

12-1-2007

Beyond foraging: behavioral science and the future of institutional economics

Alexander J. Field

Santa Clara University, afield@scu.edu

Follow this and additional works at: <http://scholarcommons.scu.edu/econ>

 Part of the [Economics Commons](#)

Recommended Citation

Field, Alexander J. 2007. "Beyond foraging: behavioral science and the future of institutional economics." *Journal of Institutional Economics* 3 (December): 265-291.

Copyright © The JOIE Foundation 2007. Reprinted with permission.

The final paper can be found at - <https://doi.org/10.1017/S1744137407000720>.

This Article is brought to you for free and open access by the Leavey School of Business at Scholar Commons. It has been accepted for inclusion in Economics by an authorized administrator of Scholar Commons. For more information, please contact rscroggin@scu.edu.

forthcoming in the Journal of Institutional Economics 3 (December 2007)

**Beyond Foraging:
Evolutionary Theory, Institutional Variation, and Economic Performance**

by

Alexander J. Field
Department of Economics
Santa Clara University
Santa Clara, CA 95053
afield@scu.edu

Institutions affect economic outcomes, but variation in them cannot be directly linked to environmental factors such as geography, climate, or technological availabilities. Game theoretic approaches, based as they typically are on foraging only assumptions, don't provide an adequate foundation for understanding the intervening role of politics and ideology. Nor does the view that culture and institutions are entirely socially constructed. Understanding what institutions are and how they influence behavior requires an approach that is in part biological, focusing on cognitive and behavioral adaptations for social interaction favored in the past by group selection. These adaptations, along with their effects on canalizing social learning, help to explain uniformities in political and social order, and are the bedrock upon which we build cultural and institutional variability.

JEL Codes: A12, B15, B52, C7, N01, Z13

Keywords: Institutions, economics and biology, group selection, game theory and experimental economics, economics and psychology

Introduction

Economic history is often defined as the historical study of economic growth, and one of the distinguishing features of the new economic history has been a more systematic application of quantitative data, as well as theoretically based tools, such as growth accounting, to this task (see, for example, Field, 2003, 2006b). But one of the central challenges of comparative historical inquiry has proved only indirectly susceptible to the application of quantitative techniques. This is the role of explicit (written or otherwise formally articulated) and informal institutions, and more broadly, non-material aspects of culture, in influencing economic outcomes.

The purpose of this essay is to reflect on what we have learned to date about this issue, and what approaches and lines of inquiry are likely to be illuminating in the future. While we are making methodological and substantive progress on macro-level explorations of the consequences of institutional variation, attempts to provide a satisfactory microanalytics, one that explains systematically how institutions are created, how they are sustained, and why they vary, have been far less successful. A recurring problem is that institutions cannot be both consequential for economic outcomes and at the same time entirely derivative of such givens as technologies and resource endowments.

The approaches commonly favored by economists have often simply provided ex post rationalizations of already known explananda. Paired with “illustrative” game theoretic models, these narratives can acquire the patina of scientific theory without satisfying its requirements: the ability to predict or explain out of sample. To the degree that we wish to root explanations of culture or institutions in characteristics of individual

human beings, we need much more systematically to incorporate findings from other fields, including evolutionary theory, behavioral genetics, neuroanatomy and neurobiology, and developmental, cognitive, and social psychology.¹

We need in particular to acknowledge – as foundational and not just peripheral -- a broader range of motivations or behavioral predispositions than has been typical in the canonical selfish actor rational choice approach, and have a better empirical understanding of when they do and do not kick in and why.² In investigating how institutions are created and perpetuated, we should recognize the significance of social learning and the cultural plasticity this enables. At the same time, we should appreciate the extent to which we are differentially prepared to learn in certain directions, and the degree to which this helps account for uniformities in human culture and provides some limits to its variability.

This essay is not simply a call for economists to be more catholic in our intellectual interests. It stresses the importance of reading in other disciplines, but reading critically. In particular, much of the developing effort to integrate the social, behavioral, and natural sciences has been hobbled by implausible restrictions on the levels at which natural selection could have operated in the past. For reasons related to four decades of contentious intellectual debate, many biologists and biologically oriented anthropologists and evolutionary psychologists reject the possibility of empirically important selection above the level of the individual organism. Their position leaves unexplained how an important set of behavioral and cognitive adaptations for social interaction could have

¹ The “we” here refers broadly to social scientists interested in understanding institutions.

² The rational choice approach can, in principle, accommodate any goals or preferences. As an empirical matter, however, preferences are often assumed to be stable, transitive, and selfish in the restrictive sense that human agents are assumed to act so as efficiently to advance their material welfare

emerged. It reinforces the restrictive motivational assumptions that limit the explanatory power of models often favored by economists. And it discourages others from even considering the possibility of important genetic or biological influences on cognition and behavior. It is a roadblock in the path of integrating the social and natural sciences, and weaning sociology and anthropology from its social constructivist traditions. Further advance requires the integration of insights from the biological sciences. But the reading of the biological literature, particularly in evolutionary theory, needs to be done carefully. In particular, social scientists need to understand the levels of selection controversy, and what is at stake in it for our own disciplines (see Wilson and Sober, 1994; Field, 2001, 2006a, c, Wilson and Wilson, 2006).

The Central Empirical Problem

The starting point is the claim a) that the institutional, cultural, or normative structure of a group influences individual behavior and economic outcomes (Field, 1984, 1991), and b) that this influence is not simply derivative of more fundamental environmental variables such as geography, climate, or technological availabilities. Both old and new institutional economists have maintained in a general sense that institutions matter, but if institutional variation could be accounted for entirely with reference to technology or geography, it would hardly be worth studying such variation in its own right. If there were a one to one mapping from current environment to normative structures, such structures could have no independent influence on outcomes.

A growing body of historical and cross national research makes it clear that institutional/cultural structures often persist for long periods in the face of major changes in recognizable features of the material environment. Yet they also sometimes change

rapidly in the absence of any easily identifiable environmental shock (Kuran, 1995). One of the principal reasons we study normative structures and their consequences is because they cannot necessarily be directly linked to currently prevailing environmental conditions.

This central empirical fact has produced a schizophrenia in institutional economics, reflected in the divide between the “old” and the “new”. Although it is difficult sometimes to say exactly what the new institutional economics is, the distinction has commonly revolved around whether the focus of inquiry should be principally upon the causes, as opposed to the consequences of institutional variation. “Old” institutional economics, as in the work of John R. Commons, assumed, without much discussion, that institutions were stable over long periods, changing only when subject to major exogenous shocks such as a Supreme Court decision, or military invasion. By assuming that such shocks were few and far between, institutional economics could be freed to focus most attention on the consequences of variation, with a case specific narrative detailing the nature of the exogenous changes.

Many believe that the old institutional economics is dead, buried under the impressive achievements of the new. But the rumors of its demise (and, arguably, about the achievements of the new) have been greatly exaggerated. The traditions of the old institutional economics are not, however, being carried forward today so much by individuals who advertise themselves as the heirs of Commons, but by a group of macro and growth economists exploiting newly available sets of cross country data. Examples include Acemoglou, Johnson, and Robinson (2001) or Hall and Jones (1999).

Like old institutional economists, these authors declaim much interest in the microanalytics of institutional formation. The basic analytical structure is this. In different parts of the world, different institutional/cultural structures emerge. These structures then persist for long periods, essentially unchanged, and influence differences in economic outcomes today. To the degree that these authors consider feedback from the current environment to institutional structure, or address how well they have measured institutional structure, the proposed methodological solution is econometric: the use of instruments to generate predicted values for the institutional variable(s) as a means of eliminating simultaneity and errors in variables bias.

