpretive Methodologies in Political Science, indicating a possible growing understanding within funding agencies, even if not (yet) among politicians and the general public, of the scientific status of interpretive work.

Moreover, thanks to those senior scholars' efforts, along with *QMMR*, this newsletter, CQRM/IQRM, the ECPR Summer and Winter Methods Schools, and so on, the research methods world today is not the same as it was in 1994, or in 2003.[2] Still, more work is needed: curricular reform to produce not only stand-alone courses on qualitative-interpretive research and interpretive methods, but also philosophy of science and research design courses that engage these ideas; and efforts to educate general methods textbook authors to more fully engage and accurately represent interpretive and qualitative methods, especially in their design chapters. We reject, however, the idea that frontal attack, through a cry (Schwarz) or a manifesto (Bueger), would have been the most effective way for us to proceed in this book (as discussed in SSY, p. 11); and we do so on pedagogic grounds. Paulo Freire, the exiled Brazilian educator, worked from the fundamental principle that people would learn most effectively when what was being taught spoke to their needs, in their language (1970, 1973). For example, teaching a group of non-literate Latin American peasants who needed water for crops, drinking, etc., he began with the word for "well." We take this principle to heart: people will be more likely to learn when teaching begins with what they already know. Certainly, we decry the hegemony of methodological positivism; but while there are times and places for manifestos, we didn't see the utility in making that the explicit center of this book—indeed, in our view, doing so would defeat our purpose. Our 2006 *Interpretation and Method* was constructed to demonstrate that interpretive methods not only deserve an equal place at the social science table, but have much to offer to knowledge production. We hope that the effort at translation in SSY advances the same claims, without strident renunciations that would block a learning process before it could begin.

We appreciate the notion advanced by Corinne Heaven that reading the book is "itself an exercise in reflexivity" in that the book calls on researchers to consider what it is that they are doing, and why. She has caught the idea that struck us as we worked through the manuscript, that the entire research process, including the writing, looks and feels different depending on the logic of inquiry underpinning one's project. It is, in other words, not merely a matter of using different methods; rather, the character of each type of research project is different, as she notes. This is, in part, why the material pushed back at us: writing and reading are linear, and that lends itself well to the linear ways of describing and diagramming how-to research processes; the much more iterative-recursive processes we describe (which may, in the end, describe not only interpretive research) could not be shoehorned into the same visual and narrative style with which one can articulate "the scientific method" and its associated forms of research.

Why did we spend a lot of book space engaging positivist research logic in a book purporting to be about interpretive scientific logic? Because, in our experience, the former is quite widely known—much more so than the latter—and it has set the stage for the research design language that largely populates course syllabi and methods textbook pages. If, as both Heaven and Schwarz posit, graduate students are not well-exposed to interpretive ideas, how are they to make sense of them when those ideas are presented in a self-standing way, without a benchmark in what they already know? The Freirean educational principle is again at work. We do remain somewhat nervous about the table Heaven mentions (SSY, p. 113) because any dichotomy runs the risk of oversimplifying and reifying what are complex, wide-ranging traditions and practices within both interpretive and positivist approaches to research; but it seemed a useful teaching device, which might be used alongside other materials that vividly demonstrate this complexity and variety.

In all writing (as in all research), authors make choices—all the more so when faced with a severe page limit, such as required by this book series, which we were already pushing. Unable to do everything, one makes judgments as to what one considers more and less significant for the development of the argument and one's imagined reader(s). We had already written extensively elsewhere about our own methodological commitments (Yanow and Schwartz-Shea 2006), including a list of some two dozen different interpretive methods (p. xx), but more importantly, this book did not seem the place for such writing, especially as the literature is massive. We opted to present a very brief list of basic readings (SSY, p. 44), with the thought that these, along with the extensive reference list for the book as a whole, would provide the interested reader with a starting point. More breadcrumbs. Still, Heaven is correct in pointing out that there is no single way to do interpretive research, something we noted on page 13, and that this book is pitched broadly (and perhaps not for a reader who is well immersed in these matters, such as a scholar in IR, where these sorts of issues and their attendant literatures have been at the forefront of discussions and debates for some two decades). Someone who knows she is a Foucauldian will, or should, have no difficulty finding original, secondary, and methods literatures to guide her research projects; SSY may indeed be too elementary in some respects or too broad-brushed for such a scholar. Our imagined reader was, in part, one who does not have those resources, having perhaps only a statistics course to work with, but who feels a mismatch between that approach and what he wants to do. We were also considering this as the initial volume in a series, other volumes of which will demonstrate the varieties of interpretive approaches.

Does interpretive research still need to have its way paved, as Heaven asks? Yes, in many U.S. graduate programs, at least, as well as in others with which we are familiar in Canada, the UK, Europe, Australia, Israel, and India (the limits of our experience and knowledge), as well as when it comes to grant money from national science foundations. This holds for various disciplines, not only for political science (and we should perhaps note that our imagined readers were social scientists broadly speaking, and not only political scientists). In organizational studies, for example, the (U.S.) Academy of Management's methods' section's journal articles, as well as its conference panels, are almost exclusively quantitative. No QMMR there! Sociology departments and leading journals in the U.S. and elsewhere are also dominated by quantitative engagements, as are urban planning, public policy and administration, educational evaluation research, and many other fields. The interpretive research community is certainly better established now than it was when either of us was a doctoral student or new faculty, and we are pleased that it is growing. At the same time, Heaven is right that the presence of and legitimacy accorded to interpretive research changes—from state to state, reflecting national research cultures, and in some cases, from department to department within a single state. (Here is an empirical research project waiting to be conducted!) That said, in those states that look to the U.S. as the epitome of what political science is and should be, for funding as well as curricula, its own positivist hegemony—something Schwarz points to—threatens to overtake "local" practices.[3]

Katherine Cramer Walsh's translation achievements should be an inspiration to the many political and other social scientists who seek to engage research fields that remain predominantly positivist, as hers does. That she has succeeded is testimony to her perseverance and imagination, along with the exceptional quality of her work. We are sympathetic with her experience, having encountered it ourselves. Translation of findings across methodological borders and education of others as to how this research is done are still the lot of many interpretive researchers whose interests draw them to such fields. We are gratified that she has found the book useful for those purposes—and that she picked up on the teaching spirit that informs it.

Walsh's thoughtful reflections give us the opportunity to re-engage an issue that is central to SSY and key for all researchers: the matter of communicating one's concerns, methods, and findings in language that might be understood by a broad range of readers, across epistemic communities. Our hope for the book has been to lessen her and others' load by explicating the logic underpinning interpretive research design, thereby not only enabling these translations but also challenging the current hegemony within research methods texts. But this is only the beginning of a conversation that, in our view, needs to happen more widely across the social sciences—not only in journal debates but also reaching into general textbook discussions of design and evaluative criteria, which have become stale in their repetition of Campbell and Stanley's views, inspirational though these may have been in 1963.

In this spirit, we would encourage interpretive scholars to develop the educating tactic with which Walsh ends her essay—working to get those who invoke standard conceptualizations of a complex concept, such as generalization, to themselves dig deeper. For example, what Walsh calls the "project of generalizability" needs discussions that push the boundaries of standard definitions. In her historical ethnography of the Challenger Shuttle explosion, Diane Vaughan has written:

> Anthropologist Clifford Geertz reminds us that having gained access to an unfamiliar universe, put ourselves in touch with the lives of strangers, and considered what the knowledge thus attained tells us about that society, cultural analysis must *extrapolate the significance* for social life as such, drawing "large conclusions from small, but very densely textured facts." (Vaughan 1996: 393, emphasis added)

Here, contextuality is foregrounded as a mode of generalizing, in contrast to positivist conceptions of generalizability. This sort of approach can be a starting point for revisiting criteria that have too often been used in unreflective and unproductive ways. Walsh describes "generalizing to theory" and generalizing processes. We would add other key questions, such as, "Who claims authority to generalize? What evidence is there about the effectiveness of various forms of generalization (for what and for whom)?" Even those general methods textbook authors who claim to cover ethnographic methods or historical approaches still write their research design chapters in ways that assume only positivist approaches (e.g., Babbie 2010; see also Schwartz-Shea and Yanow 2002).

Still, we did not intend to preach a single mode of being in the research world, let alone to fault those who pursue other paths. Martha Feldman (personal conversation, ca. 1987) once described herself as standing with one foot on either side of a bridge between the quantitative research community and the qualitative one, speaking to both. These are the sorts of choices facing interpretive researchers: to position themselves as translators, as Walsh has done, or to pursue their work on its own terms and grounds. As with all choices, there are tradeoffs. For us, seeing ourselves and others struggling to discuss our/their research using "terminology [that] is not quite right" (Walsh, p. 7) was a major driver behind this book. It is always members of the minority group, in "bicultural" mode (Bell 1990; see also Du Bois, 1990/1903, on "double consciousness"), who learn, of necessity, to converse in both idioms; they are the ones commonly required to translate their experience into the dominant group's language. We hope for a day when positivist researchers, recognizing the existence of a different logic underpinning interpretive research and its scientific character, will ask Walsh about her positionality and how she dealt with that—i.e., when they are curious about "bias" and know enough to ask her about it in terms reflecting the logic of her own (and other interpretive scholars') research, on its own terms, engaging in fruitful dialogue without her (and those other scholars) having to be the only party speaking a "foreign" language in need of translation.

In his engagement with the book's ideas, Bernhard Kittel challenges us on several grounds, and we welcome the opportunity to clarify our meaning in interacting with someone who, while pursuing other modes of research, has long been sympathetic to the need to include interpretive research at the methods table and who understands its methodological objectives and intent. We suspect that some differences in our respective views may be explained by our having taken as our foil discussions in methods textbooks, more than the deliberations of methodologists. We take up his comments at greater length than we do the other four commentators because of the greater length of his own essay and because methodological rejoinders require more space to work out the theoretical arguments.

In his final paragraph, Kittel proposes a working partnership between experimental and interpretive researchers—characterizing them as "two sisters," each of whom holds a different key to the lock that will open their grandparents' safe. The keys are described quite differently: "Experimentalists provide the regularity, interpretivists provide the unique context of a specific situation…" (p. 11). We appreciate the metaphor, as it renders them not only members of the same family, with similar interests in unlocking the store of knowledge, but also on equal footing (although, as many readers may know, siblings often squabble, in particular over inheritances!). While working in partnership is an appealing characterization, the particular key that Kittel assigns to interpretivists is too limiting for what they actually do (e.g., in his view, interpretive approaches are useful for "the coding of chats," p. 11). It is also inconsistent with interpretive understandings of "agency," which contrast in quite telling ways with those of experimentalists and experimental economists, in particular. That Heaven, too, takes up this topic shows that it is an area deserving of comment. Indeed, Heaven's reading is similar in some ways to Kittel's (e.g., where she writes that in SSY, "the subject seems to resemble a foundational figure," p. 6), implying a coherence and autonomy to that subject and his choices which seem quite consistent with Kittel's reading. Conceptualizing human agency is complex territory and, for that reason, not taken up systematically anywhere in *Interpretive Research Design*; so we welcome this opportunity to say a bit more on the topic, even if it proves insufficient given our space constraints here.

Kittel quotes our perspective: "'In interpretive research, human beings are understood not as objects, but as agents' (SSY 46), who 'actively and *collaboratively*' construct and populate their social world" (p. 9, emphasis added). In focusing primarily on individual agency, however, his essay neglects the collaborative or (in more methodological terms) inter-subjective dimensions of human meaning-making that are formative in interpretive understandings of agency (even if, as Heaven observes, we did not dwell on this point). This narrowing of the focus onto individuals likely emerges from the game-theoretic lens Kittel takes in advancing his argument. Whereas in the experimental literature cited by Kittel there have indeed been significant moves away from the "egoistic utility maximization" assumptions of early research such that experimentalists now probe "reciprocity, envy and altruism" (p. 10), the individualistic conceptualization of agency remains. He writes:

> The departure from induced preferences under the individual egoistic utility-maximizing rationality assumption is not considered irrational behavior anymore, but behavior that makes sense to the subjects. Thus, the acceptance of social preferences reinserts the subjectivity that may have been lost in early formal models of behavior, without, however, giving up the claim to search for regularities. (p. 10)

Subjectivity is indeed reinserted—but in modes of theorizing and testing that are still tethered to the original notion of a rational, individualist actor/"agent"; that is, in methodological individualism. That is why the language of "*social* preferences" is used: the marking of the noun establishes the condition as other than the normal case (as all marking does; consider "professor" vs. "woman professor").

It is only this individualistic formulation that renders the "agency" of experimental "subjects" similar to that of "subjects" acting in their own "lifeworld"—a context constructed intersubjectively with others in that culture and community, in stark contrast with an experimental laboratory setting. Put another way, methodological individualism is embedded in an experimental method that uses monetary payoffs to produce consequences that flow from experimental subjects' decisions. This individualistic reading of "agency" can be observed in the psychological language of "stimulus" and "motives" and in the economics terminology of "preferences" and "individual action."[4] The move in this experimental tradition toward "social preferences" is indicative of a mode of thinking and theorizing that is foreign to the vast range of interpretive research. It is a misreading of interpretive presuppositions and intentions to speak of such research as "looking into the subject's mind," as Kittel puts it (pp. 9, 10).

Secondly, concerning the "reality" of the experiment, Kittel writes that experimentalists "construct a world in the laboratory which is as real as a decision condition to the subjects taking part in the experiment as is the 'natural' social world to subjects in their everyday world" (p. 9). This argument has been advanced by experimental social scientists for a long time (e.g., Brewer 1985). As one of us claimed in a 1991 experimental study, the "monetary incentives are large enough that subjects regard the situation as a real one in which they are emotionally involved—occasionally quite involved, as indicated by reactions after the experiments" (Schwartz-Shea 1991: 53). This claim has been essential to obtaining grant monies to run such experiments. Granted, it is a "real" world in its consequences in the sense that subjects' payoffs are tied to their decisions in those experiments. Those monetary "incentives" as well as other portions of the experiment do attempt to "induce in subjects one specific understanding of the situation" (Kittel, p. 9); and experimenters use "manipulation checks" to assess the degree of success in that "inducement." Of course, all of that is more likely to succeed under conditions of no communication, as Kittel also knows. The usefulness of the models decreases dramatically when "subjects" get to communicate with one another—that is, to construct, intersubjectively, their understandings of the experimental conditions, which is a closer approximation to the lived experience studied "interpretively," but still a far cry from investigating how they think about and understand their own worlds.

In contrast, a study of intersubjective construction of meaning begins with human beings *already* and *always* enmeshed in their "natural" settings and language forms (and it is that humanly, *socially* constructed language that must be used to make sense of those monetary payoffs and experimental parameters/conditions; even absent communication, subjects draw on prior knowledge and experience, the meaning of which has been intersubjectively understood). Because they understand human agency as intersubjectively constructed and enacted (Fay 1996, ch. 2), interpretivists do not focus only on what is unique to a context, as the quote from Vaughan's work, above, demonstrates. Indeed, patterns and habits are part and parcel of the lifeworld, and interpretive researchers are often keen to identify not only ambiguities and contestations in meaning-making, but also the tacit assumptions at work in intersubjectively constructed worlds. Decisions, context, and agency are not "separated out," as in the experimental lab (so that the analyst can point to such and such a variable or condition as the linchpin), but are *of a piece* with the totality of the lifeworld; this is what the researcher endeavors to show, holistically, rendering subjects' actions "sensible" in their terms.

Whereas the foregoing differences of understanding concerning experimental and interpretive projects seem to be more ontological in character, others of Kittel's comments appear more epistemological. The issue here is the role and understanding of abduction and how it relates to both generating research questions and knowing when those questions have been answered. Put another way: Where do research questions come from? And how can "solving a puzzle" be "logical" if it does not involve either inductive or deductive reasoning (logic)? Although abduction has been around for quite some time, social scientists have not discussed it as much as the other logics until very recently, and therefore there is much room for exploration and discussion.

In his discussion of abduction and "puzzle finding," Kittel endorses Popper's view that this idea-generative process can never be understood via "logic," and that if we try to do so in any systematic way, that would "mean that certain possible lines of thinking are blinded out by the logic itself" (p. 10). We suspect that this comment emerges from a different understanding of "logic" (of inquiry) than the one we have been working with. For instance, Kittel does not recognize an "iterative logic" as logic. Such recognition may seem to create a contradiction, but if so, it is one *within* the practice of research: Researchers must be open, curious, and exploratory, on the one hand, and then, at the very next moment, critical! Doing research requires this kind of "schizophrenia"—which can either be ignored à la Popper, so that "discovery processes" (of possible explanations) are demarcated from falsifiability tests, or explicitly theorized, as Harding (1993) and other feminist philosophers have argued, in terms of researcher identity and societal location. The SSY chapter on where questions come from, if not providing a fully fleshed out logic (or theory), starts down that road by giving credence to prior experience—something that the Popperian position gives up on and rules out, which is why Harding (1993) calls that position "weak objectivity." In discussing why Popperian falsifiability does not fit interpretive logics of inquiry, SSY engages explanatory coherence, also describing numerous checks on researcher sense making (culled from interpretive research practices), although none of these may add up to the "definitive test" that a Popperian would see as requisite of science.

Further, Kittel claims that whereas both inductive and deductive logics involve a relationship between the known and the unknown, abductive reasoning has no logic (that is, no criterion for judging the status of a truth claim), "no anchor in something *commonly accepted* from which research reaches out" (p. 10, emphasis added). In our view, the puzzle is the unknown, its resolution the known. This formulation may seem to turn the order of the logic of inquiry on its head. In some sense, it appears more "logical" to start with what is known, rather than with what is puzzling. But consider: what makes something puzzling is its implied or explicit contrast with the researcher's expectations—which derive from what the researcher *knows*, albeit from and in a different context than the one under study. Still, what demands explanation is the puzzle, and it is in the logic of unraveling the puzzle—of figuring out what "conditions" make it more "natural" and less surprising—that researchers (try to) make their expectations more explicit. This is not strictly Popperian, perhaps, but we are not aware that conforming to Popperianism is the *sine qua non* of research logics. (On an argument for an expanded philosophy of reason, see Hawkesworth, 2006.) If there are other grounds for ruling out abduction as a logic of inquiry, we should like to learn what they are. Moreover, what criteria are "commonly accepted" as "anchors" varies by epistemic community—which is precisely the point of the book!

We disagree with Kittel's implication that processes of theorizing are separable from a logic of inquiry, let alone existing in a superior-subordinate relationship. Abduction usefully explains how interpretive research projects (at least) proceed, in ways that neither of the other two logics provides. At the same time, there is nothing to prevent interpretive researchers from using inductive logic—and in fact, we find compelling discussions in the methodological literature suggesting that abduction, induction, and deduction may work sequentially, in ways as yet to be sorted out.

One last point. We agree with Kittel's observation that interpretivists and experimentalists ask different questions and that neither alone covers the full range of possible research questions (p. 11). We emphasized in the book (p. 11) that we are pluralists, much more interested in figuring out what method(s) is (are) most useful for pursuing particular projects. That said, this formulation suggests that the framing of a research question is independent of the researcher's methodological presuppositions and orientations toward particular methods, something we think does not obtain: the two are far more intimately intertwined than most discussions of the methods- vs. question-driven research debate acknowledge. And while it is tempting to agree that interpretivists and experimentalists "generate complementary knowledge," their contrasting epistemological and ontological assumptions make that proposition questionable. But scholars like Kittel who seriously engage these conversations help to broaden and deepen methodological pluralism, and for that we are most appreciative.

Finally, Christian Bueger has captured the book's analytic strategy in its broadest context, that of a science studies (or sociology of knowledge) perspective that sees methods issues in the broader setting of the constitution of a discipline, its practices, and its knowledge claims. Pondering genres of painting as an analogy for interpretive research design is a provocative notion, something one of us previously essayed (Hatch and Yanow 2008) and which Bueger pushes further still in the context of teaching and learning new practices. Those of us who study practices know that much of what is learned and done is known tacitly. Curiously enough, it may be that this embedded tacit knowing requires the teaching of practices to draw on metaphors and analogies to communicate what cannot be said explicitly. Such tropes are rife in the music world, and Bueger not only engages in them himself, but has caught us out drawing on them extensively in trying to articulate aspects of research practicing that are largely known tacitly.

Informed by his own work in both science studies and practice theory, what Bueger writes provides us with an explanation of why our materials pushed back when we tried to write a "how-to" rules book. As many studies of practices demonstrate, reading and following the rules articulated in the operating manual does not enable one to ride the bicycle, to draw on Polanyi's oft-repeated example, because "we know much more than we can tell" (Polanyi 1966: 4). The requisite practice-knowing can only be learned in the practicing; it resists codification into a rule book. Would that we had applied this insight, familiar from our own substantive research, to this project when we started! We might have spared ourselves many months of teeth-gnashing and discarded manuscript pages.

That may explain why SSY is not more explicitly practice-centered in its theoretical engagements. We "practiced" the thinking through and writing of a methodology book; but (ironically, perhaps) we did not make its practice character itself explicit in its pages. Had we been writing an analysis of social science disciplinary practices for a science studies audience, that would have been an obvious move. Instead, we backgrounded that orientation, envisioning a reader for whom that point of reference would be less crucial than the direct engagements with methodological issues. This judgment lay, in part, in knowing that science and practice studies have not yet widely infiltrated political or other social sciences, being still fairly bounded in their engagements with physical and natural sciences, albeit widely interdisciplinary in their own right (for an exception in political science in addition to Bueger's own work, see Brandwein 2000, 2006).