In Acemoglu, Johnson, and Robinson (2001), for example, the collision of the rest of the world with European civilization created one type of institutional structure where settler mortality was low, and another type where it was high. These different structures persisted and now, the authors argue, explain most of the divergence in per capita income in formerly colonial regions today. More specifically, current levels of GDP per capita are strongly related to institutional measures of security of property, which the authors in turn instrument using nineteenth century and earlier measures of settler mortality. The argument is that where the disease environment allowed Europeans to migrate without dying in great numbers, they established European style institutions which provided security of property and protection against governmental predation. Elsewhere Europeans focused on extraction of natural resources, and built or allowed a different type of institutional structure. The varying institutional forms then persisted – strongly influencing levels of economic activity today. Thus they conclude that Africa today is largely poor not (directly) because of its geography or current disease environment, but

because of its institutions. A similar approach is reflected in Hall and Jones (1999: 86, 114), where institutional structure is largely predetermined by location, and where causality runs principally from institutions to output per worker rather than vice versa.

Many of the authors associated with the “new” institutional economics, including North and Williamson, have criticized the old economic history on the grounds that it encouraged a descriptive approach and ad hoc explanation. There are worse things for an explanation to be, however, than ad hoc, and that is for it to yield predictions systematically contravened by the data, or, worse, to be simply tautological. It is, for example, now largely accepted that the demise of dinosaurs approximately 65 million years ago was due to the collision of the earth with a large asteroid off the coast of the Yucatan peninsula. The resulting atmospheric debris led to climate cooling, a reduction of vegetation, and mass extinction. Such an explanation may be ad hoc, but as far as we can currently tell, it is not wrong.

Nevertheless, the new institutional economics has promised a more theoretically informed approach, one that would account for institutional change as well as stability, and do so without appealing to the equivalent of asteroid strikes (Mathews, 1986: 903; Williamson, 2000: 595). The first incarnations of the new institutional economics can be thought of as an attempt to rewrite Marx without the class struggle. The hope was that one could peer below the surface variation of institutional and social structure over time and space and perceive underlying causes in environmental determinants such as land labor ratios or technological availabilities (North and Thomas, 1973). This approach very quickly ran afoul of the central institutional fact described above: institutions often persist for long periods in the face of major changes in recognizable features of the

environment, yet sometimes change rapidly in the absence of any easily identifiable environmental shock (Field, 1981).

When the problems with this approach became evident, North retained an emphasis on understanding the *causes* of institutional variation, but allowed that one had to acknowledge an intervening role for political and ideological factors. This perspective was combined with the claim that new analytical tools, particularly game theory, could give us “insight” into such variation. Allowing an intermediate link for politics and ideology, however, severed any direct link between the environment and institutional structure. That was the price that had to be paid to reconcile the approach with the central institutional fact.

One of the key problems, then, is to understand when institutions change in response to changed environment, and when they don't. It is always easy *after* the fact to explain why a Supreme Court decision went the way it did. A truly scientific approach to politics and institutional change would require that we make *ex ante* predictions as well as *ex post* “explanations.” These kinds of concerns are largely ignored by authors in the macroinstitutional arena, as well as their critics (e.g., Glaeser et al, 2005), because they are not by and large concerned with causes except as largely exogenous givens. What we see instead is genuine bravado as the researchers plunge into the struggle to extract meaning from the cross national data. The benefit has been the generation of some real meat upon which we can chew. But the questions of what kinds of generalizations we can make about the persistence of different types of institutions, how resistant or susceptible they might be to changes in technological availabilities or the organized

campaigns of powerful interest groups, and why, remain relatively unilluminated in this literature.

New institutional economists have tried to be more precise about microanalytics, but the efforts don't stand up well to critical scrutiny. Williamson (2000), for example, distinguished between four levels of institutions, ranked according to how sensitive they were to environmental alterations, and claiming that as one moved through the levels, the time scale of persistence decreased, and the role of "calculative" behavior increased. By allowing that "noncalculative" behavior influenced persistence at the highest level (informal institutions, religion, and norms), he sought to preserve an unencumbered role for "calculative" behavior at levels below. This would imply that a mapping from environment to institutions might still be possible at the lower levels, even if it could not be so sustained at the highest, and that the microanalytics of such an explanatory framework could be constructed on the basis of the cognitive and behavioral predispositions that typically underlay economic analysis.

But Williamson simply asserts that "calculative" behavior increases as one moves down his hierarchy. It is not obvious, for example, that informal institutions necessarily persist for hundreds and thousands of years, and that they are on average more persistent than, for example, contract law. Suppose most informal institutions last but a short time while a few persist over longer periods. As institutional historians our attention would be drawn to those in the latter category, and we could mistakenly make the inference suggested by Williamson.

Williamson begins his essay by acknowledging that we are "ignorant of institutions". He argues, nevertheless, that much effort had gone into constructing our

current state of ignorance, maintaining that the new institutional economics has progressed “not by advancing an overarching theory but by uncovering and explicating the microanalytic features ... and by piling block upon block until the cumulative value cannot be denied” (Williamson, 2000: 596). If we remain skeptical as to what has actually been accomplished, he asks for our critical forbearance and makes an analogy with graduate school admissions:

Good but underdeveloped ideas are evidently like good but underdeveloped minds: both are precious things.... Once admitted, theories, (and graduate students) are progressively built up – moving from less formal to more formal stages of development. Finally, as with promising graduate students, we do not hold on to cherished theories indefinitely: some do flunk out. Specifically, theories that remain tautological or yield predictions that are contradicted by the data must make way for theories that yield predictions for which the data are corroborative” (2000: 607).

Williamson has the final sentence just right. The problem is that after several decades of work justified by promise, we could be excused for suggesting that what we have here is the equivalent of an n-th year graduate student. If this criticism seems too strong, ask the following question: What have we learned from applying game theory over the past quarter century to issues of institutional origin, change, and stability that is not obvious or that we did not know already? The difficulty in answering this question is one reason thoughtful scholars have backed away from more extravagant claims, embracing less ambitious enterprises such as the construction of analytic narratives (e.g., Greif, 2006).

But we must be skeptical here as well. To what degree can narratives “informed” by game theory be told equally well without it? Here is how Geertz defined the enterprise of cultural anthropology:

Believing with Max Weber, that man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs, and the analysis of it to be therefore not an experimental science in search of law, but an interpretive one in search of meaning (Geertz, 1973: 5).

Economists have historically been unsympathetic to the work of sociologists or cultural anthropologists, particularly where they declaim scientific ambition. For many economists, who fancy themselves true scientists, the Geertz position is a cop out. If what we are in fact after is interpretation and the search for meaning, we are in the realm of the humanities. At their best, analytic narratives contain good narrative history, valuable in its own right but not made more “scientific” by adding on ex post “analysis” which often doesn’t satisfy Williamson’s test. If you are sympathetic to Geertz’s position, then you should be sympathetic to the aspirations of analytic narratives, and the converse holds as well.

Microeconomics and Foraging Algorithms

Why hasn’t game theory provided more insights into central questions of institutional stability and change? One way to understand microeconomic theory is that it is the study of foraging algorithms in settings never anticipated by the forces of natural selection that crafted them (Field, 2001, 2006a). Foraging algorithms have both cognitive and behavioral dimensions. They include intuitive capabilities for applying logic and making statistical inferences in pursuit of the satisfaction of drives that lead us to seek food when we are hungry or water when we are thirsty. They have been reinforced and refined by hundreds of millions of years of natural selection and are an indisputably important component of who we are.

The behavior studied in microeconomic theory is admittedly broader than that involved in an animal’s foraging for food because it can involve the pursuit of a more

extensive range of goals, and more sophisticated connections between sensory data and behavioral output (Sterelny, 2003). Both humans and cats are driven to eat when they are hungry, drink when they are thirsty, and have sex when exposed to certain stimuli. Some of our preferences reflect drives which are universal and indeed shared with other animals. But we are also capable of adopting goals or preferences from a set far broader than anything cats probably consider. In less complex animals, moreover, there is likely to be a fairly tight nexus between sensory data input and specific behaviors. Humans have more sophisticated capabilities to form representations of their environment and choose flexibly among behaviors.

That said, the foraging algorithms can, in principle, be used in pursuit of any goal. Whereas the range of objectives in whose service they have been applied may be broader for humans, the optimization under constraint such algorithms cause us to approximate is common in other organisms as well. Foraging algorithms refer to evolutionarily refined means of using externally acquired information to guide behavior in pursuit of objectives. These algorithms have an ancient lineage, are shared with other animals, and are the foundation for the processes modeled by microeconomic theorists.