Bueger is spot on in noting the centrality of "impact" to the latest incarnation of the UK faculty output assessment "exercise," something that took final form as we were writing the book and which he would be sensitive to given his institutional affiliation (another example of practices differing across states). But what constitutes impact in that system—and how one would witness and assess it—is proving to be interestingly problematic for political and other social sciences. Is it sufficient that one's work has been read by key governmental officials? Or does it have "impact" only if Parliament revises existing policy in light of that work? And how would one demonstrate either of these? In his presidential address to the American Sociological Association a few years ago, Michael Burawoy (2005) called for members to engage in more "public sociology"; around the same time, in U.S. political science circles disgruntled with the privileging of quantitative methods, discussions of a more "engaged political science" could also be heard. What has become of these? Were there parallels in other social science fields? The notion of research impact—or, more commonly phrased, utility—is central to (participatory) action research (PAR) in its various forms (Greenwood and Levin 2007); but PAR is contested terrain in many social sciences, even though some argue that it is the only ethical way of doing research in community settings, where academic knowledge can be put to the service of improving the lot of those we study (cf. Greenwood, in press). Too, as Orsini (2013) shows us, that "service" is itself not unproblematic, often arriving with assumptions about the constitution of those communities, including at the hands of Institutional Review Boards.

We have learned much from engaging with these five interlocutors. Should we have the opportunity to develop a second edition of SSY, we look forward to revisiting these comments.

Notes

[1] Such learning is often a happenstantial discovery, as many of the personal essays introducing each of the chapters in Yanow and Schwartz-Shea (2013/2006) demonstrate.

[2] These dates mark the publication of the King, Keohane, and Verba book and the first Qualitative Methods Section panels at APSA.

[3] For example, as we write this, the Department of Politics and Government at Israel's Ben Gurion University is at risk of being closed because a subcommittee of the Council of Higher Education deemed that it had too few "positivist" courses in its curriculum and faculty to teach them and pursue that sort of research.

[4] Morton and Williams (2010) may claim that it is "only the decision to act that is of interest to the experimentalist, not the subject as such" (Kittel, p. 9), but—at least in the past—there was a game-theoretic attention to sex-of-subject (e.g., Eckel and Grossman 1996, 1998, 2008), which makes subject characteristics central to such decisions. For a critique of this literature and its individualistic assumptions, see Schwartz-Shea (2002); see also Akerlof and Kranton (2001).

References

Akerlof, George A. and Rachel E. Kranton. 2001. *Identity Economics: How our Identities Shape our Work, Wages, and Well-being*. Princeton: Princeton University Press.

Atkinson, Paul. 1992. *Understanding Ethnographic Texts*. Newbury Park, CA: Sage.

Babbie, Earl. 2010. *The Practice of Social Research*, 12th ed. Belmont, CA: Wadsworth.

Becker, Howard S. 2009. "How to Find out How to do Qualitative Research." *International Journal of Communication* 3, 545–553.

Bell, Emma L. 1990. "The Bicultural Life Experience of Career-Oriented Black Women." *Journal of Organizational Behavior* 11, 459–477.

Brandwein, Pamela. 2000. "Disciplinary Structures and 'Winning' Arguments in Law and Courts Scholarship." *Law and Courts: Newsletter of the Law and Courts Section of the American Political Science Association* 10:3, 11–19.

Brandwein, Pamela. 2006. "Studying the Careers of Knowledge Claims: Bringing Science Studies to Legal Studies." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Dvora Yanow and Peregrine Schwartz-Shea, eds. (Armonk, NY: M.E. Sharpe), 228–243.

Brewer, Marilyn B. 1985. "Experimental Research and Social Policy: Must it be Rigor Versus Relevance?" *Journal of Social Issues* 41, 159–176.

Burawoy, Michael. 2005. "Provincializing the Social Sciences." In *The Politics of Method in the Human Sciences*. George Steinmetz, ed. (Durham, NC: Duke University Press), 508–525.

Campbell, Donald T. and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin.

Denzin, Norman K. and Yvonna S. Lincoln, eds. 2011. *The SAGE Handbook of Qualitative Research*. Newbury Park, CA: Sage.

Du Bois, W.E.B. 1990 [1903]. *The Souls of Black Folk*. New York: Vintage Books.

Eckel, Catherine and Philip J. Grossman. 1996. "The Relative Price of Fairness: Gender Differences in a Punishment Game." *Journal of Economic Behavior and Organizations* 30, 143–158.

Eckel, Catherine and Philip J. Grossman. 1998. "Are Women Less Selfish than Men? Evidence from Dictator Games." *The Economic Journal* 108, 726–735.

Eckel, Catherine C. and Philip J. Grossman. 2008. "Differences in the Economic Decisions of Men and Women: Experimental Evidence." In *Handbook of Experimental Economics Results*, vol. 1. Charles R. Plott and Vernon L. Smith, eds. (Amsterdam: North-Holland), 509–519.

Fay, Brian. 1996. *Contemporary Philosophy of Social Science: A Multicultural Approach*. Oxford: Blackwell.

Freire, Paulo. 1970. *Pedagogy of the Oppressed*. New York: Herder and Herder.

Freire, Paulo. 1973. *Education for Critical Consciousness*. New York: Seabury Press.

Greenwood, Davydd J. In press. "The Organization of Anthropology and Higher Education in the United States." In *A Companion to Organizational Anthropology*. D. Douglas Caulkins and Ann T. Jordan, eds. (Oxford: Wiley-Blackwell).

Greenwood, Davydd J. and Morten Levin. 2007. *Introduction to Action Research: Social Research for Social Change*, 2nd ed. Thousand Oaks, CA: Sage.
Harding, Sandra. 1993. "Rethinking Standpoint Epistemology: What is 'Strong Objectivity?'" In *Feminist Epistemologies*. Linda Alcoff and Elizabeth Potter, eds. (New York: Routledge), 49–82.
Hatch, Mary Jo and Dvora Yanow. 2008. "Methodology by Metaphor: Ways of Seeing in Painting and Research. *Organization Studies* 29, 23–44.
Hawkesworth, Mary. 2006. "Contending Conceptions of Science and Politics: Methodology and the Constitution of the Political." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Dvora Yanow and Peregrine Schwartz-Shea, eds. (Armonk, NY: M.E. Sharpe), 27–49.
King, Desmond. 1998. "The Politics of Social Research: Institutionalizing Public Funding Regimes in the United States and Britain." *British Journal of Political Science* 28:3, 415–444.
King, Gary, Robert Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.
Larsen, Otto N. 1991. *Milestones and Millstones: Social Science at the National Science Foundation, 1945–1992*. New Brunswick, NJ: Transaction Publishers.
Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab*. Cambridge: Cambridge University Press.
Orsini, Michael. 2013. "'May I See your Color-Coded Badge?' Reflections on Research with 'Vulnerable' Communities." In *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, 2nd ed. Dvora Yanow and Peregrine Schwartz-Shea, eds. (Armonk, NY: M.E. Sharpe), in press.
Polanyi, Michael. 1966. *The Tacit Dimension*. New York: Doubleday.
Schön, Donald A. 1987. *Educating the Reflective Practitioner.* San Francisco: Jossey-Bass.
Schwartz-Shea, Peregrine. 1991. "Understanding Subgroup Optimization: Experimental Evidence on Individual Choice and Group Processes." *Journal of Public Administration Research and Theory* 1:1, 49–73.
Schwartz-Shea, Peregrine. 2002. "Theorizing Gender for Experimental, Game Theory: Experiments with 'Sex Status' and 'Merit Status' in an Asymmetric Game." *Sex Roles* 47:7/8, 301–319.
Schwartz-Shea, Peregrine and Dvora Yanow. 2002. "'Reading' 'Methods' 'Texts': How Research Methods Texts Construct Political Science." *Political Research Quarterly* 55, 457–486.
Vaughan, Diane. 1996. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago: University of Chicago Press.
Voeten, Erik. 2012. "House Votes to Prohibit Political Science Funding." *The Monkey Cage* (10 May), http://themonkeycage.org/blog/2012/05/10/house-votes-to-prohibit-political-science-funding/.
Yanow, Dvora and Peregrine Schwartz-Shea, eds. 2013 [2006]. *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*, 2nd ed. Armonk, NY: ME Sharpe.

Symposium: Reports from the Multi-Method Research Frontier

A Case for Case Studies: A Multi-Method Strategy for Ecological Inference

Anne Meng
University of California, Berkeley
ameng@berkeley.edu

Brian Palmer-Rubin
University of California, Berkeley
brianpr@berkeley.edu

Ecological inference (EI) is the process of using aggregate data to infer discrete individual-level relationships of interest when individual-level data are not available (Cho and Manski 2008). EI is one of the longest standing challenges to quantitative social science research, yet scholars continue to debate the best statistical methods to deal with this unit-of-analysis problem (Freedman 1999, King 1997). In political science research, ecological inference commonly comes into play when a scholar is interested in inferring the voting behavior of some subgroup but electoral data that distinguish among subgroups are unavailable. This problem is especially salient in studies of voting behavior analyzing developing or weakly democratic countries where reliable individual- or precinct-level polling is rare.

We propose a strategy to improve the validity of ecological inference through a combination of quantitative analysis of aggregate-level units and within-unit case studies. First we provide an overview of the ecological-inference problem and the most prevalent statistical techniques that have been proposed to address it. Next, we outline a strategy of using qualitative case studies to gather evidence that allows the researcher to evaluate the assumptions that underlie ecological inference. Finally, we describe an application of these strategies to our own study of ethnic-minority voting in Mexico.[1]

The multi-method design that we propose builds on a recent wave of methodological scholarship that promotes case studies as a commonsense alternative or complement to complex statistical models for building and testing causal theories (Brady and Collier 2004, Freedman 1999, 2009, George and Bennett 2005, Gerring 2007, Mahoney 2010). We argue that qualitative case studies allow the researcher to collect fine-grained data on the micro-level mechanisms that underlie statistical relationships observed at the ecological level and thus constitute a useful complement to quantitative EI approaches.

Aggregate Data and Disaggregated-Level Inference

A researcher faces the ecological-inference problem whenever she is interested in making inferences about the behavior of a particular sub-population, or "population of interest" (POI), such as an ethnic group, social class, or voters registered to a particular party, yet only has data that is aggregated at a higher level. The resulting "aggregation bias" is "the effect of the information loss that occurs when individual-level data are aggregated…The problem is that in some aggregate data collections, the type of information loss may be selective, so that inferences that do not take this into account will be biased" (King 1997: 17). Ecological fallacies occur when researchers assume that relationships observed at the aggregate level are

Table 1: An Ecological Inference Hypothetical

	Democratic Candidate	Republican Candidate	Precinct Population (%)
Non-Hispanic Voters	? *(p)*	?	0.70
Hispanic Voters	? *(q)*	?	0.30
Total Votes (%)	0.80	0.20	

the consequence of corresponding disaggregate-level relationships (e.g., *counties* with more Hispanic voters vote disproportionately for Democratic candidates, therefore Hispanic *voters* vote disproportionately for Democratic candidates).

Studies of minority voting behavior commonly deal with ecological-inference problems. In recent decades, institutional mechanisms such as majority-minority districts and legislative quotas have been created to promote political participation by ethnic minorities and descriptive or substantive representation of minority interests in policymaking. A number of studies use data that are aggregated at the district level (or higher) to attempt to gauge whether these institutions affect minority voting (Goodnow and Moser 2012). Due to the secret ballot, it is impossible to directly observe how given ethnic groups vote, however, and even when exit polls are available, their results are often not reliable in racially charged elections (Grofman, Handley, and Niemi 1992). Challenges abound in studies of developing or weakly democratic countries (e.g., Blaydes 2011, Chandra 2004, Goodnow and Moser 2012) or historical societies (e.g., Childers 1983, Hamilton 1982), as exit polls are rare and electoral data are usually only available at relatively coarse levels of aggregation.

Quantitative Strategies: The Status Quo

A number of quantitative techniques have been developed to generate estimates about disaggregate-level relationships when only aggregate data are available. In this section, we outline the Method-of-Bounds approach and regression-based approaches and discuss the limitations of these quantitative strategies.

Formalizing the Problem

Studies of voting behavior that face the EI problem often resemble the following hypothetical:

A congressional precinct in the United States is composed of Hispanic and non-Hispanic voters. There are two candidates running for office: a Democrat and a Republican. We have electoral data on the vote share each candidate received, as well as the percent of the precinct that is Hispanic. Due to secret ballots, we do not know which candidate each individual voter voted for. What percent of Hispanic voters voted for the Democratic candidate?

Table 1 summarizes the EI problem. Although we have data on the margins of the table, we would like to fill in the question marks—most relevantly, the question mark in the lower-left cell.

The Method-of-Bounds approach is the simplest method for estimating the missing values, providing scholars with a first-cut tool to evaluate the plausibility of their disaggregate-level inferences (Duncan and Davis 1953). It relies on the intuition that vote shares cannot be negative; therefore each missing cell in Table 1 must be bounded by a [0,1] interval. Let p denote the fraction of the non-Hispanic population that voted for the Democratic candidate and let q denote the fraction of the Hispanic population that voted for the Democratic. Since we know the fraction of the total precinct population that voted for the Democratic candidate, as well as the fractions of the population that are Hispanic and non-Hispanic, we know that p and q must satisfy the following equation:

$$0.70(p) + 0.30(q) = 0.80$$

Furthermore, since we know that p and q must be bounded by [0,1], we can derive upper and lower bounds for these terms. Let p=1. Then $0.70(1) + 0.30(q) = 0.8$. We then calculate q=0.33. Now let p=0. Then $0.70(0) + 0.30(q) = 0.80$. We then calculate q=2.6. However, since q is also bounded by [0,1], we can conclude that the percent of the precinct's Hispanic population that voted for the Democratic candidate is bounded by [0.33,1]. Bounds that are wider represent more acute EI problems, and they tend to occur as the POI represents a smaller proportion of the aggregate unit or as the outcome of interest occurs in close to 50 percent of the aggregate unit.

Regression-Based Approaches

Due to the limited inferential power of the Method-of-Bounds approach, many researchers have turned to regression-based strategies (e.g., Blaydes 2011, Chandra 2004, Goodnow and Moser 2012). Commonly used regression models for ecological inference include the Goodman regression (1953), neighborhood model (Freedman, et al. 1991), and King's model (1997). All of these techniques center on the same intuition but differ in the complexity of the statistical model and underlying assumptions required to produce estimates. The Goodman model, the most basic ecological-regression approach, is set up as follows:

Returning to our voting example, let x denote the percent of the precinct population that is Hispanic, and y denote the percent of the total precinct vote share that went to the Democratic candidate. The subscript i in the equation indexes x, y, and ε by precincts

$$y_i = a + bx_i + \varepsilon_i$$

We can use Ordinary Least Squares (OLS) to estimate the parameters a and b. We interpret a as the estimate of the fraction of the non-Hispanic voters who voted for the Democratic candidate. Then $a + b$ represents the estimate of the fraction of the Hispanic voters who voted for the Democratic candidate.[2]

Notwithstanding its popularity, ecological regression is far from a silver bullet. The three most common regression-based techniques represent imperfect choices in dealing with the tradeoff between the plausibility of assumptions, the accuracy of estimates, and the simplicity of the model. Though the Goodman regression is easy to explain and understand, this technique is vulnerable to biases inherent to OLS regression and often produces nonsensical estimates of turnout or vote share that are not bounded by [0,1]. It also relies on an unrealistic "constancy assumption" that requires voting behavior across municipalities to be identical for minorities and non-minorities. The Freedman linear neighborhood model relaxes the constancy assumption by assuming that differences in voting behavior are independent of ethnicity, but this model is fundamentally based on the same OLS regression. Finally, while King's strategy offers certain advantages, such as the use of truncated distributions to compel respect for bounds and the incorporation of covariates to control for confounders, scholars debate whether this more complicated model produces more accurate estimates than its predecessors (Cho 1998; Freedman et al. 1991; King 1999).

A New Proposal: Using Case Studies to Evaluate Ecological Inference

Given the limitations in quantitative strategies, we propose a qualitative method for evaluating the plausibility of EI in causal research: carrying out case studies of units from a previously conducted large-N analysis wherein the scholar evaluates whether the observed ecological-level statistical relationship is driven by the hypothesized effect on the POI. These case studies have the goal of neither testing a hypothesis nor generating new hypotheses: Presumably, the researcher entered into the analysis with a plausible hypothesis to be tested and has moved onto these case studies because she has already found evidence on the aggregate level that supports this hypothesis.

Compared with statistical strategies, the case-study approach has the benefit of providing evidence about the specific mechanisms that underlie a causal relationship through the use of process tracing, which allows the scholar to "make strong inferences in just one or a few cases, based on one or a few pieces of the right kind of evidence" (Bennett 2008: 718). A drawback to the case-study approach, as with all small-n strat-egies, is the challenge to generalizability; it is impossible to prove that the POI-level relationships that one uncovers in a few studied units are the same that underlie the broader ecological-level relationship.

The goal of these case studies is to gather evidence that the ecological relationship that is observed in the original large-N analysis is driven by the hypothesized effect of the independent variable on the POI, rather than some other effect. For instance, in a study of minority voting, the scholar may seek to gauge whether a change in the *Democratic Party's campaign strategy* (independent variable) increased the *fraction of the Hispanic vote going to Democratic congressional candidates* (dependent variable). To this end, the scholar is looking for evidence to validate four criteria regarding the effect of the campaign strategy variable: causality, exclusivity, consistency, and non-interaction (Table 2).

These case studies do not necessarily require a large investment of time conducting interviews or archival research. These criteria can be evaluated in as few as two case studies—both cases that confirm the hypothesized relationship, one of which has a high prevalence of the POI and the other that has a low prevalence of the POI. (Of course, the greater the number of cases analyzed, the better argument the scholar can make that the EI is valid across the dataset.) The scholar will make many "causal process observations" (Brady and Collier 2004: 277–78) within each case by gathering evidence on at least two within-unit subgroups (POI and non-POI). It is appropriate to forgo studies of hypothesis-refuting cases because the goal is not to measure the causal effect of the independent variable, but rather to observe how it operates on different subgroups within an aggregate unit.

Table 2: Ecological Inference Criteria to be Verified through Case Studies

Criterion	Description	Importance for Causal Inference	Seeking Evidence to Show that…
Causality	The independent variable affects the POI as hypothesized	Necessary	The new campaign strategy caused Hispanic voters to vote more for Democratic Party candidates than in previous elections
Exclusivity	The independent variable does not affect the non-POI	Preferable	The new campaign strategy has no effect on the vote choice of non-Hispanic voters
Consistency	The causal effect is not influenced by the proportion of POI in the aggregated unit	Preferable	The effect of the campaign strategy on Hispanic voting does not vary based on the proportion of Hispanic voters in a unit
Non-Interaction	The independent variable does not interact with some other unobserved variable	Preferable	The effect of the campaign strategy does not vary among subsets of Hispanic voters (e.g., across social classes)

Three types of evidence can be used to evaluate these criteria:

> *Direct evidence of mechanism*: First-hand information, gained through interviews or observations, demonstrating how the independent variable affects the POI. (Example: The researcher attends a Democratic Party rally and observes that issues important to the Hispanic community are touted more than at rallies in the past.)

> *Bystander account of effect*: Media or other accounts that provide clues as to the voting behavior of the POI and non-POI. (Example: An interviewed election observer reports seeing more Hispanic attendees at Democratic Party campaign rallies after the strategy change.)

> *Disaggregated within-unit data*: Medium- to large-N data within aggregated units demonstrating that the POI-level outcome corresponds to the initially observed aggregate-level relationship. (Example: A survey that breaks down voting by race within a district shows that Hispanic voters voted in greater proportion for the Democratic Party candidate since the adoption of the new strategy.)

Applying the Case-Study Method: Minority-Concentrated Districts in Mexico

In this final section we describe the application of our case-study method to our study of minority voting in Mexico. We investigate whether minority-concentrated districts (MCDs)—an institution adopted in Mexico in 2005—influence the degree to which minority voters (the indigenous population) choose to vote for opposition parties.[3] Our study approximates a natural experiment by using a redistricting reform that took place between two midterm congressional elections (2003 and 2009) to estimate the causal effect of living in an MCD on voter behavior at the municipal level.[4] We collected electoral data from the *Instituto Federal Electoral* (IFE), Mexico's national electoral institute, on congressional election outcomes. These data were recorded on the municipal level, the finest level of disaggregation available. (Districts are made up of several municipalities.) Census data reporting the prevalence of indigenous and non-indigenous populations are also only available down to the municipal level. Therefore the aggregate unit for our large-N analysis is the municipality, and the POI is the block of indigenous voters in a given municipality.

Across the municipalities in our dataset—all of which are at least 50 percent indigenous—we found a significant negative effect of MCDs on the vote share for the dominant party (the PRI). These findings remain consistent across a variety of model specifications. Based on our quantitative evidence, and intuitions gleaned from newspaper accounts and previous research, we suspect that the creation of MCDs allowed opposition parties to penetrate populations that previously voted overwhelmingly for the dominant party. In this scenario, the purported mechanism is a change in opposition parties' electoral strategies following the adoption of MCDs: enterprising party leaders capitalized on these districts by nominating minority candidates and making targeted patronage appeals to indigenous communities.

We were satisfied to find a significant negative effect on the PRI's vote share on the aggregate level; however, we wanted to confirm that these findings reflect a change in indigenous voting behavior. *Are indigenous voters living in MCDs less likely to vote for the dominant party than indigenous voters not living in MCDs*? Answering this question and providing evidence of the mechanism that underlies the effect of MCDs on indigenous voting would make our causal inferences much stronger. Specifically, our quantitative analysis left four questions unanswered, corresponding to the four EI criteria described in the previous section:

1. *Causality*: Is the low vote share for the dominant parties in municipalities assigned to MCDs explained by the effect of MCDs on indigenous voting behavior?