In certain contexts, they may be all we need to build tolerably good models of human behavior. They are useful, for example, when explaining and predicting behavior in purely competitive markets. Foraging algorithms are useful as well where implicit prices are involved, such as when we predict that individuals, if they are in a hurry to make an appointment, will, among parking lots, choose the one that minimizes the distance between their vehicle and their destination.

Foraging algorithms are equally applicable in studying the behavior of hunter gatherer societies or other animals, involve cognitive capabilities optimized for determining the best patch in which to search, and how long one should search before moving to another. Since such algorithms assist organisms in obtaining food when they are hungry, those better endowed with them faced higher probabilities of survival and reproduction in the past; it is because of this that they have been selected for and refined over hundreds of millions of years of evolutionary history.

In all of the instances described above, however, we are in the realm of decision theory, in the sense that relevant information confronts individuals parametrically. Purely competitive markets organize exchange among individuals, but where numbers are large the price mechanism strips interaction of any apparent strategic element, since neither buyer nor seller can measurably influence price. Similarly, in the case of selecting a peripheral parking lot, the parameters influencing the decision problem are largely part of the non-human environment; in particular, one is not playing a game against the free spaces.

Foraging theory is useful in the study of exchange in thickly traded impersonal markets. It's far from perfect, due to our cognitive biases and limitations, but it's a decent starting point.³ Once we move beyond impersonal markets, however, we move into the realm of game theory. Game theory can be thought of as an attempt to migrate foraging theory to the realm of strategic interaction. In this domain, however, although foraging algorithms continue to play a role, they often take a back seat to the influence of

³ A number of behavioral economists and others (e.g., Taleb, 2005) have argued that foraging algorithms perform poorly in asset markets, in which of course it is also the case that prices confront individuals parametrically. But argument from cognitive limitations and biases is a different line of criticism from that developed in this paper.

other predispositions. Due to advances in instrumentation and imaging technology we are now able to confirm literally that parts of the brain other than the prefrontal cortex are involved when experimental subjects advance money in a game of trust, or reject a positive offer in an ultimatum game. Once one is in the realm of strategic interaction, there is no longer a good case that the foraging algorithms form a sufficient grounding for behavioral modeling.

The study of the microanalytics of institutions as well as the study of political behavior more generally, inevitably involves considerations of strategic interaction, which means that we are in a realm where foraging only models often cannot provide reliable predictions. The confounding influences include not just cognitive biases and limitations, about which a great deal has been written, starting with Simon (1955), but also different behavioral predispositions. We are in a domain that elicits the operation of powerful algorithms specialized for social interaction, and these sometimes trump the counsel of foraging algorithms.

Most social interactions, whether within or outside of an organized market, involve reciprocity. Behavioral and cognitive adaptations for social exchange must therefore be those which enabled and enable the establishment and persistence of reciprocal relations. Experimental methods provide robust evidence of predispositions to play cooperate in a one shot PD, make positive transfers in a game of trust, and reduce own welfare to punish those who have violated norms of reciprocity, particularly in small numbers interactions. A predisposition to venture trust – not unconditionally – but to venture it nonetheless – is absolutely necessary in order to build continuing relationships. These two predispositions: to trust sometimes even in the absence of a prior verified record of

trustworthiness, and to punish those who violate principles of reciprocity, are at the core of a grammar of human interaction out of which most social relations are constructed. The experimental evidence overwhelmingly documents these human propensities. If the rational choice approach is to have any claim as a scientific foundation for behavioral science, it has to make deconfirmable predictions, and when these predictions are deconfirmed, it matters.

Except in the most tautological sense, it is simply not rational to play cooperate in a one shot PD, whether one knows the likely play of one's counterpart or not. It is not rational to reject a positive offer in a one shot ultimatum game,⁴ just as it is not rational to seek revenge or to retaliate after a devastating nuclear first strike. Yet many of us, if placed in such circumstances, are predisposed to do so. The best metaphor, I think, is to say that we possess behavioral and cognitive modules that shut down, short circuit, or bypass the counsel of foraging algorithms. We now have evidence, for example, that oxytocin, a hormone released during sexual intercourse and breast feeding, also plays a role in other affiliative behavior, such as making a positive transfer in the first stage of a game of trust (Fehr et. al, 2005).

With respect to fear and anger, LeDoux (1996) has shown that when humans respond quickly in response to a threat, the higher mental processes (cognition) associated with the prefrontal cortex are bypassed by different neural pathways that run through the amygdala. The release of dopamine and epinephrine shuts down access to the prefrontal cortex. We know from neuroanatomical research that one can easily, and

⁴ A Nash equilibrium is a strategy profile such that, for each player, given what the other has chosen, neither could do better. In a typical ultimatum game, A has \$10 and can propose any whole dollar division to B. If B accepts, each gets what was proposed; otherwise neither receives anything. Since any outcome in the ultimatum game can be Nash, the concept of subgame perfect rationality represents an equilibrium refinement.

at will, enrage a cat by electrically stimulating its amygdala. These examples give concreteness to the metaphor of short circuiting. Social exchange requires behavioral predispositions that rely not simply on different or more complicated calculation. *They depend, in a very real sense, on shutting it down or bypassing it entirely.*

Focusing on reciprocity as a basic principle of social organization need not detract from exploration of issues of hierarchy and power. The topics are interrelated, since reciprocity is implicated in even highly asymmetrical relationships, as in “I won’t despoil your person if you give me your assets.” Although legal systems distinguish between contracts entered “voluntarily” and those agreed to under duress, both types and the entire continuum that spans them involve reciprocity. Reciprocity can also be generalized, as when we assist those who are helpless in anticipation that others would (and will) treat us with similar compassion were we so situated.

There is a lively and ongoing debate about the nature of political and social relations prior to the Neolithic revolution. Boehm (1999), for example, argues that pre-agricultural human groups were relatively egalitarian, with the power and influence of leaders easily subject to check by collective action. What Hayek (1988) described as our primitive social instincts, and what the experimental literature has documented in modern populations, can provide support for both democratic and dictatorial orders: neither is inherently more “natural” than the other.

Social contract theorists start with the premise that cooperative enterprise among humans depends on an overlay of institutions or culture that restrains more basic (natural) impulses that would otherwise cause us to tear each other to shreds. Traditional sociology, anthropology, and political theory are clear about this, and about the

proposition that these overlays are socially constructed. Hayek, in contrast, argued that there was no “problem of order“, or if there was the solution was to be found in the primitive social instincts which had a biological substrate and had been favored by group selection. For state of nature theorists (e.g., Hobbes), it is civilization, socially constructed, that provides the collective constraints on the dangerous individualism which is our natural heritage. For Hayek and those who follow him, the individualism that in his view gives rise to greater human liberty is the consequence of advance in civilization. Thus one might say that for Hayek, individualism is socially constructed.

On the question of why humans cooperate, economists have been all over the map, one reason why the microanalytics of institutions and culture are in such poor shape. Adam Smith can be read as foreshadowing the Hayekian position; he emphasized that we are endowed with moral sentiments such as sympathy that provide a structural template within which voluntary cooperation develops. But Smith can also be read as a proponent of selfishness-only assumptions about human motivation. By far the most common approach to institutional analysis today among economists is to try and build up human cooperation from the interactions of purely selfish agents.

Why any exchange takes place at all – simultaneous, over time, or generalized -- is a real problem for foraging only models. Defection in a one shot Prisoner’s Dilemma is a strictly dominant strategy, and the subgame perfect equilibrium in a game of trust is for no money to be advanced by the first player. Game theorists typically “solve” this problem by embedding exchange in an indefinitely repeated Prisoner’s Dilemma, and

then demonstrating that contingent cooperation could form the basis for a Nash equilibrium.⁵

It is important to understand how unsatisfactory this approach is and why. First, the existence of multiple equilibria in indefinitely repeated games precludes the possibility of deducing a one to one mapping from environmental variables to any particular equilibrium. In the indefinitely repeated PD, this therefore precludes the possibility of deducing any systemic tendency toward cooperation and social order among foraging only agents. A cooperative solution might emerge, but then again, it might not. That is all the theory can tell us. But this is not the only or even the most serious problem with the approach.

Friendships, coalitions, business relationships, and marriages all begin with an interaction that could well end after one play, and sometimes does. If foraging only parties play to an unsatisfactory conclusion in their first interaction, it is unlikely they will come back for more. Thus how people behave in a one shot game is absolutely critical in understanding how continuing relations develop. This point of view stands in sharp contrast to the dismissive comments about the empirical importance of such games common in the literature (see Binmore, 1994; Bendor and Swistak, 1997).

In the one shot Prisoners Dilemma, the only Nash equilibrium is the strategy profile where both players defect. Game theory makes a clear, unambiguous prediction about how wealth maximizing foragers will play, but upwards of forty percent of subjects in

⁵ The focus here is on game theory as used by social scientists within the rational choice tradition. Evolutionary game theoretic models, often explored using simulation, don't require calculating agents and are more ecumenical in considering other regarding as well as self regarding motivations. Although they don't start by assuming indefinitely repeated interaction among dyads, they do assume at the outset a demographic structure in which agents interact repeatedly with other agents in an essentially random pattern.

anonymous one shot experimental settings play cooperate. Of the four strategy profiles in the game, (defect/defect) is the only one which is inefficient. But since defect is a strictly dominant strategy, foraging only players are driven ineluctably toward it. It is the only Nash equilibrium in the game.

It has been widely acknowledged that as good as our intuitive statistics and logic might be, our cognitive abilities are limited. But it stretches credulity to argue that the problem here is one of bounded rationality. We are, to be sure, subject to cognitive biases in the way we transform sensory inputs into mental representations, and we do not have on demand access to the equivalent of Deep Blue in thinking about the best course of action. But putting so much attention on cognitive limitations leads to the mistaken conclusion that game theoretic models of strategic interaction would necessarily fare much better predictively if the cognitive side of human foraging algorithms were less deficient. Behavior in the one shot Prisoner's Dilemma provides a strong counter case.

Among economists there have been various responses to the experimental literature and the challenges it poses for common formulations of rational choice. One approach is reflected in the literature on social preferences, intent on accommodating the experimental findings by modifying the form and arguments of individual utility functions while preserving the essentials of a constrained maximization modeling of decision making (eg. Rabin, 1993).

It is indeed encouraging to see the experimental findings taken seriously. There is reason, however, to question whether constrained maximization remains a useful metaphor for understanding why a subject, for example, plays cooperate in a one shot PD. Samuelson (2005: 490) has argued that this behavior can be interpreted as evidence

of a preference for the more egalitarian outcome that would eventuate if both players cooperated. But could one not equally well interpret the play of cooperate as reflecting a death wish, or a masochistic desire to do much less well than one's counterparty? One can only view a play of cooperate as reflecting a social preference for equal outcomes if one imputes to one's counterparty a similar (and non-Nash) predisposition to play cooperate. Feelings of empathy may well incline us to do so. But one cannot call such behavior rational in any of the non-tautological senses of the term.

A number of authors have suggested that the emotional impulses that may underlay our propensity to cooperate contribute, paradoxically, to our rationality (Frank, 1988; Damasio, 2000; McDermott, 2004). It is true that in the one shot PD those who play cooperate may end up earning more money than those who play defect.⁶ But to suggest that this outcome per se makes the play rational reflects a poor understanding of the dilemma. This claim (a similar argument is made with respect to a propensity to reject positive offers in ultimatum games) appears paradoxical but is better described as incoherent.

There is strong pressure within much of modern social science to interpret behavior, particularly that which may make us better off, as rational. But if rationality is to mean more than the tautological statement that people act in satisfaction of their desires, one has to be able to imagine behavior inconsistent with it. In the narrower and more commonly understood sense of the word, *no amount of invisible hand waving can make it rational to play a strictly dominated strategy.*

⁶ For example, a group of players who cooperated and were always matched with each other would earn more money than a group who defected, but that wouldn't make it rational for any individual to cooperate.

The findings regarding behavior in one shot PDs are part of a large body of results including rejections of positive offers in ultimatum games, positive transfers in both directions in games of trust, positive allocations in multi player public goods games, and wealth reducing expenditures to inflict punishment in public goods experiments on those who have attempted to free ride. The behavior exhibited in these experiments goes to the heart of the empirical reasonableness of foraging only models. Together, they undermine the justification for identifying the indefinitely repeated PD game framework as a satisfactory framework for explaining the origin of cooperation. Because if the predictions of game theory are correct, and if we accept the historical necessity of building repeated play out of a successful initial interaction, no wealth maximizing human driven by the logic of the foraging algorithms alone would ever cooperate on the first play.

These conclusions pose a real threat to the claim that the canonical selfish actor rational choice approach theory can provide a comprehensive framework for understanding political, social, or strategic behavior. It is a threat not mistaken by anyone who thinks seriously about it. Two rather different lines of defense/counterattack have emerged. The first is best exemplified by Rubinstein, who has argued, quite disarmingly, that game theory simply should not be “a tool for predicting or describing real human behavior” (2001: 616). This means, if taken literally, that no accumulation of evidence on actual human behavior could or should threaten the preeminence of game theory as the foundational approach in the social and behavioral sciences. This position is not defensible if one is serious about using such theory as a foundation for scientific work.

One has to admire Rubinstein for his forthrightness. I suspect his view is shared privately by many theorists. It eliminates the apparent problem posed, for example, by the large gap between the Nash equilibrium prediction in the one shot PD game and actual behavior, or the indeterminacy associated with multiple equilibria in the indefinitely repeated game. For obvious reasons, his position is not generally publicly embraced by those with a more ambitious agenda for traditional approaches.

The second line of resistance attacks the interpretation and significance of the experimental results. One of the key attractions of the experimental method has been the promise that it can control for the effects of reputation and repetition. The attack says, essentially, that the controls fail, that the results reflect the inability of subjects to shed habits which serve them well outside of the experimental environment (Samuelson, 2005). Here Samuelson is reviewing Henrich et al (2004), and Samuelson is right in suggesting that much of the variation reported in that study probably reflected differential success in implementing controls while running the experiments in widely varying field conditions (Field, 2005).

But Samuelson's criticism is also implicitly directed towards the interpretation of the average levels of behavior which show up in these as well as the original experiments in which university students from developed countries were the main subjects. In none of those studies are the results consistent with the subgame perfect equilibrium prediction in the ultimatum game. The failed controls dismissal of experimental research is often advanced in an off hand matter, but the implications go to the heart of experimentalists' claim to have told us something about human nature. The argument, if sustained, trivializes the methodology and research program.

The claim that subjects are incapable of distinguishing between one shot and repeated play games is implausible. It has, moreover, been tested and found wanting (Falk, Fehr, and Fischbacher, 2003). And it runs some serious danger of blowback: If subjects' habits are so powerful that they produce nonrationlizable behavior in the controlled environment, that is likely to be the case as well in nonexperimental contexts. The dismissal is reminiscent of John Nash's response when told the results of the world's first Prisoner's Dilemma experiment, conducted in 1952. The experiment consisted of a finitely repeated PD game. In the experiment, the parties reached mutual cooperation on 60 of the 100 plays, and although it is true they both defected on the very last play, they jointly cooperated on five of the last ten plays, and the rate of cooperation rose as the game progressed.

These results did not accord with Nash's expectations (the unique Nash equilibrium is all defect/all defect). The behavior of the subjects disappointed him, and he commented that he "would have thought the players more rational." The subjects here were Armen Alchian of the UCLA economics department, and John Williams, chair of the RAND Corporation's math department, (Poundstone, 1992). Pioneers of behaviorist psychology had an often repeated principle for interpreting experimental results that conflicted with theory. Their principle was: "The rat is always right", by which they meant that when subject behavior varied from theoretical prediction, it meant that something was wrong with the theory, not the animal.