2. *Exclusivity*: Did the assignment of municipalities to MCDs affect the voting behavior of non-indigenous voters in those municipalities? If so, did these populations respond in a way that strengthens our inference about indigenous voting behavior (voting less for opposition parties) or weakens this inference (voting more for opposition parties)?

3. *Consistency*: Did assignment to MCDs affect indigenous voting differently in municipalities with high indigenous populations (close to 100 percent) than it did in municipalities with relatively low indigenous populations (close to 50 percent)?

4. *Non-Interaction*: Did reassignment to MCDs have different effects among various subgroups within municipal indigenous populations (e.g., indigenous Catholics vs. indigenous Protestants or poor vs. non-poor indigenous)?

In order to respond to these questions, we undertook case studies that focused primarily on two municipalities in Chiapas, a highly indigenous state in southern Mexico. The two municipalities, Ocotepec and Simojovel, both received treatment (were assigned to MCDs) and both experienced significant decreases in the vote shares for the PRI from 2003 to 2009 (in Ocotepec, from 46 percent in 2003 to 29 percent in 2009 and in Simojovel from 52 percent in 2003 to 15 percent in 2009). They also belong to the higher and lower ends of the spectrum in the prevalence of indigenous populations in our dataset: Ocotepec's population is 97 percent indigenous and Simojovel's population is 71 percent indigenous. We conducted two methods of qualitative data collection: first, interviews with party leaders and indigenous authorities who reported on the mechanism underlying the relationship between MCDs and indigenous voting; and second, second-hand observations of the electoral behavior of the POI and non-POI, acquired through searches through newspaper archives and interviews with informed observers (academics and NGO workers).

The first three questions are oriented toward understanding the effect of the treatment on indigenous and non-indigenous voting. Is there evidence that reassignment to MCDs increased indigenous voting for opposition parties? Did this treatment have any effect on non-indigenous vote choice? Do these effects vary based on the proportion of indigenous voters in a municipality?

To address these questions, we interviewed indigenous authorities and leaders of the once-dominant party (PRI) and the main opposition party (PRD) in Chiapas, and gathered newspaper accounts of campaign activities by these parties. Interviews with party leaders uncovered affirmative evidence of the mechanism: PRD leaders reported that they targeted their campaigns to indigenous communities following the redistricting by nominating indigenous candidates and promoting social programs and infrastructure spending for indigenous communities. PRI leaders reported very little change in their campaign strategies following the redistricting. These reports were bolstered by newspaper reports of the intensification of indigenous-targeted appeals by the PRD in 2009. Furthermore, we found no evidence that these mechanisms operated differently in Ocotepec (our high-POI case) and Simojovel (our low-POI case).

The final question has to do with different treatment effects across subgroups of indigenous voters within a municipality. This issue is important because it is possible that the hypothesized effect on indigenous voters only occurs among a subset of the indigenous, which would suggest that we modify our initial hypothesis. For instance, scholars of indigenous activism have observed that Catholic and Protestant churches with indigenous congregations promote different forms of mobilization and respond to partisan appeals in different ways (Palmer-Rubin 2011, Trejo 2009). Reassignment to MCDs could also be more influential among poorer indigenous populations, which are likely to be more prone to clientelistic appeals than relatively better-off populations. To address these potential modifications to our findings, we interviewed subjects who were able to comment on the effect of MCDs on the electoral behavior of indigenous voters of different religions and different economic strata. Interviews with party leaders as well as both Catholic and Protestant indigenous authorities suggested that all indigenous populations were targeted by the PRD after 2005, regardless of religion. Both predominately Catholic and predominately Protestant indigenous organizations formed alliances with the PRD (although the switch to the PRD appeared to be more widespread by Catholic groups). Through interviews of leaders representing both relatively well-off (more urban) indigenous populations and relatively poor (more rural) indigenous populations, we found no compelling evidence that poverty mediated the effect of MCDs on indigenous vote choice. We also found newspaper reports of indigenous-targeted patronage in both rural and urban areas.

In sum, the case studies provided evidence to bolster our quantitative findings. The aggregate-level finding—that municipalities reassigned to MCDs demonstrated lower vote shares for the PRI than similar municipalities that were not reassigned—is grounded in the POI-level relationship that our hypothesis predicts, at least in the municipalities where we conducted case studies. We observed the mechanism that underlies the effect of MCDs on indigenous voting: namely, shifts in opposition-party strategies. We also found no compelling evidence that MCDs significantly affected the vote choice of non-indigenous voters, nor that the effect of MCDs on indigenous voting was mediated by the proportion of the municipality that is indigenous or by some other variable. Of course, we cannot be sure that these findings are generalizable to all other municipalities in our dataset. However, our case studies lend a great deal of plausibility to our causal argument by demonstrating that at least in a couple of cases, the hypothesized POI-level effect took place.

Concluding Thoughts

This paper addresses the challenges a researcher faces when working with aggregate data if she would like to make arguments about the behavior of some population at a lower level of aggregation. In response to the limitations of statistical techniques that have been proposed to alleviate the ecological-inference problem, we develop a case-study approach. Our proposed strategy does not differ markedly from other multi-method approaches that employ case studies to identify mechanisms that underlie causal relationships identified through large-N analysis. However, our approach is tailored to assist the scholar in observing the causal effect of an explanatory variable on the population-of-interest and to detect aggregation bias that may threaten the large-N findings. Compared with commonly used statistical strategies for addressing ecological inference problems, the primary advantage of case studies is that they provide the scholar with original evidence of mechanisms that underpin the hypothesized causal relationships.

While we believe that this strategy offers certain advantages over statistical approaches, we agree with Freedman's (1999: 4030) prediction that "the problems of confounding and aggregation bias…are unlikely to be resolved in the proximate future." Nonetheless, research using aggregate-level data will continue to be important and common, given the predominance of ecological-level data. Thus, in the interest of reaching the most defensible causal claims, scholars are advised to use all the evidence, both quantitative and qualitative, that it is practical to collect.

Notes

[1] The research project for which we developed this multi-method approach capitalizes on a redistricting reform undertaken between two congressional elections in Mexico to approximate a natural experiment, allowing us to measure the effect of indigenous-concentrated districts on voting outcomes. Due to space constraints, we do not go into detail about our large-N identification strategy here. The examples in this article, instead, resemble EI problems that would occur in the context of any large-N research that uses aggregate-level data.

[2] To see this, recall that a is the height of the regression line at x=0, and $a + b$ is the height of the regression line at x=1.

[3] We define the opposition parties in 2009 as the PAN and the PRD. Although the PAN won the presidency in Mexico in 2000, the PRI is the historically dominant party that ruled all of these municipalities before the insertion of the PAN and PRD in the 1990s.

[4] Treated units are municipalities that were redistricted to MCDs, and control cases are similar municipalities that were not redistricted to MCDs. To ensure balance between the two groups, we matched on

redistricting criteria, socioeconomic variables, and pre-treatment electoral outcomes.

References

Bennett, Andrew. 2008. "Process Tracing: A Bayesian Perspective." In *The Oxford Handbook of Political Methodology*. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. (New York: Oxford University Press), 702–721.

Blaydes, Lisa. 2011. *Elections and Distributive Politics in Mubarak's Egypt*. New York: Cambridge University Press.

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

Chandra, Kanchan. 2004. *Why Ethnic Parties Succeed*. New York: Cambridge University Press.

Childers, Thomas. 1983. *The Nazi Voter: The Social Foundations of Fascism in Germany, 1919-1933*. Chapel Hill: University of North Carolina Press.

Cho, Wendy K. Tam. 1998. "If the Assumption Fits…: A Comment on the King Ecological Inference Solution." *Political Analysis* 7:1, 143–163.

Cho, Wendy and Charles Manski. 2008. "Cross-Level/Ecological Inference." In *The Oxford Handbook of Political Methodology*. Janet M. Box-Steffensmeier, Henry E. Brady, and David Collier, eds. (New York: Oxford University Press), 547–569.

Duncan, Otis and Beverly Davis. 1953. "An Alternative to Ecological Correlation." *American Sociological Review* 18:6, 665–666.

Freedman, David. 1999. "Ecological Inference and the Ecological Fallacy." In *International Encyclopedia of the Social and Behavioral Sciences*. Neil J. Smelser and Paul B. Baltes, eds. (Oxford: Elsevier), 4027–4030.

Freedman, David. 2009. "On 'Solutions' to the Ecological Inference Problem." In *Statistical Models and Causal Inference*. David Collier, Jasjeet Sekhon, and Philip Stark, eds. (New York: Cambridge University Press), 83–96.

Freedman, David, Stephen P. Klein, Jerome Sacks, Charles A. Smyth, and Charles G. Everett. 1991. "Ecological Regression and Voting Rights." *Evaluation Review* 15:87, 673–711.

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

Gerring, John. 2007. *Case Study Research: Principles and Practices*. New York: Cambridge University Press.

Goodman, Leo 1953. "Ecological Regression and the Behavior of Individuals." *American Sociological Review* 18:6, 663–664.

Goodnow, Regina and Robert Moser. 2012. "Layers of Ethnicity: The Effects of Ethnic Federalism, Majority-Minority Districts, and Minority Concentration on the Electoral Success of Ethnic Minorities in Russia." *Comparative Political Studies* 45:2, 167–193.

Grofman, Bernard, Lisa Handley, and Richard G. Niemi. 1992. *Minority Representation and the Quest for Voting Equality*. New York: Cambridge University Press.

Hamilton, Richard F. 1982. *Who Voted for Hitler?* Princeton: Princeton University Press.

King, Gary. 1997. *A Solution to the Ecological Inference Problem*. Princeton: Princeton University Press.

King, Gary. 1999. "The Future of Ecological Inference Research: A Reply to Freedman, et al." *Journal of the American Statistical Association* 94:445, 352–355.

Mahoney, James. 2010. "After KKV: The New Methodology of Qualitative Research." *World Politics* 62:1, 120–147.

Meng, Anne and Brian Palmer-Rubin. 2012. "Ethnic Voting and Electoral Competition: The Effects of Minority-Concentrated Districts in Mexico." Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Palmer-Rubin, Brian. 2011. "From Frame to Campaign: Indigeneity in the Political Arena." Paper presented at the Annual Meeting of the American Political Science Association, Seattle.

Trejo, Guillermo. 2009. "Religious Competition and Ethnic Mobilization in Latin America: Why the Catholic Church Promotes Indigenous Movements in Mexico." *American Political Science Review* 103:3, 323–342.

More than the Sum of the Parts: Nested Analysis in Action

Craig M. Kauffman
University of Oregon
ckauffma@uoregon.edu

After reading Lieberman's (2005) article on nested analysis, I was eager to test the purported benefits by adopting this multi-method approach in my dissertation research (Kauffman 2012). After using it, I am even more convinced of Lieberman's assertion that both quantitative and qualitative methodologies "can inform each other to the extent that the analytic payoff is greater than the sum of the parts" (2005: 436). My purpose here is to provide an example of nested analysis in action to illustrate how quantitative and qualitative methods may be used in supplementary and complementary ways at various stages of a research design, from model specification, to case selection, to data analysis.

The puzzle that inspired my research was why some local governments in Ecuador pursued and successfully implemented a set of reforms known as Integrated Watershed Management (e.g., Gregersen, Ffolliott, and Brooks 2007), while others did not. Specifically, I was interested in the creation of two institutions designed to better manage local watershed resources: (1) a participatory decision-making structure in which multiple stakeholders jointly identify needs and develop watershed management plans and activities, and (2) a stable financing mechanism using local revenue sources to fund watershed management activities. Through previous fieldwork, I had seen how these reforms could change production practices, alter socio-political relations, and create new governance arrangements with characteristics commonly equated with good governance. I found the reforms curious in part because, while many attempts failed, they sometimes succeeded in unlikely places—cantons with poorly performing governments, high levels of poverty and inequality, corruption, clientelism, social and ethnic conflict, and little history of social organizing. I initially saw this as an opportunity to test competing hypotheses regarding improvements in local government performance.

Having identified my general research question, I used both quantitative and qualitative methods to specify hypotheses and a model, assess the validity of indicators, evaluate rival explanations, check for omitted variable bias, and search for causal mechanisms behind correlations. Following the nested analysis approach, my research design involved a two-step process: (1) a statistical analysis of all 221 Ecuadorian cantons, and (2) a qualitative comparison of six case studies.

Model Specification

For the statistical analysis I used a logistic regression model to test competing explanations of local governance reform based on an original dataset I compiled. The dependent variable was dichotomous, indicating whether or not natural resource management reforms were attempted in each canton. While my interest was in watershed management, I looked at the broader category of natural resource management reforms in order to have a population of cases of sufficient size and variation to make statistically meaningful comparisons. Between 1997 and 2008 roughly half of Ecuador's municipalities (108 of 221) pursued reforms to better manage natural resources. An assumption of the study is that the relationships applying to natural resource management reforms generally also apply to the sub-category of watershed management reforms. My list of reforming municipalities came from Ecuadorian scholars and practitioners who catalogued innovative local government reforms in Ecuador since the 1990s (e.g., Ramón and Torres 2004; Asociación de Municipalidades Ecuatorianas 2004, 2008; Garzón 2009), as well as information from personal interviews. The dataset also included indicators of explanations commonly found in the literatures on local government performance and natural resource management. These include measures of political competition, political organization, citizen participation, social trust, technical assistance, central-local government relationships, ecosystem type, level of economic development, and demographic variables, among others. As expected, some explanations were ruled out while several variables had a significant effect (e.g., measures of participation, trust, and ecosystem type).

While statistical analysis could assess the strength of partial explanations, it could not illuminate the causal mechanisms or describe how different variables interacted to produce success. For this I turned to within-case analysis. Because my interest was initially in model testing, my first step in the qualitative portion was to conduct process tracing in a typical success case to check the validity of my indicators and uncover the causal mechanisms behind the correlations in the statistical study. To identify a typical success, I calculated the predicted probabilities of success for all 221 cantons and compared this with their actual scores. I then selected Tungurahua, the case with the highest predicted probability that had implemented the Integrated Watershed Management reforms described above.

By necessity, the indicator for my dependent variable in the statistical study had been relatively crude—a dummy variable that indicated reform or no reform. But government reform is rarely black or white. There are many types and degrees of success that must be empirically verified. One advantage of small-N analysis is the ability to develop more nuanced and precise indicators of success. To measure the extent to which Integrated Watershed Management reforms occurred, I developed an index of 15 indicators. These indicators measured the degree to which there was consensus among various stakeholders on the problem and proposed solutions, the degree to which institutions were created and action plans implemented, the degree to which stakeholder groups participated in the process, and the degree to which the reforms were institutionalized in a way that provided lock-in effects. Cases were scored on a scale ranging from 0 (low success) to 15 (high success). For each of the 15 indicators, a case was awarded 1 point if the condition was present and 0 points if it was absent. A half point was awarded when a condition was only weakly present or was present at one time but not sustained.

Four months of field research in Tungurahua produced surprising results. On one hand, it confirmed that Tungurahua was indeed an example of successful Integrated Watershed Management reform; it scored 13.5 out of 15 on the index. However, process tracing revealed that, instead of the structural explanations espoused in the literature and represented in my statistical model, the main story was the role played by transnational networks in setting the agenda for reform and mobilizing coalitions of local stakeholders with the motivation and capacity to pursue and implement reforms. In short, my qualitative analysis suggested that my statistical model was misspecified and potentially contained spurious correlations.

Navigating Between Model Testing and Model Building

At this point my endeavor turned from model testing to model building. The Tungurahua case presented a new hypothesis—that variation in the implementation of Ecuador's local natural resource management reforms (including watershed management reforms) was explained by the pattern of network ties and strategies employed by transnational coalitions advocating these reforms. I ultimately wanted to test this hypothesis through comparative case studies. But first I assessed the new hypothesis using a statistical model. I wanted more confidence that Tungurahua was not an anomaly and that the investment of time and resources needed to conduct case studies would be worthwhile.

I revised my quantitative study by adding a new independent variable, the presence of transnational environmental networks. The theory was that the presence of transnational environmental networks made reform attempts more likely since international actors carry new ideas and practices regarding natural resource management to cantons through these network ties. These networks also provide important resources to local advocates who embrace the reforms and seek to implement them. I measured the presence of transnational environmental networks using the amount of external environmental aid received by each canton from international actors. While environmental aid was not a perfect proxy for transnational networks, it was the best data available that would allow a large-N analysis. I assumed that if international organizations provided aid to particular cantons, particularly in a sustained manner, representatives of those organizations were interacting regularly with local actors to implement environmental programs. Theoretically, this regular interaction created ties between local actors and a transnational environmental network through which ideas and resources flowed. I also assumed that international environmental actors provided more aid to the cantons with which they had stronger ties, and a lack of aid reflected a lack of network ties.

The specific indicator for my transnational network variable was the average amount of environmental aid (in millions of dollars) donated by international actors in each canton between 2007 and 2009, as calculated by the Ecuadorian Agency for International Cooperation (AGECI). This indicator included the money spent by multilateral organizations (e.g., the World Bank and the European Union), bilateral development organizations (e.g., U.S. Agency for International Development and German Technical Cooperation), international NGOs (e.g. The Nature Conservancy and Conservation International), and private businesses on local environmental programs. A key assumption was that levels of environmental aid during 2007–2009 (the only years for which data were available) were consistent with levels of similar aid during the previous seven years, during which most reform processes began. My six case studies later compensated for this weakness by providing evidence that the financing of Integrated Watershed Management reforms in the late 2000s typically followed a decade of similar funding on natural resource management programs.

It is beyond the scope of this article to describe all the details of my quantitative analysis. However, I want to share a few results that particularly influenced the qualitative portion of my project. The results supported the hypothesis regarding the importance of transnational networks. Controlling for all other explanations, the presence of transnational environmental networks (as measured through environmental aid) had both a large effect and was highly significant ($p = .002$). Various robustness tests confirmed these results. Many of the variables representing alternative explanations lost their significance, suggesting there may have been spurious correlations. In short, my new model allowed me to eliminate many alternative explanations and justified a narrower focus on transnational networks as the main explanatory variable of interest in my case comparisons.

The best fitting model—the one that combined parsimony with predictive power—contained three explanatory variables: the presence of transnational environmental networks (indicated by environmental aid), social trust (based on survey results conducted by the LAPOP project; Seligson et al. 2006), and a dummy variable, Sierra, indicting whether a canton is in Ecuador's mountainous Sierra region. I included this variable because the region is argued to have several conditions conducive to reform. Ecuador's Sierra region is commonly believed to have a more innovative political culture conducive to change, in part due to the presence of highly organized indigenous movements. Control over natural resources has been one of the banners around which indigenous movements have mobilized in recent decades. The region's ecosystem and pattern of water distribution are also important; water tends to be more scarce and unevenly distributed in the Sierra, particularly compared with the Amazon region.

Interestingly, the average marginal effect of environmental aid was very different between cantons inside and outside the Sierra region. Outside the region, a one standard deviation increase in environmental aid boosted the probability of reform attempts by 62 percent. By contrast, the average marginal effect of environmental aid inside the Sierra region was 34 percent. The lower marginal effect inside the Sierra region makes sense given the region's many propitious conditions. The fact that environmental aid's effect was much greater outside the Sierra than inside suggests that transnational environmental networks were able to compensate for the lack of propitious conditions in regions like the Amazon, which lacks a political culture conducive to social organizing and the water scarcity that might induce reforms. In this way, the quantitative analysis produced new hypotheses to be further tested through qualitative methods.

Case Selection for Comparative Case Studies

Once quantitative methods indicated that transnational networks mattered, I turned next to designing case comparisons that could reveal how and why they mattered. I selected six cases in four steps, using both quantitative and qualitative methods. First, I identified cases where similar Integrated Watershed Management reforms were attempted to ensure comparability. This narrowed the list to 26 potential candidates. Second, I selected six cases that ensured instances of both successful and unsuccessful reform, based on the 15-point index described above. Third, I used quantitative methods to select cases that were both "typical" and "deviant" (Seawright and Gerring 2008) from the perspective of existing explanations of local government reform. I did this using the predicted probabilities calculated from my original statistical model that excluded transnational networks.

Using these criteria, I selected six cases grouped into four categories: typical-success, typical-failure, deviant-success, and deviant-failure (see Figure 1). Two cases, Tungurahua and Celica, were predicted to successfully reform and did. Similarly, Zamora's efforts were predicted not to succeed, and they were less successful. I chose these "typical" cases to help identify the micro-causal processes leading to successful reform (George and Bennett 2005). By contrast, two cases, Pastaza and Ibarra, were strongly predicted to reform, but their attempts were unsuccessful. The sixth case, El Chaco, was predicted not to succeed, but it did. These "deviant" cases are useful for uncovering new explanations and causal mechanisms (Seawright and Gerring 2008). Comparing these four case types permitted several forms of analysis. The typical cases allowed me to evaluate the importance of transnational networks vis-à-vis alternative explanations and look for evidence that correlations in my original statistical model were spurious. The deviant cases allowed me to test whether the mechanisms relating to transnational networks facilitated reform even where propitious conditions did not exist (deviant-success), and whether their absence explained failure to reform even where propitious conditions did exist (deviant-failure). Together, these case comparisons constituted a harder test to provide more confidence in the generalizability of my theoretical model of how transnational networks explain variation in the success of local Integrated Watershed Management reforms.