If one finds Rubinstein's position unsatisfactory, and is disinclined to reject the experimental results as the consequence of failed controls, how does one interpret them? The view advanced here is that they identify a set of behavioral adaptations for social

interaction favored in our evolutionary past by group selection. One doesn't have to accept the group selection argument to accept the first part of the statement, but it is not possible to provide a plausible account of the historical (evolutionary) origin of these predispositions without it (see D. S. Wilson and Sober, 1994; Sober and D. S. Wilson, 1998; Boehm, 1999; Field, 2001, 2006a,c; Sterelny, 2003; Wilson and Wilson, 2006). These behavioral adaptations for social exchange are paired with cognitive adaptation that make us good at judging trustworthiness through reading facial expressions and make us particularly attuned to and adept at detecting violators of social rules (Cosmides and Tooby, 1992).

Unfortunately, Cosmides and Tooby and other evolutionary psychologists have been unwilling to admit group selectionist explanations for the origin of the predispositions they have helped document. This may be puzzling to those coming to these debates from economics, since the few economists who have actually considered the matter, such as P. Samuelson, Becker, Hirshleifer, and Hayek have been sympathetic to the possibility, as, more recently, with some qualifications, has been Gintis (2000).

There are several reasons for this antipathy (for full discussion, see Field, 2007). One has to do with the legacy of the debates about the empirical significance of group selection that began with Williams (1966). Another has been confidence that evolutionary theorists have available adequate explanations for the existence of apparently altruistic behavior. Altruistic behavior among close relations, such as parents and children, is now uncontroversially explained as the consequence of Hamiltonian kin selection. E. O. Wilson, suggested in 1975 that altruism among non-kin remained the

“central problem” for evolutionary theory. In 1983 he believed it had been solved by Trivers, although by the late 1990s Wilson was not so sure (see Field, 2001: 109-110).⁷

Trivers’ 1971 article on reciprocal altruism continues to be widely cited, and it is important to understand why its argument is no more satisfactory than the indefinitely repeated PD explanation of cooperation. In fact, it is the same argument, with a small modification reflecting a move from a focus on agents’ conscious decisions towards a focus on the differential replication of genes predisposing towards certain types of behavior.

The essence of Trivers’ argument is this. It may appear altruistic for me to jump in a river and save you. But, said Trivers, if I can save you at a small risk to my life, thus conferring a big benefit on you and if, several months later, you can return the favor, humans with genes predisposing to such behavior would be more likely to pass on such genes to future generations.

Trivers’ model can be viewed as a game of trust. The subgame perfect equilibrium in this game is no transfer in the first stage (no initial rescue) and consequently no transfer in the second stage (no subsequent rescue). In a world where selection takes place at levels no higher than the organism one can only make it evolutionary “rational” for genes predisposing to rescue by making interaction indefinitely repeated. But positing that one is embedded in such interaction from the start solves the problem simply by assuming at the outset much of what needs to be explained.

The problem with the Trivers’ model is that my initial behavior reveals to you that you had some positive probability of being saved even though you had made no prior

⁷ By 2006 E.O. Wilson had joined with D. S. Wilson in calling for a reconsideration of the mainstream view emerging in the 1960s that the morphological and behavioral legacy of group selection in human populations was nil (Wilson and Wilson, 2006).

deposit to the favor bank. Upon initial appearance, genes favoring such rescues would be disfavored by natural selection operating at levels no higher than the individual organism. And, assuming the initial rescue did take place, the survival maximizing strategy four months hence is for you to be deeply apologetic as you explain to me that you would like to help but can't risk the possibility you might drown.

In evolutionary theory the Trivers model is every bit as inadequate as the indefinitely repeated game theoretic explanation of cooperation. Nevertheless, both seem to be surviving just fine, judging from recent textbook treatments and survey articles. I suspect, however, that we will look back some years from now and conclude that it was in part inertia and intellectual lassitude that allowed these close relatives to continue for so long to enjoy such influence.

The behavioral and morphological legacy of group selection can usefully be thought of as the consequence of its operation over two historical epochs. Much modern discussion has focused on the environment of evolutionary adaptation (EEA), roughly the last two million years. But evolutionary biologists are increasingly receptive to the idea that during the long period prior to the EEA biological group selection (it couldn't have been cultural) played a key role in such transitions as the origin of chromosomes, the shift from prokaryotes to eukaryotes (cellular organisms) and subsequently to multicellular organisms, the origin of sexual reproduction (Maynard Smith and Szathmary 1995; Keller 1999), and selection for restraint on intraspecific harm. Violations of the last principle are the exception rather than the rule in the animal kingdom: most of the time animals don't eat or kill members of their own species. Thus the behavioral and morphological characteristics possessed by the common ancestor of the three surviving

chimpanzee species six million years ago, as well as our ancestors two million years ago, already reflected important legacies of the operation of biological group selection.

In the EEA, cognitive and other adaptations allowed for an explosion of affirmative cooperation, moving us beyond the passive altruism of failure to harm, and making us the most cooperative animal species with the possible exception of the social insects. Much attention has focused on increased cranial capacity in humans and the expanded cognitive capabilities this allowed. Engineering constraints associated with bipedalism limited how much the female birth canal and hips could widen, which forced humans to give birth to dependent infants with a consequent period of post natal dependence much longer than that observed in chimpanzees, and extended periods of close contact between infants and experienced adult(s). The result was a more favorable venue for social learning and imitation, a joint product of improved cognitive capabilities and an extended period of postnatal dependence.

The emergence of culture as a strong influence on behavior, particularly in the last half million years (Boyd and Richerson, 2004), meant that there could now be strong interdemetic (between group) as well as intrademetic (within group) selection. Learned cultural behaviors allowed boundaries of groups to be less permeable, thus allowing selection to operate more effectively via the differential survival and extinction of groups. With the emergence of culture the old and much criticized Wright (1945) model of group selection, with its focus on interdemetic selection, has been in part rehabilitated, in the sense that it may be applicable under conditions in which culture has emerged as a strong influence on behavior, even if, as most now agree, it provides an inadequate foundation for biological group selection prior to the EEA.

Starting with Williams (1966), a main attack on the likelihood of empirically important group selection emphasized the implausible demographic conditions required for interdemic selection. In particular, permeable group boundaries could, it appeared, easily defeat it. The modern restatement (see Wilson and Sober, 2004; Sober and Wilson 1998), emphasizes intrademic selection, in which groups divide into subpopulations for a period of time and then periodically merge or recombine. Periodic permeability of group boundaries has become a necessary feature of models of group selection, as opposed to their apparent Achilles heel.

Acknowledging the prior operation of group selection provides an evolutionarily coherent explanation of the survival of human behavioral predispositions that enable social cooperation. This is not to deny the importance of the foraging algorithms. It is simply to say that if we want to develop a systematic approach to institutions and culture, we have to go beyond them. In addition to our drives to eat, drink, and procreate, and the foraging algorithms that help us pursue them, we also possess cognitive machinery and behaviors specialized to social interaction, predispositions that incline us to trust once (refrain from first strike on conspecifics) and to retaliate against those who violate our trust. These predispositions are the legacy of group level selection, and underlie what Hayek referred to as our primitive social instincts. They are the bedrock upon which we build institutional and cultural variation, and help explain universal features of human culture.

Because the use of game theory typically entails a foraging only mind set, it leads to the confusion of deductively correct conclusions with those that have predictive and explanatory power in understanding human behavior. When applied to realms where

other algorithms have important influences on behavior, an approach based on foraging only assumptions often produces poor predictions. Greif (2005), for example, defends a role for game theory in comparative historical analysis. He acknowledges the problem of equilibrium selection, and grants that we cannot expect to achieve a “deductive” theory in which there will be a one to one mapping from environmental variables to normative structures. He even acknowledges how unsatisfactory it is to try to account for third party punishment by appeal to those who will punish those who don’t punish, and so on, in a recursive enterprise that merely kicks the Prisoners Dilemma to a higher level but does not resolve it.