Finally, I also selected these cases to control for alternative explanations of watershed management reform (see Table 1). These include ecosystem characteristics, quantity of avail-

Figure 1: Four Case Types

	Typical	Deviant
Successful Reform	Tungurahua, Celica	El Chaco
Unsuccessful Reform	Zamora	Ibarra, Pastaza

Table 1: Case Study Characteristics

Case	Success	Ecosystem	Water Scarcity	Landowner Location	Indigenous Stakeholders	Population*	Poverty Rate*
Tungurahua	High	Andean	High	Watershed	Yes	329,856	50%
Celica	High	Andean	High	City	No	14,468	76%
El Chaco	High	Amazon	Low	Watershed	No	7,960	65%
Zamora	Mixed	Amazon	Low	City	No	25,510	61%
Pastaza	Low	Amazon	Low	City	No	62,016	67%
Ibarra	Low	Andean	High	Watershed	Yes	181,175	40%

*Source: 2010 Census, Integrated System of Social Indicators in Ecuador (SIISE), www.siise.gob.ec

able water, land use patterns (e.g., reflected by whether or not landowners live in the watershed, which affects their interest in and use of watershed resources), and demographic information such as the poverty rate, population size, and whether or not indigenous groups are among the watershed's stakeholders. This latter variable is important because indigenous groups tend to have higher levels of social organization, potentially making reform attempts easier. The cases also controlled for local political and economic conditions. Both successful and failed cases contained variation in political organization (e.g., relative party affiliations among mayors, municipal council persons, and the national government); the level and form of social organization (e.g., the existence of irrigation councils, indigenous movements, and environmental and development associations); as well as the particular stakeholders involved (e.g., whether hydroelectric companies, indigenous groups, or biodiversity conservationists, among others, were present). In sum, both quantitative and qualitative methods contributed to my case selection in complementary ways and allowed me to control for a wider array of alternative explanations.

Data Analysis

Having identified transnational advocacy networks as my main variable of interest, I needed both qualitative and quantitative methods to analyze if and why these networks explained variation in Integrated Watershed Management reform. The two methodological approaches were complementary in that they addressed different ways in which transnational networks matter. Transnational networks are a useful analytical concept because they combine elements of both structure and agency (Keck and Sikkink 1998). As a structure, they pattern the interactions and relationships among individuals and organizations. As an agent, they articulate and advocate specific policy changes. Quantitative methods can be useful for examining the structure of networks. For example, the positive correlation between environmental aid and reform attempts implied that the pattern of ties within transnational environmental networks—specifically, their geographic reach into different cantons—explains why reforms were attempted in some cantons but not others. Qualitative methods supplemented this quantitative analysis by verifying the validity of indicators and the correlations among them, as well as revealing the causal mechanism behind the correlations. Through network analysis I documented the variation in transnational network connections linking local watershed stakeholders to advocates of Integrated Watershed Management reform, as well as the information and resources flowing through these network ties. Process tracing revealed how differences in network ties produced variation in the reform processes and their outcomes.

Qualitative methods also complemented the quantitative analysis by analyzing the agency of transnational advocacy coalitions. Understanding the agential component of transnational networks was important because, while the diffusion of ideas and practices through network ties provided local actors the opportunity to reform, it did not guarantee success. Qualitative methods were needed to test hypotheses relating to the strategies used by transnational coalitions of advocates to explain why some reform attempts were more successful than others. Without going in to the details of my case comparisons, I used process tracing, network analysis and framing analysis to show that the relative success of reform attempts was explained by variation in the network construction and framing strategies employed in each reform campaign.

Another example of how combining quantitative and qualitative methodologies strengthens data analysis is the way both methodologies were used to overcome the endogeneity problem inherent in my statistical study. While there was undoubtedly a strong and robust correlation between external environmental aid and the tendency to attempt natural resource management reforms, the statistical model could not determine the direction of the causal arrow. It is possible that international actors chose to invest in cantons where they perceived reform attempts to be most likely, or where reforms were already initiated. The argument that external environmental aid led to reforms rather than the other way around rests on two crucial assumptions. First, that the idea to reform came from international rather than local actors. Second, that international environmental actors decided where to invest using criteria that were independent of local actors' desire to reform.

The six case studies provided strong evidence that these assumptions were valid. Process tracing showed that while local actors identified problems related to natural resource management, in each case it was external actors who identified IWM reforms as the best solution to these problems. Furthermore, these international actors often worked for years persuading local politicians and social groups before the reforms were implemented. The case comparisons demonstrated how Integrated Watershed Management reforms grew out of environmental projects financed by international actors a decade earlier. Furthermore, they showed that the financing of these reforms in the late 2000s generally followed many years of similar funding, increasing confidence that my environmental aid indicator was valid.

The counter argument—that the idea to reform came primarily from local actors—was tested through quantitative analysis. A dummy variable in my dataset, *Decentralization Requested*, indicated those cantons where local governments requested environmental decentralization, showing a desire to take on additional environmental management responsibilities. It is reasonable to expect that these cantons would be more likely to pursue natural resource management reforms if such reforms primarily emerge from local actors. If this argument were true, we would expect the variable *Decentralization Requested* to be positively correlated with reforms being attempted. However, the data did not support this.

Regression analysis similarly undermined the argument that transnational actors allocated environmental aid based on which cantons were most likely to reform. If this were true, we would expect *Decentralization Requested* to be a significant predictor of environmental aid. That is, we would expect transnational actors to steer their investments toward local governments that expressed an interest in taking on environmental management responsibilities. Again, the data do not bear this out.

The case comparisons provided further evidence that international actors selected their sites based primarily on criteria other than local desires to reform. While local political will was important, the evidence showed that international actors financing natural resource management activities selected their sites based primarily on the importance and vulnerability of the ecosystem. Other criteria included social needs, such as poverty and water conflicts.

In sum, both quantitative and qualitative data indicated that the assumptions behind interpreting environmental aid as an explanation for reform attempts were valid. All the evidence showed international actors chose where to promote environmental programs based on factors that were independent of a canton's inherent propensity to pursue reforms. It also showed that the impetus to pursue reforms came primarily from international rather than local actors. Together, the evidence indicated that the likelihood of an endogeneity problem was small.

Conclusion

I initially viewed nested analysis as a linear process in which quantitative and qualitative methods were used sequentially, each adding different pieces to a single puzzle. In some sense this is true—seen, for example, in the way each methodology provided evidence related to a different aspect of my explanatory variable and helped resolve issues of endogeneity. However, I have come to realize that the real power of combining quantitative and qualitative methodologies is the way they can inform each other through an interactive process that produces analytic insights greater than the sum of its parts. For me, the power of this interaction was most evident in the way it informed my model specification and case selection, producing a research design that would have been unlikely had I used only one methodology or the other.

References

Asociación de Municipalidades Ecuatorianas. 2004. "Premio a las Mejores Prácticas de la Gestión Pública Seccional-Primera Edición (2003–2004)." Quito, Ecuador: AME.

Asociación de Municipalidades Ecuatorianas. 2008. "Reportes de Gestión-Premio MPS." Quito, Ecuador: AME.

Chen, Xiao, Phil Ender, Michael Mitchell, and Christine Wells. 2007. "Logistic Regression with Stata." Los Angeles: UCLA: Academic Technology Services, Statistical Consulting Group.

Garzón, Andrea. 2009. "Estado de la Acción sobre los Mecanismos de Financiamiento de la Protección y Recuperación de los Servicios Ambientales Hidrológicos." In *Informe Final de la Síntesis Regional Sobre Servicios Ambiental Hídricos en Los Andes*. Lima, Peru: Consorcio para el Desarrollo Sostenible de la Ecorregión Andina.

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

Gregersen, Hans M., Peter F. Ffolliott, and Kenneth N. Brooks. 2007. *Integrated Watershed Management: Connecting People to their Land and Water*. Cambridge: CABI Publishing.

Kauffman, Craig. 2012. "Global Governors and Local Governance: Transnational Networks and the Decentralization of Watershed

Management in Ecuador." Ph.D. diss. The George Washington University, Washington, DC.

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca, NY: Cornell University Press.

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

Ramón, Galo and Victor Hugo Torres. 2004. *El Desarrollo Local en el Ecuador: Historia, Actores y Métodos*. Quito, Ecuador: Abya-Yala.

Seawright, Jason and John Gerring. 2008. "Case Selection Techniques in Case Study Research: A Menu of Qualitative and Quantitative Options." *Political Research Quarterly* 61:2, 294–308.

Seligson, Mitchell A., Juan Carlos Donoso, Daniel Moreno, Diana Orcés, and Vivian Schwarz-Blum. 2006. "Democracy Audit: Ecuador, 2006." Nashville: Vanderbuilt University.

A Horse of a Different Color: New Ways to Study the Making of Citizens

Calvert W. Jones
Yale University
calvert.jones@yale.edu

To what extent can contemporary state leaders use social engineering[1] to produce the citizens they want? Under what conditions do they succeed or fail? The question of how states shape citizens is a classic one, yet it is also one that has taken on additional complexity and alternate "colors" in the contemporary era. Most scholars understandably link the challenge of citizen-building with the effort to foster a common national identity, among the first and most fundamental tasks in state-building. Today, however, and in contrast to earlier eras, billions of people already recognize themselves as citizens of a state, and take its authority for granted. Shifts in the international system, such as the decline of major war and intensification of global economic competition, may also be influencing state leaders' priorities. In these conditions, the willingness and ability of citizens to fight in battle for the state may be less important than their willingness and ability to "fight" in markets by contributing to their nations' economies.

My dissertation ("Bedouins into Bourgeois") investigates how state leaders are re-interpreting the challenge of citizen-building, what outcomes they are achieving, and the conditions for their success and failure. To answer these questions, I use the United Arab Emirates (UAE) as a data-rich empirical laboratory for the contemporary "making of citizens." As in many countries, UAE state leaders are struggling to build more *entrepreneurial citizens*: individuals who will demand less from the state in terms of social and economic welfare, while showing a greater willingness to contribute to market-driven economies and take risks to build vibrant private sectors. With a strategy of "soft" social engineering through education reform, state-sponsored spectacles, and other instruments, they are hoping to create what they see as a more enlightened citizenry, motivated to achieve but still loyal, without provoking unrest of the sort we have seen by way of the Arab Spring as well as the Greek crisis. In the international community, these efforts to build a "new Arab citizen," one who is better educated and equipped to compete in the global economy and less susceptible to radicalism, have been heralded as a way forward for the troubled region (UNDP 2003; World Bank 2007; Faour and Muasher 2011).

Based on my findings, however, "pro-globalization" social engineering is failing in some intriguing and unexpected ways: instead of building entrepreneurial citizens, the data suggest it is giving rise to "*super*-entitled citizens," cultivating youth with expectations of elite status and little interest in private sector work. To explain this outcome, I highlight the role of the middle-men carrying out the social engineering campaign on behalf of autocrats. Largely Western-educated professionals embedded in transnational networks of expertise, they operate in a political context, I argue, that produces behaviors on their part that undermine macro-level strategy. The counter-intuitive result has been an intensified culture of rentierism among the Arab youth who are the target of the campaign, quite the opposite of what was intended.

To substantiate these arguments, I use several methods uncommon in the existing literature, aiming to bring new approaches and data to bear the classic question of how states shape citizens. In addition to conducting over a hundred interviews, for example, I surveyed 5,076 Arab youth across the UAE across treatment and control groups. I also gathered data through rare in-palace ethnography, interviews with ruling elites (including a monarch), interpretive analysis of state-sponsored spectacles, focus groups, and four experiments with random assignment. In my contribution to this symposium, I discuss this mix of methods, focusing on their integration in the service of an overarching research design, the substantive payoffs I see arising from that design, and the surprises and lessons learned along the way. A key goal is to illustrate the added value of multi-method research designs for addressing classic questions in new and multi-faceted ways. I suggest that such approaches, like lens focal lengths, can help scholars to see phenomena from different angles, promoting creativity and innovation within broader research streams as well as methodological thoroughness.

Multi-Method Research Design

For the state, the making of citizens is a classic challenge. In the modern era, state leaders have placed particular emphasis on molding the citizen, defined at a minimum by a sense of national identity, recognition of the state as a legitimate political authority, and willingness to obey its rules (Merriam 1931; Bendix 1964; Gellner 1983; Hobsbawm 1992). Beyond these fundamentals, state leaders have also sought to mold the citizen in rather more specific and sophisticated ways, aiming to enhance state power, speed up economic growth, or conform to certain ideas of progress.[2] One of the key comparative questions in these areas concerns the extent to which state leaders succeed or fail in their efforts to shape citizens, engaging in what Rogers Smith has called the politics of "people-building" (Smith 2001). Any attempt to answer this question, however, raises several key challenges.

Figure 1: Integrating Methods

Research Questions	Examples of Data Collected
1. What are state leaders' **goals** for social engineering?	Palace ethnography Interviews with ruling elites, including a monarch Intepretive analysis of political speeches
2. What **outcomes** are they achieving?	Survey of Arab youth (n = 5,076) across control and treatment groups
3. Why? What are the **causal mechanisms** leading to these outcomes?	Focus groups with youth Interviews with middle-men Four experiments with random assignment

One challenge involves how to conceptualize and measure success and failure. Although the success or failure of social engineering is often presented in historical hindsight as self-evident, the question is a deceptively simple one, and the answer is rarely obvious. For example, Soviet Russia is frequently given as an example of the failures of social engineering (Alexander and Schmidt 1996; Scott 1998). Yet scholars also argue that many of the attitudes, values, and behaviors forged by the Soviet system have persisted, suggesting Soviet social engineering to create the "new Soviet man" may have succeeded in important ways (Kenez 1985), even if the Soviet regime did not. Another challenge is identifying the effects of social engineering. Makiya (1998), for instance, suggests that Baathist social engineering played an important role in stamping out the ability of Iraqi citizens to think independently, but he also highlights the "culture of fear," driven not by social engineering but by coercion and violence by the state, which contributed to that same outcome. Eugen Weber's famous account of social engineering and cultural change in nineteenth century France (1976) provides another excellent example. In what ways did "peasants" really become "Frenchmen"? What role did the deliberate activities of elites involved in the Paris-based social engineering campaign play in this transformation, as opposed to other causal forces, such as modernization?

In my dissertation, I use a multi-method research design to respond to these challenges, aiming to investigate classic questions of citizen-building in novel ways. A central goal of the study is to help build a more up-to-date and nuanced theory of social engineering for the purposes of citizen-building. To do this, I argue that social engineering goals, outcomes, and causal mechanisms connecting goals to outcomes must be identified with greater conceptual and methodological precision than is typical. A more systematic investigation is needed, in other words, of what state leaders in different contexts *intend* as well as what they *achieve* in their efforts to mold the citizen. For this purpose, I use the United Arab Emirates as a multi-method empirical laboratory (Figure 1). The UAE is a valuable, data-rich context for the study of top-down social engineering, allowing opportunities for ethnography, experiments, process-tracing, congruence-testing, and other modes of analysis. Not only are state leaders unusually open to policy experimentation in the making of citizens, but they also have the political and budgetary flexibility to move beyond rhetoric about change, providing a rich testing ground for theory development. I investigate the following three research questions:

> What are state elites' goals for social engineering, and how are those goals being pursued?
>
> What outcomes have been achieved, thus far?
>
> What causal mechanisms best explain why and how these outcomes have occurred?

First, to identify goals, I use a conceptual framework adapted from the literature on nationalism and citizenship (Bendix 1964; Marshall 1964; Kymlicka and Norman 1994). As Figure 2 illustrates, the framework "unbundles" the concept of the citizen and his or her relationship to the state into four component parts: economic, national, political, and civil. Disaggregation allows more precise goals in citizen-building to be identified, taking into account the possibility of multiple, complex, and conflicting goals at the macro-level.

What kind of a citizen do state leaders want to create? To answer this question, I combined evidence from rare in-palace ethnography, interviews with ruling elites, and content analysis of government strategy documents. Official documents can be an excellent source of information about the type of citizen that state leaders want to produce, especially in times of reform when the "citizen of the future" is often described in rich detail. Yet official documents rarely tell the whole story, especially within secretive authoritarian regimes. Thus, I gathered additional ethnographic evidence as a frequent guest at palace dinners, meetings, and other events attended by ruling

Figure 2: "Unbundling" The Citizen

Dimension	Empirical Referents
1. Economic	Demands for state-provided welfare
2. National	Attitudes toward national identity/allegiance
3. Political	Demands for political participation
4. Civil	Demands for civil liberties, individual freedom

elites, where the problems of the youth and strategies for developing them in new ways were a common topic of conversation. I also conducted semi-structured interviews with ruling elites, including one of the country's seven ruling monarchs. In addition, I used an interpretive approach to examine the regime's use of symbolism and spectacle to motivate and model the new entrepreneurial citizen. For example, I visited new schools built as part of the student-centered education reform movement to interview reformers on the ground and attend classes. Following in the footsteps of George Mosse (1975) and Lisa Wedeen (1999), I analyzed major events targeted at youth as political spectacles, including the Young Entrepreneurs Competition, Festival of Thinkers, the Abu Dhabi Science Festival, the Summer of Semiconductors, and the Celebration of Entrepreneurship. At such events, concepts such as "work" are being invested with new meanings that serve a political purpose. For instance, the idea of private sector work is being tied to personal fulfillment and self-discovery in an effort to create a more market-friendly culture with hints of Weber's spirit of capitalism. I also collected over a hundred photographs of installations, artistic exhibits, posters, slogans, and other forms of visual propaganda that reflect the goals of the social engineering campaign.

To answer my second research question, I surveyed over 5,000 Emirati Arab youth across the country, and used a quasi-experimental design to build knowledge about the micro-level outcomes of the campaign. I selected as the "treatment" a new and celebrated public high school, which has served as an important policy experiment in the fostering of the new citizen. My purpose in designing the survey was to uncover the effects of social engineering in a more nuanced and precise way than is typical of the existing literature on the making of citizens. To estimate these effects, I surveyed Arab youth in control and treatment school types as well as earlier ("pre-treatment") and older ("post-treatment") grade cohorts. Using a difference-in-differences (DD) causal identification strategy, I compared differences in students' attitudes across younger and older cohorts within control schools against the differences in those same attitudes across the same cohorts in treatment schools.[3] By identifying micro-level outcomes in the four areas of citizenship reflected in my framework, I was able to test hypotheses about the intended and unintended effects of social engineering.

Finally, I used several methods to explore causal mechanisms in response to my third research question, aiming to unearth the reasons for success and failure in top-down social engineering. I used congruence-testing and process-tracing to examine the "fit" of existing theory in these areas, exploring the extent to which prominent explanations in the literature actually explain the outcomes that I identified. I also triangulated evidence from focus groups with students and interviews with the Western-educated middle-men actors who are carrying out the campaign on behalf of rulers. To supplement this qualitative evidence, I conducted four experiments with random assignment, which I designed to help disentangle the role played by potential causal mechanisms.

Substantive Payoffs

Several substantive payoffs emerge from this approach. First, combining methods allowed me to paint a richer and more empirically accurate portrait of leaders' goals than would have been possible, I think, through the use of one method. When I present my work, an excellent question that often arises is, "Do autocrats really want to change citizens like this? Our models don't predict that, since we assume autocrats want to maintain social stability." Combining palace ethnography, interviews with ruling elites, and document analysis allowed me to "see" that, in fact, UAE autocrats *do* desire changes in citizens that may be surprising from the perspective of existing theory. For example, it was only through ethnography at the palace and other venues in which ruling elites could speak their minds less publicly that I realized the role played by factors like embarrassment. Poor performance by young citizens in school and on international tests—and the idea that the world may not sufficiently respect the UAE as a result—have caused embarrassment and concern about status. Along with other factors, such as over-reliance on hydrocarbons, the questions of how to obtain respect in the world and project modernity, despite authoritarianism, have helped motivate social engineering. I suspect that, on their own, document analysis and semi-structured interviews may not have revealed such rich information about macro-level goals and motivations.

A second substantive payoff has been a set of insights about what leads to success and failure in the use of social engineering to mold the citizen in these ways. In pursuit of authoritarian neoliberal enlightenment, UAE ruling elites have sought to create a citizen who is economically, socially, and culturally more conscious, self-reliant, and hard-working, but nevertheless remains politically passive. Having hired armies of Western-educated professionals to design and implement "soft" social engineering, they have undertaken a variety of ambitious reforms, ranging from major investments in research

Figure 3: Actual Outcomes of Social Engineering—The *Super*-Entitled Citizen

Dimension	Intended Change	Theorized Change	Actual Change
1a. Economic: *Entrepreneurialism*	↑		
1b. Economic: *Rentierism*	↓		↑
2. National *Nationalism & cultural pride*		↑	↑
3. Political *Demand for political participation (self)* *Demand for political participation (others)*		↑ ↑	↑ ↓
4. Civil *Demands for civil liberties*		↑	

and science-based innovation to the building of a new cultural and educational district featuring branches of the Louvre and the Guggenheim, to a radical overhaul of public schools. These types of changes are, in many ways, exactly what critics both inside and outside of the Middle East have long said are necessary for the region's renewal and revitalization.[4]

Yet I find that social engineering is failing and even backfiring in unexpected ways. Instead of building entrepreneurial citizens as desired, the data suggest it may be giving rise to "*super*-entitled citizens," thus reinforcing the very rentier citizenship mentality and sense of entitlement that rulers wish to change. Rather than displaying greater entrepreneurialism and self-reliance, youth subject to school-based social engineering reported *stronger* economic claims on the state, especially through the perceived right to a government job. In the sample, such "treated" youth were also less willing to pay an income tax to support the country's development than were their same-age counterparts across the same grade levels in regular government schools that are not part of the reform movement. In addition, the "treated" youth reported heightened levels of national and cultural pride, and higher levels of interest in political participation for themselves. At the same time, however, they reported lower levels of support for the right of *all* UAE citizens to have a say in government policymaking. I summarize these changes in Figure 3, illustrating the growth of the *super*-entitled citizen with expectations of elite status and little interest in private sector work.

The use of a multi-method research design has allowed not only a more precise and nuanced investigation of social engineering goals and outcomes, but also a means of identifying causal mechanisms. Why has social engineering failed in these ways, and why might it, unintentionally, be producing *super*-entitled citizens? To explain this outcome, I offer an alternative to the conventional wisdom that top-down planners of all stripes fail because they lack local knowledge (Scott 1998; Mitchell 2002). Rather, I argue, the reasons for failure and perverse outcomes in citizen-formation can be found in the political context of implementation by a professional class of middle-men, who are carrying out the social engineering campaign on behalf of rulers. Largely Western-educated experts specializing in areas such as education, business training, and youth development, these middle-men are key players, both whispering into the ear of monarchs and operating at the local level as teachers, principals, and trainers. They operate in a political context producing behaviors on their part that undermine macro-level strategy. At the local level in schools, for instance, they deliver undue praise, inflate grades, and flatter the culture, both pursuing self-interest and job security and conforming to cosmopolitan norms of political correctness. Rarely, moreover, do they "speak truth to power" by telling the autocrats who hired them what they really think. As a result, although wealth, political will, and expertise are not lacking, macro-level strategy is failing.

Evidence from different sources has been crucial in building and substantiating these arguments about causal mechanisms. First, focus groups and interviews with youth helped to illustrate the substantive significance of survey findings. When treated students report stronger support for the right of citizens to receive a government job, why do they feel this way? When they report heightened interest in political participation for themselves, but *less* support for the right of all citizens to participate in politics, what do they mean, exactly, and why? In response to such questions, subjects gave very revealing answers: "Because we are leaders," "I want to have a good position [in government], a high one, and to have a good salary that fits me," and "We work hard, we put extra effort in studying. So why would we be equal to them?" Interviews with the Western-educated middle-men also served to clarify the behaviors on their part, influenced by the political context, which are leading to such micro-level outcomes. For instance, these interviews made it clear that the risks of being downgraded, transferred, or even deported were leading new teachers, trainers, and principals to offer excessive praise and flattery, without requiring students to earn it. In a tribal authoritarian context such as this, one text message from an angry high school student to a relative in government can get a teacher fired.

Finally, four experiments with random assignment also helped to elucidate causal mechanisms. The results suggest that even small doses of praise and the related increase in one's perceived status can affect attitudes in these areas, in this case helping to foster what I call *super*-entitled citizens. Since the large-N survey did not differentiate between students' attitudes toward their *own* right to a government job and the right of *all* citizens to a government job, these experiments and qualitative data enriched my overall findings about the unintended effects of social engineering, reasons for failure, and the causal mechanisms leading to perverse and unexpected outcomes at the micro-level.

Surprises and Lessons Learned Along the Way

Using multiple methods like this can be a challenge. Yet, while gathering data of different types can increase complexity and attract charges of dilettantism, it can also add value in very significant ways. As I suggest above, a multi-method research design can help to develop, qualify, and enrich findings, keeping the researcher anchored to the overarching question of interest rather than any particular methodology. Beyond this, however, I suggest that multi-method research designs can offer more than, so to speak, the sum of their parts. Such an approach, like experimenting with lens focal lengths, can help scholars to "see" phenomena in new ways, promoting creativity and innovation within broader research streams.

For example, once I had committed myself to a multi-method research design for my dissertation, I was surprised by the relative ease with which I was able to formulate and test interpretive hypotheses about symbolism and spectacle, both important tools within the UAE social engineering campaign. As many a graduate student knows, despite talk of a new era in which all methodologies are treated as equal, significant divides remain. Arguments about potentially differing logics of causal inference are very much alive. Before I embarked on my fieldwork, I recall experimentalists arguing that "You can't test interpretive claims. They're not falsifiable." I also recall consulting interpretive social scientists, who frowned at the use of a large-N survey and experiments to investigate symbolism and spectacle.

As a result, when I began my fieldwork, one surprise was that linking these methodologies was nowhere near as challenging as I had been led to imagine. For me, interpretive methods were essential for understanding what state leaders wanted and how they defined the ideal citizen. They were also critical for understanding how evolving symbolism, regime rhetoric, and spectacles aimed at the rising generation were being put to use for the purposes of social engineering. To employ the typical three-part model of communication, these approaches were most valuable in uncovering (1) the message leaders wish to communicate and (2) the ways in which that message has been communicated to the Arab youth who are the "objects" of the social engineering campaign. Yet, how has this message been (3) received and interpreted? For interpretive political scientists, I believe that a key challenge is how to determine, with rigor and precision, how the symbols, spectacles, and other ideational phenomena involved in the manipulation of meaning actually affect targeted audiences. In this area of inquiry, I found quasi-experimental and experimental approaches especially helpful, supplemented by focus groups and interviews.

Another surprise was how a multi-method approach helped me to "see," not just the payoffs of combining methods, but the challenge of citizen-building itself in new ways. Before I embarked on my fieldwork, a common question was, "But is this really citizen-building? It's not a democracy." That query made me realize how very theory-laden and potentially narrow our conceptualization of this phenomenon may be. For instance, we understandably link the term "citizen" with the Western historical experience of liberal democracy. However, as Aristotle pointed out, the relationship between the individual and the state can be conceived in far more multi-dimensional ways. Second, we associate the challenge of citizen-building with the effort to foster a common national identity, among the first and most fundamental tasks in state-building. While this task remains important, circumstances today are different from earlier eras, since billions of people already have a national identity. A key question for citizen-builders, then, is what comes next? The challenge of citizen-building today may thus be a "horse of a different color," similar in fundamentals to what it was before but imbued with new meanings. Multi-method research can help us explore such classic questions in new ways, facilitating not just methodological rigor but conceptual creativity and theoretical risk-taking.

Notes

[1] I define "social engineering" as activities consciously undertaken by state leaders to shape the hearts and minds, and ultimately the culture, of their own citizens. My definition restricts these activities to socialization and the management of meaning through educational initiatives, state-sponsored spectacles, media campaigns, and other non-violent methods. (Of course, some regimes have also used force to shore up their efforts at "making" citizens, but I treat coercive power as a separate causal factor that can be distinguished conceptually from social engineering.)

[2] See, for example, Kenez (1985) on Soviet Russia, Hanioglu (2011) on Turkey under Ataturk, Garon (1997) and Dower (1999) on post-1945 Japan, and Alston (1969) on Tsarist Russia.

[3] This approach has several advantages for causal inference. First,

it controls for selection bias in treatment assignment. Selection bias is a well-known challenge to causal inference in these areas; students, of course, are not randomly assigned to schools. The DD approach removes this type of selection bias by subtracting out initial differences in outcomes between control and treatment populations, preventing any unobserved factors that remain constant over time, which correlate with treatment assignment and affect the outcome variables, from biasing treatment effect estimates. Such factors may include income levels, levels of parental education, and other demographic differences. Another advantage is the removal of bias stemming from aggregate factors that would cause change in the outcome variables over time or across grade cohorts even in the absence of the treatment. Such factors include age or maturation, broad socio-economic changes, and national or regional political context.

[4] See, for example, the UN's Arab Human Development Report, "Building a Knowledge Society" (2003) and Nasr (2009).

References

Alexander, Jon, and Joachim K.H.W. Schmidt. 1996. "Social Engineering: Genealogy of a Concept." In *Social Engineering*. Adam Podgórecki, Jon Alexander, and Rob Shields, eds. (Ottawa: Carleton University Press), 2–19.

Alston, Patrick L. 1969. *Education and the State in Tsarist Russia*. Stanford: Stanford University Press.

Bendix, Reinhard. 1964. *Nation-Building and Citizenship*. New York: Wiley.

Dower, John. 1999. *Embracing Defeat: Japan in the Aftermath of World War II*. London: Penguin Books.

Faour, Muhammad and Marwan Muasher. 2011. "Education for Citizenship in the Arab World: Key to the Future." In *The Carnegie Papers*. Carnegie Endowment for International Peace.

Garon, Sheldon M. 1997. *Molding Japanese Minds: The State in Everyday Life*. Princeton: Princeton University Press.

Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca: Cornell University Press.

Hanioglu, M. Sükrü. 2011. *Atatürk: An Intellectual Bibliography*. Princeton: Princeton University Press.

Hobsbawm, Eric J. 1992. *Nations and Nationalism since 1780: Programme, Myth, Reality*, 2nd ed. Cambridge: Cambridge University Press.

Kenez, Peter. 1985. *The Birth of the Propaganda State*. Cambridge: Cambridge University Press.

Kymlicka, Will and Wayne Norman. 1994. "Return of the Citizen: A Survey of Recent Work on Citizenship Theory." *Ethics* 104: 352–381.

Makiya, Kanan. 1998. *Republic of Fear: The Politics of Modern Iraq*. Berkeley: University of California Press.

Marshall, Trevor H. 1964. *Class, Citizenship, and Social Development*. New York: Doubleday.

Merriam, Charles Edward. 1931. *The Making of Citizens: A Comparative Study of Methods of Civic Training*. Chicago: The University of Chicago Press.

Mitchell, Timothy. 2002. *Rule of Experts: Egypt, Techno-Politics, and Modernity*. Berkeley: University of California Press.

Mosse, George L. 1975. *The Nationalization of the Masses: Political Symbolism and Mass Movements in Germany from the NapoleonicWars through the Third Reich*. New York: H. Fertig.

Nasr, Vali. 2009. *Forces of Fortune: The Rise of the New Muslim Middle Class and What It Will Mean For Our World*. New York: Free Press.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. New Haven: Yale University Press.

Smith, Rogers M. 2001. "Citizenship and the Politics of People-Building." *Citizenship Studies* 5:1, 73–96.

UNDP. 2003. *Building a Knowledge Society: The Arab Human Development Report*. New York: UNDP.

Weber, Eugen Joseph. 1976. *Peasants into Frenchmen: The Modernization of Rural France, 1870–1914*. Palo Alto: Stanford University Press.

Wedeen, Lisa. 1999. *Ambiguities of Domination: Politics, Rhetoric, and Symbols in Contemporary Syria*. Chicago: University of Chicago Press.

World Bank. 2007. "The Road Not Traveled: Education Reform in the Middle East and North Africa." In *MENA Development Report.*

Using Formal Theory in Multi-Method Research: National Democratic Support for Subnational Authoritarianism

Juan Rebolledo
Yale University
juan.rebolledo@yale.edu

Beginning my third year of graduate school, I faced the daunting task of going from a scrutinizer of academic work to a producer of such work. Coming from Mexico shaped my general areas of interest, but narrowing my project down to a manageable yet still relevant question was a tough task. I was concerned with the political development of young democracies and intrigued by the existence, in Mexico, of sub-national governments with authoritarian characteristics that had successfully maintained power despite national democratization. After reading V.O. Key's canonical text *Southern Politics in State and Nation* (1949), I realized this was not a Mexico-exclusive phenomenon. Upon further research, I recognized that similar dynamics were occurring in countries as diverse as India, Brazil, the Philippines, and Argentina. The very existence of these authoritarian pockets in democracies was puzzling, yet it is even more puzzling that a national democratic government seems to be unable or unwilling to challenge subnational authoritarian practices, even when those regions are controlled by the opposition. I successfully narrowed the topic to an intriguing question: In young democracies, why do democratic national governments, which recently fought authoritarian abuses, seem to be unwilling or unable to act against regional autocrats?

Given the relatively little work (Gibson 2005, Gervasoni 2010, Giraudy 2010) done in this area of research, and the complications that empirical scholars have faced for years when studying democracy, it became clear to me that demonstrating an unequivocal causal relationship would be nearly impossible. There would be no perfect instrument for level of democracy. Nor would I find a natural experiment in which some states were "as if" randomly assigned to retain authoritarian characteristics while others became more democratic. In addition, the complexity and many moving parts of the question made it likely I would miss interesting and important implications of the argument, or even worse run the risk of falling into logical inconsistency, if I forced a single answer onto the question. At times the issue felt intractable, but it became clear that I could only make the project both theoretically manageable and causally explanatory by bringing multiple methods to the task. I thus combined a formal model, quantitative analysis, and qualitative case studies in the project.

The Formal Model

Timing presented the biggest challenge of using a multi-method approach. On one hand, I wanted every aspect of the research project to be informed by the different methods; on the other hand, I wanted the project to be deductive. My knowledge of the cases had to illuminate the assumptions I was prepared to make, but not the hypothesis itself. I had spent time doing fieldwork in Latin America on other projects, which helped me get a sense of the relevant political actors and the strategies available to them. Extensive background reading into other country cases suggested these phenomena were generalizable. However my theory at this point had an excessive number of moving parts. I began working on the formal model, which brought clarity to my theory. I knew who the key actors were: the central government, the governors, and the electorate. The model helped me make my assumptions explicit, and forced me to sort those assumptions I was willing to make from those I was not. It allowed me to incorporate aspects I felt had to be part of the story (governmental resources, resource exchange among different levels of government, expenditure on electoral cycles, the level of democracy and policy making), while forcing me to leave aside inessential aspects of the theory. While many of the inessential aspects are still substantively important, they had to be removed to make manageable an otherwise unmanageable and irreducibly complex reality. In the end, the model generated interesting, counterintuitive, and testable propositions that I would not have uncovered without it.

Let me briefly describe the model and the main propositions derived from it. To understand the endurance of regions with high levels of authoritarian practices, we must pay close attention to the strategic interaction between the central government and the different regions. In particular, in the dissertation I propose that because of support the central government needs from the opposition to get its preferred policy approved, it tolerated states controlled by the opposition with high degrees of authoritarian practice persistence. The formal model consists of subnational regions belonging to different parties, the central government from one of these parties, and the citizens. In my model, regions can vary in their levels of authoritarian practices.[1]

In this probabilistic voting model, citizens can vote for either party. Because a citizen can cast only one vote in the model, citizens from the region in which the central government and the regional government are controlled by different parties must support one at the expense of the other. The central government has two distinct concerns: its party's electoral fortunes (obtaining a large vote share in both regions) and its party's policy agenda. The central government has resources it can spend in any region to sway the electorate to support its party and must choose how to divide these resources among the regions. When the central government spends resources in a region controlled by an opposing party, the governor of that region and his party will have to spend more resources trying to stay in power.

A central government wishing to advance its policy agenda must go through the political process (e.g., a legislative bill in the national congress) and will, therefore, need the support of some members of the opposing party. A central government that does not need support from an opposing faction to implement its desired policy is beyond the scope of this theory. However, I contend that, in most cases of national democracies, no single faction can determine policy unilaterally.

Regional governors are able to offer such support. Governors often have undue influence over legislators for a variety of reasons: because legislators owe their nomination to the governor, as is the case in modern Argentina (Jones and Hwang 2005, Behrend 2011), or because support for systematic disenfranchisement in a region assisted the power holders in being elected, and unites them across an array of policy issues, as was the case in the U.S. South during the Jim Crow era.[2] Crucially, the central government can achieve its preferred policy by negotiating with regional opposition leaders, rather than opposition leaders at the national level who have already expressed opposition to the policy in question. If the government cares about policy, it may offer to reduce electoral intervention in regions controlled by an opposing party in exchange for the political support of that region in advancing its policy agenda.[3]

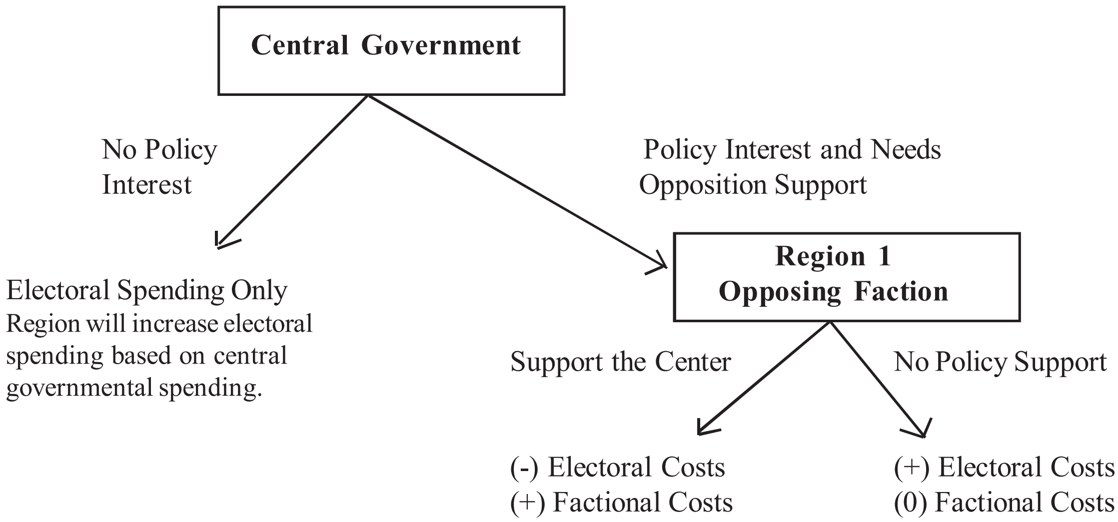
Regional governors face a lose/lose situation. On the one hand, they face an electoral threat from the central government, which, unless they offer their support to the opposing party's agenda, will spend high levels of resources in their region attempting to oust them. Even if the central government does not succeed, its spending resources in this manner entails higher electoral costs for the regional governors, who must react to increased resources spent against them by increasing the resources used to defend themselves. On the other hand, regional governors also face a cost imposed by their party if they decide to support a central government led by an opposing party. The cost imposed by their own party for defecting can manifest itself in a variety of ways, including resource flow, public perception, and electoral support.[4]

A region controlled by an opposition party will make a cost-benefit comparison between the party-imposed costs of supporting the central government and the electoral costs of party loyalty.

Figure 1 presents a stylized version of the formal model based on the current presentation of the argument.

The model shows that an equilibrium exists (out of two possible equilibriums) in which the central government offers to reduce electoral spending in a region controlled by an opposing party in exchange for policy support at the national level. For this equilibrium to exist, two conditions must hold: First, as the chart shows, the central government must have policy interests; second, trade-offs will occur only if the re-

Figure 1



gion can overcome the cost imposed by its national party. What makes these agreements more likely, especially if the electoral and party costs are held constant? Leaders of regions with low levels of democracy have a higher probability of electoral victory, inducing them to spend less on elections. As a result, they can compensate for costs imposed by their own national party more easily than their counterparts in high-democracy regions. In addition, low-democracy regions are willing to accept a lower concession from the central government precisely because their entrenched position and control allows them to more easily overcome party costs. Ceteris paribus, regions with lower levels of democracy have a higher probability of absorbing their national party's costs, are less affected by central government intervention, and are thus more likely to meet the second condition for defection in a much broader set of cases.

The security and control of the governing group in a low-democracy region allows it to be more independent of the national party, and therefore more likely to reach an agreement with the central government. Nevertheless, party defection by a regional governor to support the central government is still a rare event. If party costs are sufficiently high, no degree of authoritarian continuity will be enough for an agreement to be reached.

The theory indicates not only that low-democracy regions have a higher probability of supporting the central government, but also, precisely because they supported the central government, electoral efforts by the central government's party against the regional governor from an opposing party will decrease. Regions governed by an opposing party that supported the central government parties' agenda will observe lower levels of expenditures against them by the central government in their region. We are left with two very unlikely allies, with the national democratic party allying itself with the regional low-democracy governors from an opposing party.

Quantitative Analysis

Having used the formal model, I developed clear propositions which were not obvious at first. Testing the theory clearly required evidence not only of the causal claim I was making, but also of the assumptions that led me to the prediction in the first place. This meant generating a new dataset. The reinforcing nature of the multi-method project became very handy. Just as my knowledge of the key strategic actors and relevant variables, based on previous fieldwork, were fundamental to designing the model, now I began with a concrete list of variables, with specific meanings in the context of the model. I began searching for proxies of these variables. Using multiple methods helped set a clear direction to the data collection process.

For the empirical analysis, I constructed a panel dataset of all 32 states in Mexico for the period 1997–2008. To the best of my knowledge this is the only study that includes all 32 states and covers not just one presidential administration but three distinct administrations, including both PRI and PAN national governance. Mexico lends itself to the study of cross-regional variation in the persistence of subnational authoritarian institutions. Mexico is a federation composed of 32 subunits, each with its own executive, legislative, and judiciary branch. Nationally the country had an autocratic political system, characterized by a single-party hegemonic regime, which began crumbling in the late 1980s. There is much literature on how the Institutional Revolutionary Party (PRI) stayed in power for so long. However, a rough consensus holds that a combination of good economic performance in the 1960s and 1970s, in conjunction with a rupture-preventing strategy, helped the PRI maintain hold (Magaloni 2006, Ames 1970). After the 1993 and 1996 round of electoral reforms that "leveled the playing field" for the opposition, there was finally a guarantee of free and fair elections. The advent of national democracy in Mexico is usually attributed to the growing support for the opposition after recurrent economic crises during the 1980s and 1990s.

These crises, in conjunction with economic development that had occurred in the previous two decades, freed citizens from reliance on the state and allowed them to make ideological investments. Regardless of the theory of the national democratization of Mexico, the important point is that it was an exogenous shock to subnational units' political power.[5]

The democratizing reforms crystallized in 1997 when the PRI lost majority control of the congress and needed the support of the opposition to pass legislative initiatives. In 1997, under President Ernesto Zedillo of the PRI, Mexico was living in a new democratic reality, in which the President needed opposition support to pass his legislative agenda. However subnational alternation in office had begun some years before. In 1989, an opposition governor first took office in Baja California after a series of post-electoral conflicts. Research has already been conducted on subnational alternation in an authoritarian setting, and it is generally understood that alternations in gubernatorial offices before 1994 were not the product of elections (e.g., Eisenstadt 2004). Since national democratization, some states followed the national trend and became democratic. In other states, governors successfully concentrated their power and were able to sustain authoritarian practices and institutions in the newly democratic setting. Thus, as is widely acknowledged, Mexican states vary significantly in their level of democracy (Fox 1994, Giraudy 2010).