Yet, surprisingly, he endorses the traditional sociological and anthropological view that “preferences and rationality are socially constructed,” a view that he suggests “is now common in economics” (2006: 22). I appreciate the considerations that have led Greif to go this route, but doubt a) that the field has in fact become quite so postmodern and b) that it is the right route to go. Certainly, much of the variable component of human culture is socially constructed, in the sense that it is the result of imitation and social learning. But not all: there is a biologically influenced foundation, a legacy of group selection, which accounts for universalities in human culture and is the bedrock upon which cultural variability is built.

Note that I have assumed a rather mild interpretation of social constructivism here, one roughly consonant with the blank slate assumption. There are far more radical variants which aggressively reject the idea that there is an external objective (“ontological”) reality and adopt what can really only be interpreted as an anti-scientific stance. I don’t think Greif wishes to endorse these views.

The prospect of building bridges between the social and natural sciences is one that terrifies many sociologists and anthropologists as well as philosophers and others in the humanities, and one which the philosophy of social constructivism is intended in part to forestall. The problem with social constructivism is not the claim that experience and interactions with environment influence aspects of human motivation and cognition. Of course they do. The problem is the implicit and sometimes explicit claim that genetics and biology play no role in this process (see Pinker, 2002).

If one tries *ex ante* to use game theory to account for cooperation and social order, one ends up again and again either with indeterminate predictions (because of the existence of multiple equilibria) or determinate predictions which are simply contravened by observed behavior (as in the one shot PD game). If one uses game theory *ex post*, it is sometimes possible to find a model based on foraging only assumptions that appears to be consistent with the observed behavior. But the sense of explanatory satisfaction one gets with such a model is too often a reflection of hindsight bias.

People will continue to generate game theoretic models, because the activity can be intellectually challenging and fun. But although the results carry an aura of scientism about them, they are often scientifically barren, in the sense that they do not produce models with robust, unambiguous, and correct predictions. That is why Rubinstein wishes to absolve the enterprise of any responsibility to do so. Often the models are not falsifiable, or in the cases where they are falsifiable, they are in fact falsified.

To the degree that we desire a better microanalytics of institutional formation and stability, we are going to have to move beyond foraging. What we need to do is to better circumscribe the domains within which foraging only assumptions do pretty well, and

constrained maximization models provide reasonable metaphors for what is going on, and other domains where other, often non-calculative algorithms specialized for social interaction kick in.

Beyond Foraging

Amidst all the uncertainty, there is one proposition around which there is almost universal agreement: we lack a satisfactory framework for understanding institutional change and stability. The question is, how can we best go about developing one? If we are to build brick by brick, as Williamson suggests, we need to know where good structural materials are being generated. The restricted range of behavioral predispositions typically entailed by the rational choice game theoretic approach is too limited to provide the foundations for a broader theory.

The initial step is to acknowledge that the wide body of experimental results (Camerer, 2003) is attributable to something other than simple failure of the controls: that it reveals and documents the operation of a set of species typical behavioral predispositions at variance with the foraging only assumptions that typically underlie rational choice/game theoretic models. Humans possess a set of behavioral and cognitive adaptations for social exchange, premised fundamentally on reciprocity. They include propensities to cooperate in individual or multiparty prisoner's dilemma when the strictly dominant strategy is to defect. They include cognitive adaptations to detect cheaters, paired with behavioral propensities to punish those who violate principles of reciprocity, including third party violators. Because genetic predisposers to these behaviors could not upon initial appearance have been favored by organism level selection, higher level

selection must have been involved in their evolution to higher frequency (Wilson and Sober, 1994; Field 2001; 2006a, c; Wilson and Wilson, 2006).

The second step is to move beyond social constructivism. Suppose we consider a moderate blank slate version that doesn't reject the idea of objective reality. Pinker has been a forceful critic of this approach, and there is increasing agreement among cognitive scientists, developmental psychologists, and linguistics scholars that our brains arrive from the factory with some preformatting. We are differentially prepared to learn in certain directions, and to classify sense impressions into preencoded conceptual categories (Atran 1990; Boyer, 2001; Pinker 2002).

Moral philosophers are increasingly sympathetic to the view that there is a universal normative grammar, with innate contextually dependent human predispositions against killing other human beings, lying to them, or cheating them (Harman, 1998). The contextual dependence is that humans typically view these rules as applicable only within their own group. Experimental research shows that ingroup/outgroup dichotomization is quite plastic (Tajfel and Billig, 1974), and as a practical matter it can make a tremendous difference how one is classified by those with whom one is interacting.

Understanding what underlies that plasticity is critical in developing a positive theory of politics. Exploring the role played by propaganda, ideology, and other belief systems, and how these build upon shared cognitive and behavioral predispositions, is essential. But in studying variability it is important not to lose sight of the widespread prevalence of something akin to the golden rule in moral and ethical systems. This near universality is suggestive of hard wired adaptations favoring reciprocity rather than a simple coincidence of purely cultural achievements in restraining baser, more "natural"

impulses. Our behavioral and cognitive adaptations favoring reciprocity have as much to do with social construction as our predispositions against sexual relations with those with whom we grew up in close proximity. In other words, less than we might presume.

North defined institutions as “a set of rules, compliance procedures, and moral and ethical behavioral norms designed to *constrain* the behavior of individuals in the interests of maximizing the wealth or utility of principals” (1981: 201-202; my italics). The emphasis on constraints presupposes that the basic, universal, species typical behavioral predispositions of humans are foraging only. If we take a different starting point, however, we begin to see institutions as reflections of permutations of a limited range of genetically encoded, environmentally mediated cognitive and behavioral predispositions.

The view that there is a fundamental societal “problem of order”, as Parsons (1937) put it, a problem that culture and institutions “solve”, is a recurring theme within literature, philosophy, and social science (Huxley, 1902; Freud, 1962; Golding, 1962). But perhaps the problem of social order is something that we have, on our own, constructed. There is no fundamental problem if normative structures are built upon individual cognitive and behavioral predispositions, if the affiliative and retributive predispositions that lead to otherwise puzzling or anomalous experimental results have a genetic/biological substrate. On this interpretation, the broad range of cultural universals, often overlooked by a focus on what makes us different, is not to be understood as multiple independent solutions to similar environmental challenges, but rather a reflection of a shared genetic architecture.

On this reading, social and normative structures do not arise from the decisions of freely interacting foraging only agents. Nor are they simply a reflection of a

superorganismic deus ex machina, whether we call it culture, social structure, or institutions. They are built upon a universal human psychology that includes not only foraging algorithms, but behavioral and cognitive adaptations for social interaction that may short circuit the counsel of foraging algorithms.

Empirical evidence relevant to these claims can be found in the experimental studies of behavior already referenced. Reinforcement comes from neuroanatomical and neurobiological research that reveals more about mechanism. A rethinking of evolutionary theory now provides a plausible account of origin. And corroborative evidence continues to emerge from other sources, in particular the field of behavioral genetics.

It has been a continuing puzzle for foraging only theorists, for example, why most people bother to engage in anything but small group political activity, develop or research positions on issues of national policy, or vote. The logic of free riding seems inexorably to restrict effort to private advancement, rather than influencing public policy. Scholars influenced by the selfish actor rational choice paradigm have, however, been looking in the wrong places. The explanation is to be found elsewhere, in a set of cognitive and behavioral inclinations – largely those implicated in social interaction -- that predispose us to do so. It has been evident for some time that a wide range of personality and other traits are heritable in human populations. Recent work shows that heritability extends down to the level of political attitudes. Systematic evidence that a tendency towards political and social conservatism is partially genetic (Alford and Hibbing, 2004: 714) is a remarkable finding, and it is important to understand upon what it is based.

The gold standard for measuring heritability are twin studies. The standard index of heritability is twice the difference between the trait correlation for monozygotic (MZ) and dizygotic (DZ) twins. If heredity played no role in determining individual variation, the correlations for the two types of twins should be about the same, yielding a heritability index of 0. If twin similarity were entirely the result of genetic similarity, one would have a correlation of 1 for the MZ twins and .5 for the DZ twins, yielding a heritability index of 1. These measures refer to the share of individual variability that can be attributed to genetic factors.