The 31 Mexican states, and the federal district, enjoy a great deal of autonomy. They all drafted their own constitutions. Governors are all elected for six-year terms and local legislators for three-year terms, with no reelection allowed in either branch. Since 1997, all gubernatorial candidates have been products of some sort of real electoral process taking place within the state. Fiscally, states also enjoy a fair degree of autonomy. They receive both an automatic transfer from the central government and a potential discretionary transfer; since the 1990s, the *automatic* transfers have been large enough that states are not financially dependent on the central government. This arrangement means that local taxation is quite low and deficient (around 80% of a state resources come from federal transfers).

Taking advantage of state fixed-effects, I find evidence that, as the theory suggests, lower levels of democracy in an opposition state mean a higher likelihood that state will defect from party lines and support the central government. The evidence is not only statistically significant, but substantively significant as well: States that retain the most authoritarianism are 37.9% more likely to support the central government than states with a mean level of democracy.

I also find evidence that opposition governors who defected from party lines to support the central government experience lower levels of hostile spending from the central government in their regions (measured as media expenditures by the federal government in the region). The measure is also both statistically and substantively significant. The effect is large: having supported the central government in the year prior to an election decreases the media expenditures by the central government in the state by about 64%. Both of these results are robust in terms of distinct specifications and conceptualizations of the variables, adding certainty to the results.

Though unequivocal causation cannot be established by quantitative evidence alone, these findings invite further investigation into both the conditions under which these unlikely alliances are made, and the benefits to the parties involved. This study also makes explicit that researchers should pay attention to the legislative incentives behind such agreements.

Qualitative Case Studies

The quantitative evidence I present in the dissertation clearly shows a link between opposition regions with low levels of democracy and the likelihood they would defect from party lines to support the democratic national government. Nonetheless, more evidence needed to be presented to show that the mechanism driving this association was the one I proposed through the aforementioned formal model. Just as I attempted to bring the advantages of qualitative work to my quantitative analysis—by emphasizing concept development and choosing proxies based on a deep understanding of recent political history—I tried to bring the rigor of quantitative analysis to my qualitative research. To that end, I pursued both within-case analysis and across-case variation.

To better understand the mechanism and the different equilibriums, it was important to look at cases that present variations in the key independent and dependent variables, in other words variations in level of democracy, support for the central government, and level of central intervention in the region. Additionally, because elections in Mexico are staggered by state, I wanted to examine states facing the same relevant actors and actions during their electoral years. For this reason I chose four states that had elections in 1998, 2004, and 2010. Puebla, Veracruz, and Oaxaca have low levels of democracy, whereas Chihuahua is widely held to have a high level of democracy.

First, I undertook within-case analysis of each state, presenting evidence both that the assumptions of the model hold, and that the suggested causal mechanism is at work. For this I conducted over 50 interviews with elites and non-elites in the states of Puebla, Veracruz, Chihuahua, Queretaro, Michoacan, and Mexico City. Among them were: the heads of all important parties in the region, the campaign managers, heads of newspapers, journalists, members of the electoral commission, presidents of local universities, former local legislators, current local legislators, and former national legislators, among others. Archival research was conducted to analyze newspaper clippings in the different states.

Using a combination of the primary evidence from fieldwork and research on the recent political history of each state, I presented a detailed account of the process that takes temporality, as well as the sequence of events, as a crucial part of the evidence. After presenting within-case studies for all four states, I pursued a case study of a particular bill, the 2003 attempted fiscal reform, to analyze cross-state variation in gubernatorial defection and central government intervention during the next election. Careful case selection of all four states allowed me to isolate the variables of interest by holding other factors constant: In addition to having elections in 2004, all four states were governed by the PRI while Vicente Fox of the PAN was president. This bill case study allowed me to identify cross-state variation, which had so far been absent from the quantitative analysis.

Conclusion

In the end, while perhaps no individual piece of evidence was a silver bullet, the combination of methods presents a strong case for the theory proposed. Different types of evidence all pointing in the same direction reduce skepticism. However, a word of caution is warranted. Issues of timing in the use of the different methods were a constant concern. I was especially concerned about producing deductive work. I did not want to make the process of theory generation and testing identical by generating a theory based on Mexico, and then testing that theory using Mexico as a case.

Researchers using multiple methods that include quantitative and qualitative work confront a tension between needing to understand cases in order to identify the relevant actors and their possible strategies, and generating a theory independent of evidence gleaned from fieldwork—*unless* the researcher plans to test the resulting theory in a different context. The testable predictions I presented were not generated by my original fieldwork, which informed the model but could not have anticipated the comparative statics that it generated. In addition, I incorporated Argentina as a shadow case in the concluding chapter of the dissertation. This both tests the generalizability of the theory and serves as an out-of-sample test.

Moving from the dissertation to a book project, I intend to continue using a combination of quantitative and qualitative methods to fully analyze the theory in three different countries: Brazil, Argentina, and the United States. I have begun collecting data for these countries, which I plan on complementing with in-depth case studies in each. In addition, now that I will be making explicit comparisons across countries, some of the variables of the model that were fixed for the case of Mexico will vary, allowing me to test further implications of the theory.

Notes

[1] Examples of regions that maintain high levels of authoritarian practices are, in Argentina, La Rioja, San Luis, Santiago del Estero, Santa Cruz, Rio Negro, and Formosa; in Mexico, Oaxaca, Hidalgo, Veracruz, and Puebla; and in Brazil, Maranhao, Para Piauy, and Bahia until 2006.

[2] Brazil is another example of a country in which the central government goes through the regional bosses to obtain congressional support. Samuels (2003) defends at length this position, on the other hand, Cheibub et al. (2009) find that the regional effect is not as large as expected. The theory in this present work is consistent with Cheibub because we would expect parties to have the greatest influence and that state effect would be rare, as predicted by the model.

[3] Central government intervention in a region can take a variety of forms, including declarations of federal intervention that remove the governor, resources spent garnering votes, deploying the central government's intelligence, and influencing the Supreme Court in sanctioning (or not) governors that continue authoritarian practices.

[4] This was the case with the UCR in Argentina from 2006–2010 in threatening to expel from the party legislators and governors who defected from party lines to support the central government. They also were threatened with intervention in the provincial party committees.

[5] Perhaps the only theory of Mexican democratization that would be problematic is presented by Lujambio and Segl (2000) who claim that national democratization in Mexico was only possible via local democratization as a first step. However, this view is not widely held. Local democratization in a few municipalities is usually seen as temporary exceptions.

References

Ames, Barry. 1970. "Bases of Support for Mexico's Dominant Party." *American Political Science Review* 64:1, 153–167.

Behrend, Jacqueline. 2011. "The Unevenness of Democracy at the Subnational Level: Provincial Closed Games in Argentina." *Latin American Research Review* 46, 150–176.

Cheibub, José Antonio, Argelina Figueiredo and Fernando Limongi. 2009. "Political Parties and Governors as Determinants of Legislative Behavior in Brazil's Chamber of Deputies 1988–2006." *Latin American Politics and Society* 51:1, 1–30.

Eisenstadt, Todd A. 2004. *Courting Democracy in Mexico: Party Strategies and Electoral Institutions*. London: Cambridge Universtiy Press.

Fox, Jonathan. 1994. "Latin Americas's Emerging Local Politics." *Journal of Democracy* 5:2, 105–116.

Gervasoni, Carlos. 2010. "A Rentier Theory of Subnational Regimes: Fiscal Federalism, Democracy and Authoritariansim in the Argentine Provinces." *World Politics* 62:2, 302–340.

Gibson, Edward L. 2005. "Boundary Control: Subnational Authoritarianism in Democratic Countries." *World Politics* 58:1, 101–132.

Giraudy, Agustina. 2010. "The Politics of Subnational Undemocratic Regime Reproduction in Argentina and Mexico." *Journal of Politics in Latin America* 2:2, 53–84.

Jones, Mark P. and Wonjae Hwang. 2005. "Provincial Party Bosses: Keystone of the Argentine Congress." In *Argentine Democracy: The Politics of Insitutional Weakness*. Steven Levitsky and María Victoria Murillo, eds. (University Park, PA: Pennsylvania State University Press), 115–138.

Key, V. O. 1949. *Southern Politics in State and Nation*. Knoxville: University of Tennesee Press

Lujambio, Alonso with Horacio Vives Segl. 2000. *El Poder Compartido: Un Ensayo sobre la Democratización Mexicana*. Mexico: Oceano.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. London: Cambridge University Press.

Samuels, David. 2003. *Ambition, Federalism and Legislative Politics in Brazil*. London: Cambridge University Press.

Multi-Method Fieldwork in Practice: Colonial Legacies and Ethnic Conflict in India

Ajay Verghese
Stanford University
ajayv@stanford.edu

It seems odd these days for a young graduate student to suggest a dissertation project which does not include a multi-method research design. The multi-method movement in political science has become central to the way we structure our research. A carefully executed multi-methods project is the holy grail to which many a scholar aspires. And yet, graduate students have only limited advice in carrying out this kind of research. My goal in this article is to lay out the inner workings of a multi-method project. By presenting the components of my dissertation, which combined a large-N analysis of 589 districts in India with 15 months of archival research and elite interviews carried out in six case studies, I explore: (a) how a multi-method project is created and implemented, and (b) its potential pitfalls and payoffs.

My project (Verghese 2012) examines the puzzle of why ethnic conflicts in multi-ethnic states revolve around one identity rather than another. So, for example, why do conflicts sometimes revolve around religion but at other times around language? What explains *patterns* of ethnic conflict in a multi-ethnic state? This is an important question for plural states around the world struggling to limit ethnic violence. I had chosen this research question in part because of important recent work published on ethnic violence in India (Brass 1997, 2003; Varshney 2002; Wilkinson 2004). While these books had advanced our knowledge of ethnic conflict in the country considerably, I was rather surprised that they all omitted a large potential explanatory variable: the legacy of British colonial rule. After all, India was the crown jewel of the British Empire, and the political science literature linking colonialism and contemporary ethnic violence is immense, much of it focused on African cases (Horowitz 1985; Laitin 1986; Young 1994; Mamdani 1996; Posner 2005).

Furthermore, India presented a unique opportunity to study the impact of colonialism on ethnic conflict because only three-fourths of the population of the country ever came under direct colonial rule. These areas, known as *provinces*, were governed by British administrators. But the rest of the country remained under the control of independent native kings in territories called *princely states*. By comparing conflict outcomes across provinces and princely states, I hoped to isolate the effects of colonialism on ethnic conflict.

But first: why use multi-methods in this project at all? I admit that my initial intentions were hardly honorable; I wanted to utilize multi-methods because that's what you were supposed to do, especially if you wanted to write a well-received dissertation. Luckily, I later came to realize that a project as broad as this would have suffered had it utilized merely one methodology. Imagine only a statistical analysis—you would instantly say, "But you never spent a day in the field or in the archives!" Likewise, imagine fieldwork in two case studies; you might here rightly ask: "But does the argument travel?" Multi-methods have achieved a place of prominence in political science research not because it is simply the newest fad, but because it allows researchers to examine questions in a more complete and exhaustive fashion.

But the larger problem with combining methodologies was that the kind of work I really admired and wanted to do—comparative-historical analysis—was rarely married together with statistics.[1] What interested me was colonialism in India. But how exactly did one go about combining, for instance, archival research on the colonial period with a regression analysis? The answer was not obvious to me. An especially vexing problem was collecting data. I quickly realized that finding reliable figures on ethnic conflict *during* the colonial period in India would be almost impossible, so I would have to restrict my statistical analysis to the contemporary (post-independence) period.

However, there is no inherent contradiction between doing comparative-historical work *and* statistically-oriented research. I planned to run a statistical analysis of the broad pattern of ethnic conflict in contemporary India, but the comparative-historical analysis, on the other hand, would be situated within a number of targeted case studies, aimed at uncovering the mechanisms at work in producing specific violence outcomes. The two methodologies seemed complementary rather than conflicting, and together could help explain not only contemporary outcomes but also their historical causes.

Dataset Construction and Statistical Analysis

Once I had decided that I wanted to pursue a multi-method dissertation, I then set about figuring out how to actually do it. I started with the large-N analysis. I wanted to examine the broad pattern of ethnic conflict throughout modern India and its potential causes, and statistical analysis offered the best opportunity to see the big picture. I began by spending an inordinately long time collecting data on a variety of variables for my study.

I first considered the unit of analysis which I wanted to examine for this research project. Because I was interested in the effects of colonial rule, I decided to pursue a district-level analysis. The entire system of district administration in present-day India was a legacy of the British period; districts, for the most part, were either completely part of a former British province or a former princely state. Looking at states, on the other hand, was much more problematic: a state like Kerala in southern India, for example, was half-British and half-princely, which posed a major coding problem. But looking at the districts *within* Kerala made identifying British and princely areas much easier. I used the list of 2001 districts from the Indian census and ended up with a total of 589 for the analysis.

Then I coded the primary independent variable: the type of colonial rule. This was the most painstaking process of all: I had to determine whether every district in India was either part of a former province or princely state.[2] This necessitated re-searching district websites, reading British colonial reports on individual districts, and comparing geographical coordinates between the two. I coded the type of colonial rule in two ways: a dummy variable (1 if a province), and a variable recording the number of years a district was under British rule (0 for all princely states). I hoped that using two different measures would increase the confidence in my coding.

I then began to compile figures on ethnic conflict using two different existent datasets. The first source of data was the Worldwide Incidents Tracking System (WITS), which used national and international press reports to provide figures on caste and tribal conflict throughout India during the period 2005–2009. The second source was the Varshney-Wilkinson dataset on Hindu-Muslim riots in India covering the period 1950–1995. Together, these two sources of data gave me a broad view of ethnic conflict throughout the post-independence Indian republic. I couldn't document what had happened during the colonial period, but I could at least know about ethnic conflict today, and with smartly selected case studies, I hoped to be able to uncover the deeper causes behind it.

I finally compiled dozens of control variables from various sources: the Indian census, a private statistical firm (IndiaStat), and the Indian Human Development Survey, carried out by researchers at the University of Maryland. Examples of some of these variables were contemporary data on population, geography, the economy, and infrastructure. These variables allowed me to account for a number of alternative arguments about the causes of ethnic violence. Because the conflict data for my dependent variables were count variables (i.e., number of deaths and injuries in ethnic conflicts), I utilized a negative binomial regression model.

I hoped that by constructing this highly detailed database, both the specific independent variable in which I was interested (the type of colonial rule) and the specific dependent variable of interest (ethnic conflict) could then be used by other scholars in future studies. My hope was that a scholar interested in, for example, the effect of colonial legacies on Indian political parties could use my coding scheme of provinces and princely states; or, a scholar interested in the effect of poverty on caste riots might find my compilation of WITS conflict data to be helpful in that regard.

The results of my analysis confirmed that colonial rule had a major impact on ethnic violence outcomes in modern India, but not in the way I had initially expected. I found that in former provinces, *caste and tribal conflict* was the major problem; however, in former princely states, *religious conflict* was endemic. I had simply expected that *all* British provinces would be worse in terms of ethnic violence, but there was an important dichotomy in violence outcomes which I had not anticipated. My statistical analysis therefore confirmed that my chief independent variable of interest—colonial rule—was important, and it had likewise allowed me to rule out a number of potential alternative explanations. It gave me support for my working hypothesis, which I could then further investigate using qualitative fieldwork in India.

Comparative Case Study Fieldwork

The next step in my project was qualitative fieldwork. The statistical analysis alone was not enough. How could I explain the result which I had found? What *specifically* about colonial rule created this apparent dichotomy in ethnic violence outcomes? What were the mechanisms at work? Although you could certainly test mechanisms using certain kinds of advanced quantitative techniques, I felt the need to get my hands dirty and spend some time in the field. I wanted to unpack the logic at work that drove ethnic violence. Furthermore, I believed that qualitative fieldwork would lend a certain credence and believability to the project which a statistical analysis alone could never do.

First, I needed to carefully pick cases to study. This proved a rather daunting task. My working hypothesis, supported by my statistical analysis, was that variations in colonial rule effected contemporary patterns of ethnic conflict. So ideally what I wanted were paired comparisons—that is, two cases which were similar in almost every regard except for variation on a key independent variable of interest: colonial rule. I had the image in my head of exactly what I hoped to find: one princely state situated right next to a British province.

As I had spent some time in a famous princely state of north India while learning Hindi, this naturally became my first case (Jaipur). And lo and behold, right next door was a former British province (Ajmer). I called these two cases a "paired historical comparison," adapting terminology used by George and Bennett (2005: 151). I began my research in this area in the fall of 2010. The first thing I needed to do was figure out whether the pattern from my statistical analysis was also evident at the small-N level of analysis. That is, did Jaipur, as a former princely state, experience more religious conflict than Ajmer? And did Ajmer, as a former British province, experience more caste and tribal conflict than Jaipur? This is what my theory would predict, so I viewed qualitative fieldwork as both an opportunity to investigate mechanisms *and* an opportunity to reconfirm my broader hypothesis.

It is worth pausing for a moment to acknowledge the anxiety you feel in the field when you realize that your hypothesis may not be supported. And this was how I felt while collecting data in Jaipur and Ajmer about contemporary patterns of ethnic conflict. I had a sneaking suspicion every case I studied might turn out to be a deviant case. My qualitative research in the area consisted of two components: elite interviews (to figure out the state of violence in the contemporary period), and archival research (to figure out the underlying cause of this violence).

So I began with interviews, and I had been told that you always begin with journalists. No one knows more about the broad politics of an area than a journalist. Then I expanded my interviews to include police officers, government officials, NGO workers, ethnic group leaders, and a wide variety of other respondents. What I found was that the same pattern of ethnic conflict which I had uncovered at the large-N level of analysis was also evident when looking at Jaipur and Ajmer. Jaipur was indeed a major area of religious conflict, but in Ajmer the viol-

Table 1: Verghese Dissertation Case Studies

Case Study	Colonial History	Selection Criteria	Predominant Violence
Jaipur	Princely	Northern Case	Religious
Ajmer	Province	Northern Case	Caste and Tribal
Malabar	Province	Southern Case	Caste and Tribal
Travacore	Princely	Southern Case	Religious
Bastar	Princely	Deviant Case	Caste and Tribal
Hyderabad	Princely	Deviant Case	Caste and Tribal

ence revolved around caste. A wealth of interviews provided strong evidence that this was the case.

I then shifted to archival research. This really got to the crux of the issue—what made Jaipur experience more religious conflict? After spending weeks at the Jaipur City Palace archives, I uncovered a long colonial history of religious riots, most of them due to the discriminatory policies of the Hindu kings who ruled over the princely state. Muslims had been brutally repressed in the area, leading to long-term antagonisms between the two communities. In Ajmer, however, British administrators enforced discriminatory policies not toward Muslims, but lower castes and tribal groups. New land policies increased rural taxes and strengthened the power of local landlords. Therefore, there was little religious conflict in Ajmer, but a lot of violence which revolved around caste and tribal identities.

So in short, my quantitative analysis had confirmed that colonial legacies did matter, but my qualitative analysis finally gave me a plausible *mechanism*: contemporary patterns of ethnic conflict were caused by legacies of discriminatory policies dating from the colonial period. Combining the two methodologies together, it finally started to make sense.

My next step was to carefully select another paired historical comparison, this time from south India. As I had already worked in the north, I traveled southward to try to account for the enormous regional diversity of India. I found that in the small southern state of Kerala, the entire northern region (Malabar) had been under the control of the British, but the entire southern region (Travancore) had remained under the control of a Hindu dynasty. Better yet, the British themselves had called this political system an "accident" of history. I then embarked on the same fieldwork which I had carried out in Jaipur and Ajmer: a number of interviews to determine the contemporary pattern of conflict, and then extensive archival research to address underlying historical causes. Just as in the north, I found that British Malabar experienced more caste and tribal conflict whereas princely Travancore experienced more religious conflict. And again, archival research revealed long legacies of discrimination which continued to reverberate into the modern period. Because I had found similar patterns in north and south India, I felt reassured that I was onto something.

Finally, I selected two deviant cases: princely states with enormous amounts of caste and tribal conflict. These two cases, Hyderabad and Bastar (both located in eastern India), posed a major problem for the theory underlying my dissertation. And when I set foot in Hyderabad, I realized that I really had no explanation whatsoever to account for the deviant nature of these cases. Why should Hyderabad and Bastar experience such immense caste and tribal bloodshed, especially when no other princely states were similar? Interviews in the region were helpful in explaining contemporary violence, but I still couldn't understand why the two regions were so violent, which is exactly the opposite from the outcome that my theory would predict.

The major breakthrough came after spending a lot of time in regional archives. I discovered that Hyderabad had initiated the same land reforms which had occurred throughout British India, and was one of the few princely states to do so. Similarly, Bastar had come under heavy British intervention during the colonial period, much more than most other princely states of similar size. That's why these cases were *idiosyncratic*. In both cases, it also looked like the British were the culprits behind the scene. Therefore, I at the very least had an explanation for why these were deviant cases for my theory.

By the fall of 2011 I was ready to return home to America. I had spent a year in the field, had visited five archives, and conducted around 75 interviews. Only after a brief period of not thinking about political science at all was I then able to return to my project and begin to unpack what I had discovered, and how it all fit together.

Pitfalls and Payoffs of Multi-Method Research

Critics of multi-method projects often note that using multiple methodologies is quite different from using them *well*. This is a good point. Most people who use multiple methodologies do not become experts on two kinds of methodologies; rather, they learn basic competency in two areas. And it's an open question as to whether or not that is preferable to proficiency in one.

While I felt quite proficient at carrying out interviews in Hindi and poring over centuries-old archival documents, I felt somewhat less confident in my statistical analysis. What if I had omitted a critical variable and re-running the regression with said variable changed everything? What about endogeneity? Or robustness checks? I recall sitting down with a professor of history in India to explain my project, and after I detailed my statistical results she interrupted and said: "In history we don't really use statistics, but I'm guessing you are confident your results are correct?" Not without reservations.

I suppose that every scholar to some extent must grapple with this question. But those undertaking multi-method projects open themselves up to criticism from all fronts—the ethnographically-inclined are not pleased with only six months of fieldwork, the historically-inclined might like further archival work, while the statistically-inclined are similarly unimpressed with your rather basic model. So, what to do?

Considering the diversity of the political science discipline, I hardly think the answer is to completely ignore either qualitative or quantitative work in our projects. That is simply no longer a tenable position. So rather than accept competency in two methodologies as the basic criteria for doing a multi-method project, strive to do two well. Certainly this is easier said than done. I never intended to do any statistical analysis when I got to graduate school, so having a large-N chapter in my dissertation was a challenging but good step in the right direction. It is far from perfect, but getting further quantitative training under my belt is entirely within my control. There's no reason I cannot become as skilled in quantitative methods as I am with qualitative methods, and there's no reason a multi-method project can't make more than one of the methodologically diverse political science audiences (relatively) happy.

Part of the problem may also be that multi-method work and the multi-method movement are relatively new within political science. Therefore, departments still are in the process of adapting and ensuring that graduate students receive adequate training in how to carry out both qualitative and quantitative research. As multi-methods continue to gain popularity, more graduate students will be equipped with the tools to carry out these kinds of projects successfully.

The payoff of a smart multi-method dissertation is obvious: You have a variety of evidence that bolsters the strength of your central argument. My belief that colonialism matters in promoting patterns of ethnic conflict in India is borne out not merely by a large-N statistical analysis, but also by interviews, archival research, and lots of time spent in the field. By triangulating various techniques, I feel more confident in my argument than I would had I used merely one kind of methodology. This is not to disparage the work of those who do—but scholars are always left answering one of the questions which I stated earlier: Why didn't you go into the field or archives? Or, does your argument travel? It seems like the only way to offer sufficient answers to these questions (whether the questioner is a colleague, committee member, or potential employer) is to employ multi-methods.

I have an idea about my next research project, although I have not yet thought in detail about its methodology. But I do know that if I have enough data available, I'll employ a multi-method research design. And I'll continue to work at getting better at any kinds of methods I utilize. This is the best way to persuasively tackle the research problems that face us as political scientists.

Notes

[1] Most of the work in the comparative-historical tradition (see Mahoney and Rueschemeyer 2003) does not use statistical analysis. In fact, scholars like James Mahoney (2004) have argued that statistical analysis is poorly suited to comparative-historical research for a variety of reasons. Recent work by economists has sought to combine econometric analysis with historical research; see, for example, Acemoglu et al. 2001. However, the historical research in question is almost always limited to the brief use of secondary sources, and rarely entails in-depth archival or case study fieldwork. As Marcus Kreuzer notes about this kind of work, "the quality of quantitative research directly depends on the closeness of its *dialogue* with historical knowledge" (2010: 383, emphasis added). Too often, quantitative scholars use history only to grasp for and sketch out plausible causal mechanisms.

[2] Iyer (2010) also constructed a dataset of colonial India. However, she largely compared colonial and post-colonial maps, whereas I used actual geographical data and district reports from both the colonial and contemporary period to match districts.

References

Acemoglu, Daron, Simon Johnson, and James Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation" *American Economic Review* 91:5, 1369–1401.

Brass, Paul R. 1997. *Theft of an Idol: Text and Context in the Representation of Collective Violence*. Princeton: Princeton University Press.

Brass, Paul R. 2003. *The Production of Hindu-Muslim Violence in Contemporary India*. Seattle: Washington University Press.

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

Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Iyer, Lakshmi. 2010. "Direct versus Indirect Colonial Rule in India." *Review of Economics and Statistics* 92:4, 693–713.

Kreuzer, Marcus. 2010. "Historical Knowledge and Quantitative Analysis: The Case of the Origins of Proportional Representation" *American Political Science Review* 104:2, 369–392.

Laitin, David D. 1986. *Hegemony and Culture: Politics and Religious Change Among the Yoruba*. Chicago: University of Chicago Press.

Mahoney, James. 2004. "Comparative-Historical Methodology" *Annual Review of Sociology* 30, 81–101.

Mahoney, James and Dietrich Rueschemeyer. 2003. *Comparative Historical Analysis in the Social Sciences*. Cambridge: Cambridge University Press.

Mamdani, Mahmood. 1996. *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton: Princeton University Press.

Posner, Daniel. 2005. *Institutions and Ethnic Politics in Africa*. Cambridge: Cambridge University Press.

Varshney, Ashutosh. 2002. *Ethnic Conflict and Civic Life: Hindus and Muslims in India*. New Haven: Yale University Press.

Verghese, Ajay. 2012. *Colonialism and Patterns of Ethnic Conflict in Contemporary India*. Ph.D. Dissertation, The George Washington University.

Wilkinson, Steven I. 2004. *Votes and Violence: Electoral Competition and Ethnic Riots in India*. Cambridge: Cambridge University Press.

Young, Crawford. 1994. *The African State in Colonial Perspective*. New Haven: Yale University Press.

Introducing the QCA Package: A Market Analysis and Software Review

Alrik Thiem
ETH Zurich, Switzerland
thiem@sipo.gess.ethz.ch

Adrian Dusa
University of Bucharest, Romania
dusa.adrian@unibuc.ro

The increasing popularity of Qualitative Comparative Analysis (QCA) as a tool for social-scientific inquiry has also led to a proliferation of tailored software. Researchers now have the choice between three graphical interface (GUI) and three command line interface (CLI) solutions. In this article, we first present a brief overview of the QCA software market, following which we introduce the **QCA** package for the **R** environment by drawing operational parallels to **fs/QCA**, the most common GUI software.

The QCA Software Market

In this section, we present a concise analysis of the QCA software market. On the supply side, three GUI and three CLI programs are now at researchers' disposal, none of whose capabilities equals those of another program. The demand side analysis reveals that **fs/QCA** holds a clear monopoly. For a number of reasons, CLIs have failed to win any market shares so far.

The Supply Side

Table 1 provides an overview of the existing QCA programs and their individual functionality with respect to the processing of different QCA variants, the derivation of distinct solution types and a number of instrumental procedures. The left side of the table lists the three GUIs **Tosmana** (Cronqvist 2011), **fs/QCA** (Ragin and Davey 2009), and **KirqST** (Reichert and Rubinson 2012), the right side the three CLIs **fuzzy** (Longest and Vaisey 2008), **QCA3** (Huang 2012), and **QCA** (Dusa and Thiem 2012). Three symbols designate the scope of functionality: A filled dot indicates high functionality, a half-filled dot moderate functionality, and an empty dot almost no or no functionality.

All programs are capable of deriving solutions for the most basic QCA variant—crisp-set QCA (csQCA), but this is not the case for multi-value QCA (mvQCA) and fuzzy-set QCA (fsQCA). Unlike **KirqST** and **fs/QCA**, **Tosmana** cannot process fuzzy sets, but it is the only GUI for mvQCA. This variant can also be handled by the two CLIs **QCA** and **QCA3**, both of which are equally capable of processing csQCA with temporal information, usually referred to as tQCA.[1] Among the three GUIs, this is possible only with **fs/QCA**.

Complex and parsimonious solutions can be derived by all GUIs and CLIs alike, but the automated generation of intermediate solutions for the variants covered by each alternative are only possible with **fs/QCA** and **QCA**. **KirqST** and **QCA3** give users flexibility concerning the choice of those remainders which are made available for minimization, but they provide no functionality for the automatic elimination of difficult counterfactuals. Particularly in research contexts with large numbers of remainders, manual elimination becomes a cumbersome and error-prone task.

With regard to the provision of various QCA-related procedures, the two CLIs **QCA** and **QCA3** cover the broadest spectrum. In particular, **QCA** offers the widest range of calibration tools, while it shares with **KirqST** the possibility to conduct automated analyses of necessity relations. **Tosmana** factorizes solutions by default, whereas **QCA** offers a separate function for this purpose. Simplifying assumptions can be identified by **Tosmana**, **QCA3** and **QCA**. With the exception of **Tosmana**, all programs also compute usual parameters of fit. However, statistical tests are possible only in **fuzzy**, which has been explicitly developed with a probabilistic focus, and **QCA3**.

The Demand Side

Figure 1 charts the trend in the publication number of peer-reviewed journal articles. Between 1993 and 2011, 123 applied articles whose content can be subsumed under the headings of *comparative politics*, *international relations*, and *sociology* have appeared, some of them in top-ranking periodicals such as the *American Journal of Sociology*, the *American Sociological Review*, the *European Journal of Political Research*, *International Organization*, and the *Journal of Conflict Resolution*. Despite a moderate decline following the peak year in 2009 with 19 articles, over half of this total has been published after 2007.

Out of all 123 articles, 65 have referenced the software the authors drew on for their analyses. The **fs/QCA** program has enjoyed the highest popularity with 54 articles (83%) citing its use. Ten articles (15%) employed **Tosmana** and a single article relied on **fuzzy**.[2] In contrast to GUIs, CLI programs have thus won virtually no market shares whatsoever, but this finding should hardly come as a surprise. As the first CLI program, **QCA** still had extremely limited capabilities at the time of first release. It was only in 2008 that **fuzzy** appeared, and not until after 2009 when **QCA3** became available. But time is not the sole explanation. At least as importantly, GUIs possess some major advantages over CLIs from the perspective of many end-users in the social sciences: The program window immediately provides a feature overview; navigation between functions and objects is facilitated by menus and mouse-clicks; and no knowledge of programming syntax is required. CLIs do not offer these amenities. They often look surprisingly unspectacular yet intimidating to unaccustomed users, but exactly this plainness hides their light under the bushel.

R as well as *Mata* (**Stata**'s programming language) offer much more than is possible with menu entries and dialogue boxes. For regular users of QCA in particular, the advantages of CLIs are considerable. First, any repetitive task performed in a GUI is very time-consuming, whereas the same task can be performed almost instantaneously in a CLI. And second, while users see both the command and the result in a CLI, any click-

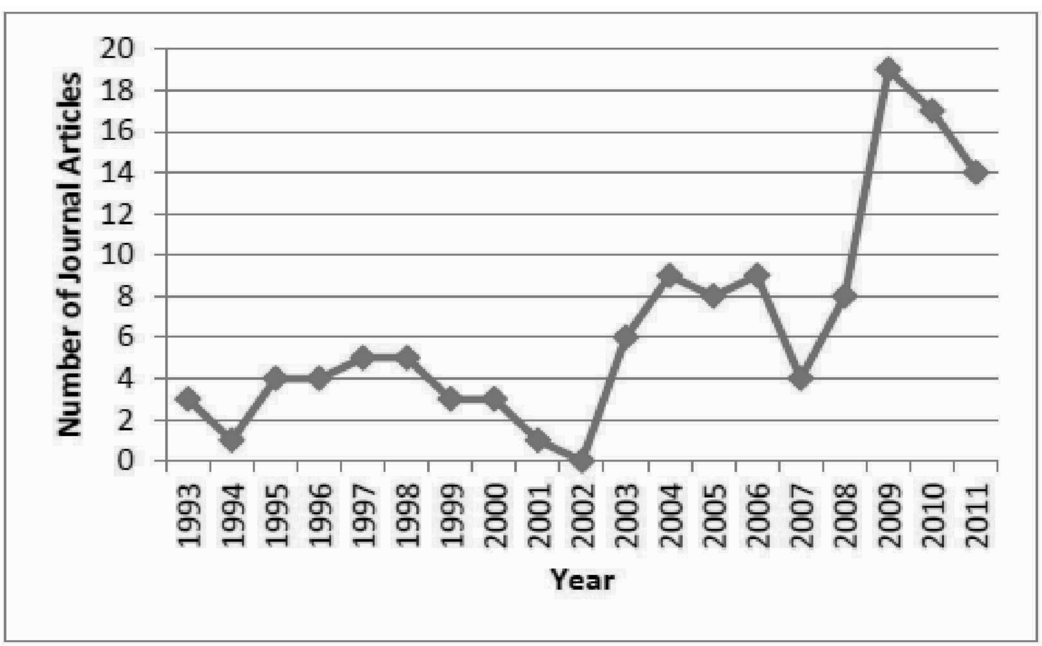
Table 1: Overview of QCA Software

	Graphical User Interface			Command Line Interface		
Function*	Tosmana[a]	fs/QCA[b]	KirqST[c]	fuzzy[d]	QCA3[e]	QCA[f]
				Variant		
csQCA	●	●	●	●	●	●
tQCA	○	●	○	○	●	●
mvQCA	●	○	○	○	●	●
fsQCA	○	●	●	●	●	●
				Solution Types		
complex	●	●	●	●	●	●
intermediate	○	●	◉	○	◉	●
parsimonious	●	●	●	●	●	●
				Procedures		
Calibration	●	◉	◉	◉	◉	●
Necessity	○	◉	●	◉	◉	●
Factorization	●	○	○	○	◉	●
(C)SA	●	○	○	○	●	●
Parameters of fit	○	●	●	●	●	●
Statistical tests	○	○	○	●	●	○

*Circle Legend: ● Full functionality; ◉ partial functionality; ○ almost no / no functionality

[a] Version 1.3.2.0; [b] Version 2.5; [c] Version 1.8; [d] Version st0140_2; [e] Version 0.0-5; [f] Version 1.0-4

Figure 1: QCA Journal Articles in Comparative Politics, International Relations, and Sociology Per Year



type mistake would typically pass unnoticed in a GUI. On the other side, however, the learning curves of GUIs are considerably lower for CLI-literate users than those of CLIs for GUI-accustomed users. The application of QCA is thus not "made much easier because by now it is possible to perform QCA in **Stata** or **R**" as Grofman and Schneider (2009: 670) argue. The appearance of CLIs just creates a possibility of choice for users between a flat learning curve and little flexibility or a steep learning curve and much flexibility.

With the **QCA** package, we have sought to strike a workable balance, reconciling the power of CLI software with a minimum of programming knowledge. As much of the programming part is hidden behind the commands, sessions early in a user's **R** career need not be frustrating, whereas more advanced users may want to increasingly exploit the full capabilities of the package. In the remainder of this article, we briefly illustrate **QCA**'s workflow by comparison with common analytical steps in **fs/QCA**.

The QCA Package

In this section, we reanalyze the data from Bochsler's (2012) csQCA study on intra-group ethnic party competition.[3] The reader shall be referred to the origi-nal article for all details. Here, we simply list the condition and outcome sets. The outcome indicates whether several political parties representing the same ethnic minority exist in the national parliament (BI: 1-yes, 0-no), and the conditions measure whether national electoral thresholds exist (TH: 1-yes, 0-no), whether votes take place in special ethnic districts by proportional representation (SP: 1-yes, 0-no), whether minority groups form a majority at the local or regional level (MA: 1-yes, 0-no), whether the minority group is territorially concentrated (CO: 1-yes, 0-no), whether the population share of the minority corresponds to two or more seats in parliament (PA: 1-yes, 0-no) and whether the minority population share equals two or more seats in an average district (DI: 1-yes, 0-no).

The command structure of **QCA** has been designed with a view to presenting end-users with only as few commands as necessary, but offering them as much flexibility as possible in how to exploit and combine the objects returned by these commands. The core of **QCA** therefore includes only five functions: *calibrate* for crisp and fuzzy set calibration, *superSubset* for automated superset and subset analyses, *truthTable* for the construction of truth tables, *eqmcc* for the minimization procedure and *pof* as a generic function for computing common QCA parameters of fit. In this brief replication, we illustrate the use of *calibrate*, *truthTable*, and *eqmcc*.

Bochsler begins his analysis by calibrating TH, DI and PA from each set's continuous base variable. In **fs/QCA**, the menu for recoding variables can be used as shown in the left panel of Figure 2. **QCA** provides the extremely flexible *calibrate* function for this purpose as shown in the right panel of Figure 2. This function can be used to calibrate binary and multi-level crisp sets and offers a vast array of different membership functions for calibrating fuzzy sets, including linear, logistic, exponential, and empirical distribution functions. In order to achieve the calibration of TH, DI, and PA, it only requires two arguments: the base variable and the threshold. As the number of such operations increases, calibration in **QCA** will take less time than in **fs/QCA** because code can be easily copied and adapted in CLI software.

After all conditions and the outcome have been specified, **fs/QCA** presents users with options to code the truth table as shown in the left panel of Figure 3. Most efficiently, the "Delete and code…" command in the Edit tab allows the quick categorization of configurations into logical remainders, positive configurations, and negative configurations. Building truth tables is as straightforward in **QCA**. The *truthTable* function permits the coding of logical remainders, negative and postive configurations, as well as contradictions in all QCA variants. The coding of contradictions is made easy by specifying the cut-off for positive configurations—*incl.cut1*—plus the cut-

Figure 2: Calibration of Crisp Sets in fs/QCA and QCA

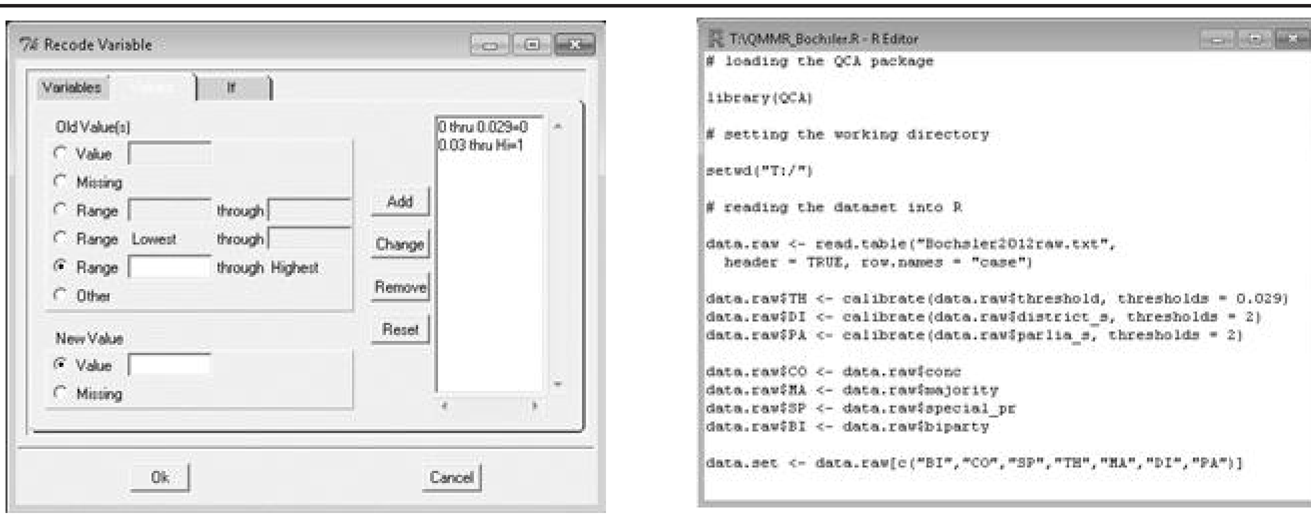
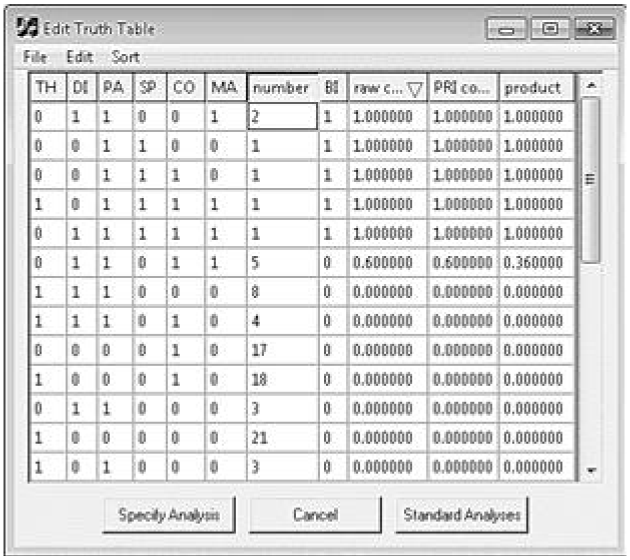


Figure 3: Truth Table Construction in fs/QCA and QCA

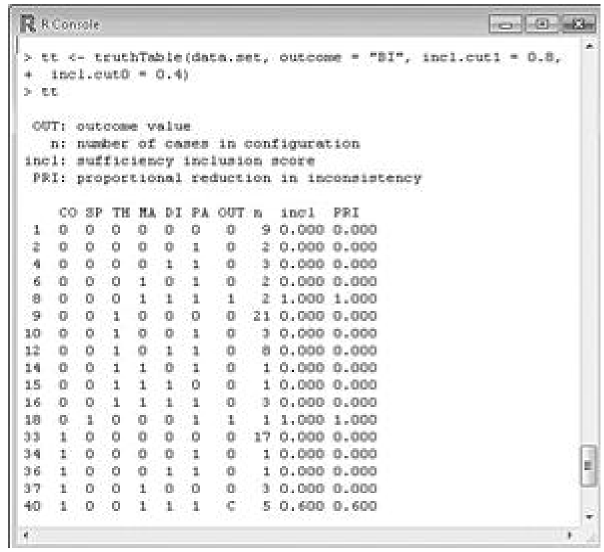


Edit Truth Table

File Edit Sort

TH	DI	PA	SP	CO	MA	number	BI	raw c...	PRI co...	product
0	1	1	0	0	1	2	1	1.000000	1.000000	1.000000
0	0	1	1	0	0	1	1	1.000000	1.000000	1.000000
0	0	1	1	1	0	1	1	1.000000	1.000000	1.000000
1	0	1	1	1	1	1	1	1.000000	1.000000	1.000000
0	1	1	1	1	1	1	1	1.000000	1.000000	1.000000
0	1	1	0	1	1	5	0	0.600000	0.600000	0.360000
1	1	1	0	0	0	8	0	0.000000	0.000000	0.000000
1	1	1	0	1	0	4	0	0.000000	0.000000	0.000000
0	0	0	0	1	0	17	0	0.000000	0.000000	0.000000
1	0	0	0	1	0	18	0	0.000000	0.000000	0.000000
0	1	1	0	0	0	3	0	0.000000	0.000000	0.000000
1	0	0	0	0	0	21	0	0.000000	0.000000	0.000000
1	0	1	0	0	0	3	0	0.000000	0.000000	0.000000

Specify Analysis | Cancel | Standard Analyses



R Console

```
> tt <- truthTable(data.set, outcome = "BI", incl.cut1 = 0.8,
+  incl.cut0 = 0.4)
> tt

 OUT: outcome value
   n: number of cases in configuration
incl: sufficiency inclusion score
 PRI: proportional reduction in inconsistency

    CO SP TH MA DI PA OUT n  incl  PRI
 1  0  0  0  0  0  0   0   9 0.000 0.000
 2  0  0  0  0  0  1   0   2 0.000 0.000
 4  0  0  0  0  1  1   0   3 0.000 0.000
 6  0  0  0  1  0  1   0   2 0.000 0.000
 8  0  0  0  1  1  1   1   2 1.000 1.000
 9  0  0  1  0  0  0   0  21 0.000 0.000
10  0  0  1  0  0  1   0   3 0.000 0.000
12  0  0  1  0  1  1   0   8 0.000 0.000
14  0  0  1  1  0  1   0   1 0.000 0.000
15  0  0  1  1  1  0   0   1 0.000 0.000
16  0  0  1  1  1  1   0   3 0.000 0.000
18  0  1  0  0  0  1   1   1 1.000 1.000
33  1  0  0  0  0  0   0  17 0.000 0.000
34  1  0  0  0  0  1   0   1 0.000 0.000
36  1  0  0  0  1  1   0   1 0.000 0.000
37  1  0  0  1  0  0   0   3 0.000 0.000
40  1  0  0  1  1  1   C   5 0.600 0.600
```

off for negative configurations—*incl.cut0*. In our example, these two cut-offs are set to 0.8, 0.4 respectively, thereby producing the contradictory configuration in line 40 of the truth table that is shown in the right panel of Figure 3. These configurations can later be included into or excluded from Boolean minimization.