The shared environmental influence for twins raised together is calculated by doubling the correlation for DZ twins and subtracting from this the correlation for MZ twins. For a purely hereditary trait, the trait correlation should be .5 for DZ and 1 for MZ, leading to a shared environment contribution of 0. If there's no genetic influence, the correlations should be the same (say .3), which is also the shared environmental effect ($.6 - .3 = .3$). Whatever remains after subtracting the effects of heredity and shared environment represents the role of unshared environment.

Such a calculation is without controversy when applied to height, which has a heritability index of about .8 in modern populations. But it runs against our intuitions that MZ twins should have correlations in their political attitudes so much higher than those observed for DZ twins. But they do. In the Wilson-Patterson inventory, subjects are given short prompts, such as Socialism or Death Penalty, and asked whether they agree, disagree or are uncertain about them. The responses of MZ twins are systematically more similar than those of DZ twins. With respect to School Prayer, for example, the trait correlation in a recent study was .66 for MZ twins and .46 for DZ

twins, leading to a heritability index of .41. Heritability for other items was somewhat lower, but the trait correlation for every single one of 28 items examined was higher for the MZ than the DZ twins (Alford, Funk, and Hibbing, 2005).

The samples in this study were large, DZ pairs were limited to single sex to remove any confounding effects of gender, and these differences hold up whether or not twins were raised in the same family, and whether or not families mischaracterized their offspring as MZ or DZ. In contrast, actual political party affiliation is much less heritable.

The main implication I wish to draw from these twin studies requires understanding a subtle point. There are two cases that will yield a heritability index of 0. The first, as discussed above, is if heredity has nothing to do with the trait. The second is if the trait is genetically determined, but there are no relevant genetic polymorphisms. If one records the number of eyes possessed by MZ and DZ twins, one will find a trait correlation of close to 1 for both groups, and a heritability index of 0. The genetic predisposers to having two eyes have evolved to fixation. Although eye number is inherited, it is not heritable.

There is, in contrast, only one interpretation of a positive (and statistically significant) heritability index, and that is that genes are in fact implicated in transmitting the trait from generation to generation. A positive heritability coefficient is therefore sufficient although not necessary to implicate heredity in shaping an aspect of behavior, cognition, or morphology. If it is true that individual differences in political and social attitudes are partly the result of genetic variation, it must perforce follow that the central tendencies in these attitudinal variables have a genetic underpinning as well. These

recent studies show that heredity plays a role in species typical political attitudes, and consequently, species typical behavior in this realm.. They are important evidence in favor of the view that our genetic architecture limits the scope for political variation just, as Chomsky's work suggests, it constrains the grammatical structures that human languages can employ, and as Boyer (2001) has demonstrated, it provides limits on the range of successful supernatural ideas. We are born with broad templates that influence what our social and political relations look like. We are collectively free to choose (or learn or imitate), but the choice set from which we select is not unbounded.

Conclusion

Most serious scholars now acknowledge the reality of the central institutional fact. A corollary is that there can be no presumption that Pareto improving institutional structures will necessarily take hold, or that at any moment of time, institutions and culture will be "efficient" given prevailing technologies and geographical/climactic conditions. The explanation given is typically political and ideological: elites, classes or interest groups may successfully fight to establish or maintain structures that benefit a small group even if the societal pie is smaller than it otherwise might be. It is partly that transitions always involve winners and losers, and there is no easy way to guarantee that Kaldor Hicks compensation will in fact be paid.⁸ But it is not just about cash: Ideological factors play an important role in the success or failure of political efforts.

If political beliefs and ideology and the forces that produce them are a key part of explaining what is responsible for the central institutional fact, we must unavoidably incorporate them into institutional analysis. Central to making progress here will be a

⁸ In welfare economics, a Pareto improvement makes someone better off and none worse off. A Kaldor-Hicks improvement allows for losers, provided the winners could compensate them and still be better off.

better understanding of the ways in which political rhetoric successfully exploits cognitive and behavioral predispositions that shut down or bypass the calculative modes essential for foraging, understanding as we do this that these dispositions lie at the foundation of all social exchange and social order. They are central to who we are, and cannot simply be dismissed, decried, or ignored. By political configuration I mean, broadly, the overlapping patterns of ingroup/outgroup delineations, and where groups collectively position themselves on the authority/democracy dimension. It is possible to gain insights by treating these configurations as largely exogenous using techniques of narrative history, or cross country regressions using instrumented variables. But if we wish to move forward on the microanalytics, there is no avoiding the development of a behavioral science that explores what lies behind them. Unfortunately, one can expect only limited help from political science. The selfish actor rational choice approach is at least as firmly rooted in this discipline as it is in economics and political scientists are having almost as much fun with indefinitely repeated games as are economists.

In the last part of this essay I made the case that universal features of institutions and culture have their roots in species typical behavioral and cognitive predispositions with genetic/biological substrates. The challenge is to understand how cultural variability in its broadest sense reflects the playing out of these predispositions within the context of coalitional and group plasticity. In terms of both uniformities and variation in institutions and culture, social learning plays an important role. Humans learn from each other through imitation and direct instruction. We absorb factual knowledge about the world, through this and other interactions with our environment. But what is remarkable from a

foraging only perspective is how easily we are prepared to absorb deontological instruction.

Members of an infantry squad, for example, can be and are taught that if a grenade rolls into a foxhole, and if there is no time to eject it, each member of the squad should scramble to cover it with his body. Members of squads which have fully absorbed this teaching will, on average, have higher survival rates than members of squads where the teaching has not taken hold. But the behavior is clearly a dominated strategy, and the fact that we are ready to absorb such instruction so relatively easily is a puzzle if we think of humans as foraging only agents. For such an agent deontological instruction would have no effect on behavior, and if the world were foraging only, there would be little point in providing it. Both logic and statistics would lead targets to the conclusion that adopting the advocated behavior would be a bad idea. Logically, because the behavior can be seen as a dominated strategy; statistically because empirical observation would show that those who won the race ended up dead.

As Wilson and Wilson (2006) argue, our receptivity to such instruction can be understood as a mechanism that suppresses selection within groups; that receptivity itself must have evolved as a consequence of group level selection. The content of deontological teaching can vary, as can learned definitions of in group/out group boundaries, which helps explain what lies behind the central empirical problem. Those attempting to enable or prevent large scale social or political change have always understand the importance of controlling the provision and content of such instruction. If their efforts have met with mixed success, it is because it is not always easy to do so. As we build on Kuran's (1995) pioneering exploration of factors leading to the collapse

of institutions, we will need to focus more attention on how such pathways can be and are disrupted.

Nothing in this essay is intended to dispute the role played by foraging algorithms in political, social, and economic behavior. There is a reason we legislate conflict of interest rules, a reason why bribery is such a persistent corrosive influence in political systems. In many domains, material incentives matter a great deal. Humans tend to choose more over less, and life over death, particularly when optimizing against non-human aspects of the natural world. But the foraging algorithms provide an insufficient foundation upon which to build a comprehensive behavioral science. Experimental methods are critical in delineating the nature and strength of other predispositions enabling social exchange, and what contexts elicit them. Research in neuroanatomy and neurobiology, and behavioral and molecular genetics, fields all currently benefiting from breakthroughs in instrumentation and imaging technology, complement this work..

The advantages of basing inquiry upon systematic analysis of data cannot be too strongly stressed. Hypotheses can be tested and empirical findings replicated, which allows an empirically progressive enterprise in which findings can build upon what has gone before. The fixed costs of contributing in these areas is, however, often high, and the enterprise can benefit from intellectual specialization. In particular, the task of interpreting, evaluating, and integrating findings from these multiple research programs is too large and too important to be left only to those producing the individual building blocks. There is room, and indeed need for, methodologically sophisticated contributions from throughout the social and natural sciences, and perhaps the humanities as well. We all – including those making the bricks - have a stake in the outcome.