The derivation of solutions via the "Standard Analyses" button in **fs/QCA** triggers the window for the specification of directional expectations with regards to each individual condition. These expectations determine which logical remainders will be made available for reduction. If no logical remainders are used in the minimization, the result is usually referred to as the complex solution. If all logical remainders which allow the further simplification of the solution are used, the result is called the parsimonious solution. And for all situations between these two extremes, the result is called an intermediate solution.

The derivation of solutions is achieved in **QCA** with the *eqmcc* function (enhanced Quine-McCluskey) (Dusa 2007). This is shown in the right panel of Figure 4. The truth table object *tt* created by *truthTable* is passed to *eqmcc* together with all options the user wants to specify. Here we also include logical remainders with *include = "?"*. Although Bochsler's argument is more complex, suppose that all conditions are expected to contribute to a positive outcome value when present, apart from TH, which is expected to contribute when absent. The specification of these directional expectations is easily achieved by providing the optional argument *direxp = c(1,1,0,1,1,1)*, where each value indicates the Boolean value that corresponds to the expectation: 0 stands for "absence," 1 for "presence," and –1 for a "don't care."

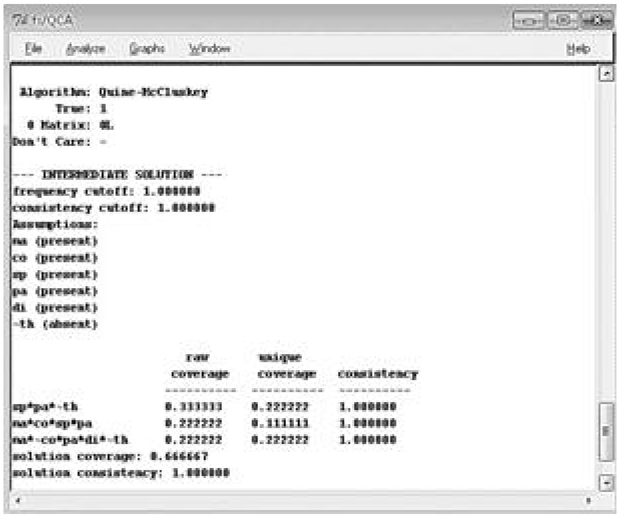
Other optional arguments are available. If *details = TRUE*, the printed output will also include the parameters of fit for each prime implicant (PI) and the minimal sum. In our case, the solution only consists of one minimal sum, but often two or even more PIs cover a specific truth table row none of the other PIs covers, while they are not essential for covering other rows. In these cases, **fs/QCA** pops up a PI chart window in which the user is asked to choose the inessential PI(s) that complete the solution. **QCA** simplifies this procedure insofar as it derives all minimal sums if *rowdom = FALSE*, and only minimal sums with the row-dominating PIs if *rowdom = TRUE*. It is not revealed in **fs/QCA** whether or not some PI dominates another.[4]

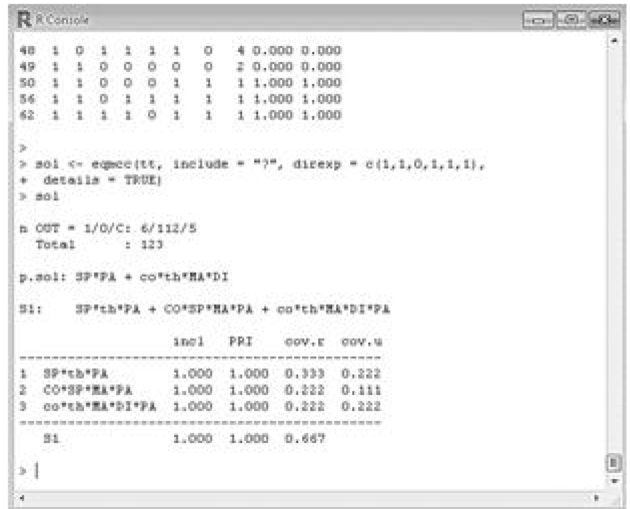
For each PI, the cases with membership above 0.5 can also be shown if *show.cases = TRUE* is passed to eqmcc.[5] Further arguments to *eqmcc* include *omit* for the explicit omission of specific configurations (remainders or even positive configurations), *neg.out* for the negation of the outcome, *use.tilde* and *use.letters* for formatting Boolean negations by tilde or lowercase letters, and *all.sol* for deriving all non-overlapping minimal sums irrespective of the complexity of their PIs or their number.

Conclusion

In this article, we have provided an overview of the QCA software market and an introduction to the **QCA** package for **R**. It is to be highly welcomed that users now have the choice between six different programs, three GUIs and three CLIs. And although **fs/QCA** will very likely remain the clear market leader for years to come, we expect a gradual diversification towards CLIs, particularly among the more technically-inclined user segment. Moreover, the proficiency of course instructors and the individual willingness of users to invest in acquiring the skills necessary to operate a syntax-based program represent the main factors in the future distribution of software market shares. CLIs should be well-positioned in this respect. Their scope of functionality now exceeds that of all GUIs, textbooks about how to perform QCA in **R** have been published (Thiem and Dusa 2013), extensive documentation replete with

Figure 4: Derivation of Intermediate Solution in fs/QCA and QCA





many practical examples accompanies all CLIs, and their transparency as open-source software allows peer-review and increases developer responsiveness.

Notes

[1] **QCA** also takes care of excluding auxiliary conditions for temporal information from the computation of parameters of fit.

[2] It is very likely that those studies which did not cite the software they used also relied on **fs/QCA**, or any of its predecessors, in particular before **Tosmana** was released in 2004.

[3] The dataset is available on www.compasss.org.

[4] One PI P_1 is said to dominate another P_2 if all fundamental products covered by P_2 are also covered by P_1 and both are not interchangeable.

[5] The same argument can also be passed to the *truthTable* function.

References

Bochsler, Daniel. 2012. "When Two of the Same Are Needed: A Multilevel Model of Intragroup Ethnic Party Competition." *Nationalism and Ethnic Politics* 18:2, 216–241.

Cronqvist, Lasse. 2011. *Tosmana: Tool for Small-N Analysis* [Version 1.3.2.0]. Trier: University of Trier.

Dusa, Adrian. 2007. "Enhancing Quine-McCluskey." COMPASSS: Working Paper 2007–49. Available at www.compasss.org/wpseries/Dusa2007b.pdf.

Dusa, Adrian and Alrik Thiem. 2012. *QCA: Qualitative Comparative Analysis*. R Package Version 1.0–4.

Grofman, Bernard and Carsten Q. Schneider. 2009. "An Introduction to Crisp Set QCA, with a Comparison to Binary Logistic Regression." *Political Research Quarterly* 62:4, 662–672.

Huang, Ronggui. 2012. *QCA3: Yet another Package for Qualitative Comparative Analysis*. R Package Version 0.0–5.

Longest, Kyle C. and Stephen Vaisey. 2008. "fuzzy: A Program for Performing Qualitative Comparative Analyses (QCA) in Stata." *Stata Journal* 8:1, 79–104.

Ragin, Charles C. and Sean Davey. 2009. *fs/QCA: Fuzzy-Set/Qualitative Comparative Analysis* [Version 2.5]. Tucson: University of Arizona.

Reichert, Christopher and Claude Rubinson. 2012. *KirqST* [Version 1.8]. Houston: University of Houston-Downtown.

Thiem, Alrik and Adrian Dusa. 2013. *Qualitative Comparative Analysis with R: A User's Guide*. New York: Springer.

Announcements

Interpretive Methods for Grant Proposal Development Workshop at University of California, Irvine May 30–31, 2013

This two-day workshop for graduate students and junior faculty will focus on how to write grant proposals for research projects that utilize ethnographic methods and/or discourse analysis. Led by UCI faculty, administrators, and non-UCI specialists in interpretive methodologies, the workshop will focus on how to employ ethnography and discourse analysis in research, basic principles of grant-writing and how they apply to interpretive research, common mistakes found in grant proposals, and how to address review criteria. Participants will be expected to submit a short draft grant proposal based on ethnographic or discourse analytic methods prior to attending the workshop. Further information on the workshop, including application details, will be available on the UCI International Studies website http://internationalstudies.ss.uci.edu/ on January 2, 2013. The final submission deadline is March 1, 2013. Feel free to contact Tanya Schwarz at tschwarz@uci.edu with any questions.

Best Qualitative and Multi-Method Submission to the *American Political Science Review* in the Preceding Calendar Year

Recipients: Jeremy Menchik, Boston University, "The Origins of Intolerance in Islamic Institutions"; and

Paul Staniland, University of Chicago, "States, Insurgents, and Wartime Political Orders."

Committee: Kathleen Thelen, Massachusetts Institute of Technology (chair); Colin Elman, Syracuse University; and Elisabeth Wood, Yale University.

This section award was created to interrupt the self-reinforcing cycle whereby the absence of qualitative and multi-method research in the *APSR* led to the widespread belief that such work was not welcome. The lack of submissions then of course confirmed that understanding. A dialogue with the previous lead editor, Ron Rogowski, resulted in the section deciding to offer an additional incentive to authors to submit to the journal, in an effort to break that negative cycle. The committee unanimously agreed that the two recipients of this year's awards were excellent manuscripts, thoroughly deserving of the award.

Staniland's article manuscript on "States, Insurgents, and Wartime Political Orders," offers a conceptual typology of political orders amidst civil war, focusing on the different types of relationships between insurgents and states in conflict zones. The essay convincingly argues that these relationships need to be problematized and theorized. The committee was taken with Staniland's argument that recognizing and articulating differences in these relationships across time and space allows the posing of new questions and/or old questions in new ways. This agenda-setting paper strongly advances our understanding of civil war and state building as well as to the broader fields of conflict and its resolution. It will be a much-cited article. Among other achievements, it demonstrates the importance of understanding the dynamics of war in a broader setting, broader in both the sense of drawing on relevant material in various disciplines and broader in the sense of analyzing little known cases. While the author modestly makes no claim to advancing theory, the typology and its exposition have clear theoretical implications.

Menchik's article manuscript on "The Origins of Intolerance in Islamic Institutions" is an analysis of variation in levels of tolerance toward non-Muslim populations across three different Islamic institutions in Indonesia. Menchik argues that theology is a poor predictor of attitudes and behaviors, but more importantly he traces variation in attitudes/behavior to the kinds of social cleavages and conflicts that dominated at the time of the founding of these different institutions. It is an argument about path dependence, but unlike many such arguments that do not provide much insight into how events in the dim past continue to influence outcomes later, he suggest two mechanisms, namely the way in which Islamic jurisprudence becomes institutionalized and patterns of political alliances that cement and reinforce particular cleavages. The committee was particularly taken with Menchik's use of evidence, which resulted from first-rate archival work among primary sources in a challenging setting.

Giovanni Sartori Award for Best Book on and/or Using Qualitative Methods

Recipient: Alan M. Jacobs, *Governing for the Long Term: Democracy and the Politics of Investment*. (Cambridge University Press 2011).

Committee: Lauren Morris MacLean, Indiana University (chair); Evan Lieberman, Princeton University; and Timothy Crawford, Boston College.

The committee has decided to confer the Sartori Award for 2012 on *Governing for the Long Term: Democracy and the Politics of Investment* by Alan M. Jacobs. In *Governing for the Long Term*, Jacobs asks a critical question about the extent to which political actors pursue long-term policies that may not reap tangible benefits within an election cycle. Jacobs brilliantly capitalizes on prior work in comparative political economy and other domains of political science, showing how leader's economic beliefs and schemas help them navigate the inherent uncertainty and complexity of economic life. As Jacobs shows us, such uncertainties are central to the politics of long-range social policies, and leaders' "mental models" critically define both the nature of the problems and the potential solutions at stake.

Jacobs' theoretical insights into the institutional, inter- and intra-interest group politics that make it possible for governments to adopt far-sighted policies are original and heuristically powerful. Their implications reach beyond the specific domain of policy investments in social welfare programs, suggesting new ways to understand domestic political conditions that foster (or hinder) commitments to many other kinds of long-range projects with potentially large benefits over time, but certain and painful up-front costs to constituents, such as those that have arisen or might arise in international economic, environmental, and security affairs. We considered Jacobs' work to be outstanding in terms of his implementation of case study and comparative-historical methods. The book is based on a theoretically informed research design, effectively comparing the origins and subsequent dynamics of public pension-making in four countries: the U.S., Canada, UK, and Germany. His analyses are based on a mix of original interviews, primary archival sources, and secondary literature. In presenting these materials, he is extraordinarily transparent in providing the reader with the likely quality of those sources. For example, when citing various scholars, he goes way beyond the oft-practiced vague citation of author-year, and provides a clear summary of the nature of the data and analyses upon which other authors have generated their conclusions. Through his theoretical framework and compelling use of empirical evidence, Jacobs is able to reveal the process of causal inference used by policymakers and empirically substantiate the role of ideas in policymaking. The committee concluded that this work is an extremely important piece of comparative political analysis, an insightful study of the inter-temporal tradeoffs in public policymaking, and an exemplar of the use of qualitative methods.

Alexander George Award for Best Article or Book Chapter on and/or Using Qualitative Methods

Recipient: Anna Grzymala-Busse, "Time Will Tell? Temporality and the Analysis of Causal Mechanisms and Processes." *Comparative Political Studies* 44:9 (September 2011), 1267–1297.

Committee: Melani Cammett, Brown University (chair); Hillel Soifer, Temple University; and Gerardo Munck, University of Southern California.

Anna Grzymala-Busse's article, "Time Will Tell: Temporality and the Analysis of Causal Mechanisms and Processes'' (*Comparative Political Studies*, 2011), is an ambitious and broadly applicable analysis of one of the most central yet elusive issues in the social sciences: how to factor time into explanations. Grzymala-Busse demonstrates convincingly that most research—even the best comparative historical work—conflates distinct dimensions of temporality and lumps them under the category of "history'' or "path dependence,'' thereby undercutting our ability to isolate causal mechanisms. Through careful conceptual work grounded in clear empirical examples, her analysis shows that temporality should be subdivided into four different dimensions—duration, tempo, acceleration, and timing—and that each of these dimensions has distinct implications for causal processes. An understanding of the varied dynamics of these processes and how they play out in specific historical contexts can allow researchers to predict causal mechanisms, identify sequences of events and social phenomena, and generate more precise explanations. Grzymala-Busse's article breaks new ground in conceptual analysis and methodology by demonstrating the imperative of disaggregating temporal processes in social science research, advancing efforts to think systematically about comparative historical analysis and honing in on issues germane to causal explanation. The insights of this article will surely be felt on researchers for years to come.

Sage Award for Best 2011 APSA Paper on and/or Using Qualitative Methods

Recipients: Derek Beach and Rasmus Brun Pedersen, "What is Process Tracing Actually Tracing? The Three Variants of Process Tracing Methods and their Uses and Limitations." Paper prepared for Presentation at the Annual Meeting of the American Political Science Association, Seattle, Washington, September 1–4, 2011.

Committee: Jennifer Hadden, University of Maryland (chair); Jonathan Githens-Mazer, University of Exeter; and Erica Townsend-Bell, University of Iowa.

The selection committee is pleased to award the 2012 Sage Award to Derek Beach and Rasmus Brun Pedersen for their paper "What is Process Tracing Actually Tracing? The Three Variants of Process Tracing Methods and Their Uses and Limitations.'' This paper highlights three variants of process-tracing methodology and their uses:

(1) Theory-testing process tracing, which uses process tracing to see if the observable implications of a theorized causal mechanism are present in a case;

(2) Theory-building process tracing, which builds theory about causal mechanisms from process tracing within a case; and

(3) Outcome explanation process tracing, which uses process tracing to craft a minimally sufficient explanation of a case of interest.

It elaborates each of these, and provides illustrations of how they can be appropriately used in political science research. The committee appreciated the creativity and clarity of presentation of this paper. In particular, we consider that the paper provides useful guidance for how to use process tracing methods to accomplish very different research goals. The authors carefully consider how one might use process-tracing methods on either side of the nomothetic-idiographic divide, as well as in the context of mixed-method research designs. We would like to congratulate Beach and Pedersen on producing a thoughtful and theoretically useful contribution on this important topic.

David Collier Mid-Career Achievment Award

Recipient: Colin Elman, Syracuse University

Committee: Gary Goertz, University of Arizona (chair), Dan Carpenter, Harvard University, Rose McDermott, Brown University.

The committee has awarded Colin Elman the 2012 David Collier Mid-Career Achievement Award. The Award honors David Collier's contribution to substantive research, methodological publications, and institution building. Elman has major achievements in all three of these areas. These successes are more typically found in someone much further advanced in his career, and the choice of Elman was an easy one for the committee. On the institutional front, Colin was a co-founder of two APSA Sections, International History and Politics, and Qualitative and Multi-Method Research. Elman has served as the QMMR section's President and Secretary-Treasurer, and he has managed many activities of the section, including APSA short-courses, business meetings, and its annual awards. In terms of leadership as well as operational activities the section would certainly not be the success it is (with membership usually in the top five of APSA sections) without the time and energy that Elman has invested in institution building. Hundreds of young qualitative methods scholars know Colin for his leadership (with Andrew Bennett and David Collier) of the Institute for Qualitative and Multi-Method Research. Now in its twelfth year, well over 1,000 students have participated in these annual institutes. Elman is also one of the founders of the new Qualitative Data Repository. Both of these endeavors have been supported by substantial funding from the National Science Foundation. Elman is also co-chair (with Arthur Lupia) of APSA's working group on Data Access and Research Transparency, and (with John Gerring and James Mahoney) a general editor of Cambridge University Press' new series *Strategies for Social Inquiry*. Elman is an accomplished qualitative scholar and methodologist. He has published on Lakatosian metatheory, explanatory typologies, the role of historiography in political science, case studies, causal complexity, and qualitative data and replication. These articles, as well as his substantive research, have appeared in *Political Analysis*, *Comparative Political Studies*, the *Annual Review of Political Science*, *International Organization*, and the *American Political Science Review*.

2013 Nominating and Award Committees

2013 Nominating Committee: Tim Buthe (chair), Duke University; Deepa Prakash, Depauw University; Elizabeth Saunders, George Washington University; and Joachim Blatter, University of Lucerne.

Sartori Book Award Committee: Ingo Rohlfing, Cologne Graduate School of Management (chair); Jason Seawright, Northwestern University; and Richard Ned Lebow, Dartmouth College and King's College London.

George Article-Chapter Award Committee: Anna Grzymala-Busse, University of Michigan (chair); Candice Ortbals, Pepperdine University; and Ariel I. Ahram, Virginia Tech.

Sage APSA paper Award Committee: Markus Kreuzer, Villanova University (chair); Isabella Alcaniz, University of Maryland; and Anne-Marie D'Aoust, Université du Québec à Montréal.

Qualitative and Multi-Method Research
The George Washington University
Department of Political Science
Monroe Hall 440
2115 G Street, N.W.
Washington DC 20052
USA

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

Qualitative and Multi-Method Research (ISSN 2153-6767) is edited by Robert Adcock (tel: 510-301-4512, fax: 202-994-7743, email: adcockr@gwu.edu). The production 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.