BIBLIOGRAPHY

- Acemoglu, Daren, Simon Johnson, and James Robinson (2001), 'The Colonial Origins of Comparative Development', *American Economic Review*, **91**, December: 1369-1401.
- Alford, John and John Hibbing (2004), 'The Origin of Politics: An Evolutionary Theory of Political Behavior', *Perspectives on Politics*, **2**, December: 707-23.
- Alford, John, Carolyn Funk and John Hibbing (2005), 'Are Political Orientations Genetically Transmitted?' *American Political Science Review*, **99**, May: 154-168.
- Atran, Scott (1990), *Cognitive Foundations of Natural History*, Cambridge: Cambridge University Press.
- Bendor, Jonathan and Piotr Swistak (1997), 'The Evolutionary Stability of Cooperation', *American Political Science Review*, **91**: 290-307.
- Berg, J. J. Dickhaut and K. McCabe (1995), 'Trust, Reciprocity, and Social History', *Games and Economic Behavior*, **10**: 122-42.
- Binmore, Ken (1994), *Game Theory and the Social Contract I: Playing Fair*, Cambridge: MIT Press.
- Boehm, Christopher (1999), *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*, Cambridge: Harvard University Press.
- Bouchard, Jr. T. J., M. McGue (2003), 'Genetic and Environmental Influences on Human Psychological Differences', *Journal of Neurobiology*, **54**: 4-45.
- Boyd, Robert and Peter Richerson (2004), *Not by Genes Alone: How Culture Transformed Human Evolution*, Chicago: University of Chicago Press.

- Boyer, Pascal (2001), *Religion Explained*, Berkeley and Los Angeles: University of California Press.
- Camerer, Colin (2003), *Behavioral Game Theory*, Princeton: Princeton University Press.
- Cosmides, Leda and John Tooby (1992), 'Cognitive Adaptations for Social Exchange', in *Barkow et. al* (1992: 163-228).
- Dennett, Daniel (2006), *Breaking the Spell: Religion as a Natural Phenomenon*, New York: Viking.
- Falk, Armin, Ernst Fehr, And Urs Fischbacher (2003), 'Reasons for Conflict: Lessons from Bargaining Experiments', *Journal of Institutional and Theoretical Economics*, 159, 171–187.
- Fehr, Ernst, Urs Fischbacher and Michael Kosfeld (2005), 'Neuroeconomic Foundations of Trust and Social Preferences', *American Economic Review* **95**, May: 346-51.
- Field, Alexander J. (1981), 'The Problem with the New Institutional Economics: A Critique with Special Reference to the North-Thomas Model of pre-1500 Europe', *Explorations in Economic History* **18**, April: 174-98.
- Field, Alexander J. (1984), 'Microeconomics, Norms, and Rationality', *Economic Development and Cultural Change*, **32**, July: 683-711.
- Field, Alexander J. (1991). 'Do Legal Systems Matter?' *Explorations in Economic History* **28**, January: 1-35.
- Field, Alexander J. (2001), *Altruistically Inclined? The Behavioral Science, Evolutionary Theory, and the Origins of Reciprocity*, Ann Arbor: University of Michigan Press.

- Field, Alexander J. (2003), 'The Most Technologically Progressive Decade of the Century', *American Economic Review*, **93**, September: 1399-1414.
- Field, Alexander J. (2005), 'Review article on *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small Scale Societies*', *Quarterly Review of Biology*, **80**, December: 453-459. .
- Field, Alexander J. (2006a), 'Group Selection and Behavioral Economics.' In *Foundations and Extensions of Behavioral Economics: A Handbook*, ed. Morris Altman, New York: M. E. Sharpe: 165-82.
- Field, Alexander J. (2006b), 'Technological Change and U.S. Economic Growth during the Interwar Years', *Journal of Economic History*, **66**, March: 203-36.
- Field, Alexander J. (2007). 'Why Multilevel Selection Matters', *Journal of Bioeconomics*, **9**: (forthcoming).
- Freud, Sigmund (1962), *Civilization and its Discontents*, New York: W.W. Norton.
- Geertz, Clifford (1973), *The Interpretation of Cultures: Selected Essays*, New York: Basic Books.
- Gintis, Herb (2000), *Game Theory Evolving*, Princeton: Princeton University Press.
- Glaeser, Edward L., Rafael de la Porta, Florencio Lopez de Silanes and Andrei Shleifer (2005), 'Do Institutions Cause Growth?' NBER Working Paper.
- Glaeser, Edward L., David I. Laibson, Jose A. Scheinkman and Christine L. Soutter (2000), 'Measuring Trust', *The Quarterly Journal of Economics*, **115**, August: 811-846.
- Golding, William (1962), *Lord of the Flies*, New York: Coward-McCann.

- Greif, Avner (2006), *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade*, Cambridge: Cambridge University Press.
- Hall, Robert E. and Charles Jones (1999), 'Why do Some Countries Produce So Much More Output per Worker than Others?' *Quarterly Journal of Economics*, **114**, February: 83-116.
- Harman, Gilbert (1998), 'Moral Philosophy and Linguistics.' In J. Fodor, J. Keyser, and A. Brand (eds.), *Celebration: An Electronic Festschrift in Honor of Noam Chomsky*, Cambridge, Mass.: MIT Press.
- Hayek, Friedrich (1988), *The Fatal Conceit: The Errors of Socialism*, Chicago: University of Chicago Press.
- Huxley, Thomas (1902), *Evolution and Ethics, and other Essays*, New York; D. Appleton and Co.
- Kuran, Timur (1995), *Private Truths, Public Lies*, Cambridge: Harvard University Press.
- LeDoux, Joseph (1996), *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, New York: Simon and Schuster.
- Mathews, R.C.O. (1986), 'The Economics of Institutions and the Sources of Economic Growth', *Economic Journal*, **96**: 903-18.
- McDermott, Rose (2004), 'The Feeling of Rationality: The Meaning of Neuroscience for Political Science', *Perspectives on Politics*, **2**, December: 691-706.
- North, Douglass C. (1981), *Structure and Change in Economic History*, New York: Norton & Co.
- Parsons, Talcott (1937), *The Structure of Social Action*. New York: McGraw Hill.

- Pinker, Steven (2002), *How the Mind Works: The Modern Denial of Human Nature*, New York: Viking.
- Poundstone, William (1992), *Prisoner's Dilemma*, New York: Doubleday.
- Rabin, Mathew (1993), 'Incorporating Fairness into Game Theory and Economics', *American Economic Review*, **83**, December: 1281-1302.
- Rubinstein, Ariel (2001), 'A Theorist's View of Experiments', *European Economic Review*, **45**, May: 615-628.
- Samuelson, Larry (2005), 'Foundations of Human Sociality: A Review Essay', *Journal of Economic Literature*, **43**, June: 488-97.
- Simon, Herbert (1955), 'A Behavioral Model of Rational Choice', *Quarterly Journal of Economics*, **69**: 99-118.
- Sober, Elliott and David Sloan Wilson (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge: Harvard University Press.
- Sterelny, Kim (2003), *Thought in a Hostile World: The Evolution of Human Cognition*, Oxford: Blackwell.
- Tajfel, Henri and Michael Billig (1974), 'Familiarity and Categorization in Intergroup Behavior', *Journal of Experimental Social Psychology*, **10**: 159-70.
- Taleb, Nassim (2004), *Foiled by Randomness: The Hidden Role of Change in Life and in the Markets*, New York: Random House.
- Trivers, Robert (1971), 'The Evolution of Reciprocal Altruism', *Quarterly Review of Biology*, **46**: 35-57.
- Williams, George C. (1966), *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*, Princeton: Princeton University Press.

Williamson, Oliver (2000), 'The New Institutional Economics: Taking Stock, Looking Ahead', *Journal of Economic Literature*, **38**: 595-613.

Wilson, David Sloan (2006), 'Human groups as adaptive units: toward a permanent consensus', in *The Innate Mind: Culture and Cognition*, P. Carruthers, S. Laurence and S. Stich, eds., Oxford: Oxford University Press.

Wilson, David Sloan and Elliott Sober (1994), 'Reintroducing Group Selection to the Human Behavioral Sciences', *Behavioral and Brain Sciences*, **17**: 585-654.

Wilson, David Sloan and Edward O. Wilson (2006), 'Rethinking the Theoretical Foundation of Sociobiology,' Paper submitted to *Science*.