Introduction

Rules alone can unite an extended order. . . . Neither all ends pursued, nor all means used, are known or need be known to anybody, in order for them to be taken account of within a spontaneous order. Such an order forms of itself. . . .

Hayek (1988, pp. 19–20)

. . . the realist . . . turns his back on the whole he cannot grasp and busies himself with a fragment.

Gibran (1918; 2002, p. 55)

Experimental economics is good at measurement, testing, and discovery in studying the microeconomics of human behavior governed by the informal norms of social exchange and the more explicit rules of exchange in institutions. It has not been good at integration and interpretation within the broader context of human social and economic development. The learning from a half-century of experimental discovery will be particularly significant if we can find a way to leverage that learning into a broader understanding of the human career; otherwise, the rewards from the range of our research will be too narrowly drawn, fragmented, and of passing interest, as scholars move on to the intricate details of whatever is next. This book is an outgrowth of my struggle to obtain a larger vision of meaning in social and market economic behavior, and to communicate whatever value that process might contribute to a larger community. I know that others have similar concerns because we have shared them from time to time in passing and in depth. The picture I see is still blurred. Its outlines, however, are unmistakable; it remains for others to sharpen or change that picture even if most just pursue their business in their own way without it.

If we are to confront the challenge of meaning, we must begin by recognizing that the phenomena that underlie our subject matter arise from the remarkable capacity of human sociality and culture to discover forms

of interaction and organization that have enabled impressive expansions in human betterment. The situations we model and study emerged naturally from individual interactions, associations, businesses, and collectives. The agents active in this process were naïve in economic understanding, but had deep personal experiential knowledge that served them well (Polanyi, 1962, 1969).

These considerations have heightened my interest in F. A. Hayek's important distinction between two kinds of rationality. I shall try to relate all of this book's discussion and examples – experimental, field empirical, descriptive – to the following two concepts of rationality:

Constructivist rationality, applied to individuals or organizations, involves the deliberate use of reason to analyze and prescribe actions judged to be better than alternative feasible actions that might be chosen. When applied to institutions, constructivism involves the deliberate design of rule systems to achieve desirable performance. The latter include the "optimal design" of institutions, where the intention is to provide incentives for agents to choose better actions than would result from alternative arrangements.

Ecological rationality refers to emergent order in the form of the practices, norms, and evolving institutional rules governing action by individuals that are part of our cultural and biological heritage and are created by human interactions, but not by conscious human design.

The two concepts are not inherently in opposition; the issues are emphatically *not* about constructivist *versus* ecological rationality, as some might infer or prefer, and in fact the two can and do work together. For example, in evolutionary processes, constructivist cultural innovations can provide variations while ecological fitness processes do the work of selection. We will encounter many examples in which the two kinds of rationality coincide, and others in which they diverge or at least are still seeking convergence.

To illustrate, people were specializing through trade in markets with asymmetric information before the agricultural revolution. Where the problem was not too intractable, our forebears long ago also discovered and solved some common problems and found private arrangements enabling needed public goods to be built. They overcame defection incentives to cooperate effectively, developed effective auction systems before the Christian epoch, and in time extended them to selling everything from art to securities. All these remarkable developments occurred in the midst of negative reciprocity, inhumane forms of punishment and violence, and persistently sharp in-group versus out-group differentiation in moral practices. Although as economists we have articulated rational models of public goods problems,

such ways of thinking were not necessary in the past for societies to create emergent solutions out of human interactions uninformed by formal economic analysis. Similarly, in hundreds of market experiments, economically unsophisticated and naïve but proficient individuals produce rational outcomes without in fact having any knowledge of the rationality and efficiency of the outcomes they produce. Their effectiveness is perhaps less surprising once we recognize that their human forbearers and contemporaries used their cultural and biological inheritance to create the institutional forms that we study in the experiments, but our neoclassical models (since the 1870s) failed to anticipate or even to appreciate this important development as we proceeded to construct the concept of an "institution free core" of economic analysis.

As theorists, our first cut at constructivist problem definition and "solution" leads quite appropriately to concerns about incentive failure, but abstract approaches to incentives may omit significant features. Consider the problem of public goods provision. Initially we thought and taught that public goods could not be produced efficiently by private means. Yet the canonical example, the lighthouse, emitting signals that all ships could observe at zero marginal cost, was privately financed before economics had become a well-defined profession. The problem of supplying incentives for private investments and aborting free riders was solved practically by lighthouse owners who contracted with port authorities to charge docking ships for lighthouse services (Coase, 1974). These contracts allowed the capital cost of lighthouses (a discrete variable cost before it is incurred) to be prorated among ship dockings, since dockings provided an effective and practical measure of lighthouse service utilization and value in consumption. For "efficiency," it is argued, the so-called "fixed" cost, once incurred, should not affect the price of lighthouse services. However, this argument is a fallacious nonstarter because it omits the inefficiency that results if the lighthouse is not built.

And the famous "tragedy of the commons" in grazing cattle was decidedly not necessarily a tragedy for the high alpine Swiss cheese makers who for each summer at least since 1224 A.D. pastured their cows on the commons. Entry to summer pastures was controlled by a property right rule that "no citizen could send more cows to the alp than he could feed during the winter" (Netting 1976, p. 139).[1] These economic design problems were solved by

[1] In the solution that Coase found for the lighthouse, note that one might paraphrase Netting that no shipping company could pass more of its ships past the lighthouse than it paid for as part of the ship docking charges.

people completely unschooled in free-rider theory, but experienced enough in their behavioral coordination problem to seek solutions that might work. Somehow they perfected them by trial-and-error "natural experiments" over time.

Constructivist analysis enables us to see that these were examples of excludable public goods, and in all such cases the question is whether there are feasible ways of limiting use to avoid or internalize external costs, or of assuring payments that cover investment cost. Not every such institutional design problem has a solution that people are able to fashion out of their experience. Ostrom (1990) examines a variety of different common property resource problems around the world and the emergent self-governing institutions that solved or failed to solve the governance issues that they addressed.[2] The solutions, as in the preceding examples by Netting and Coase, are often ingenious beyond the imagination of our pencil-and-paper theories, whose primary value is in enabling us to see why there is and were problems that require solution but did not facilitate solutions such as those that emerged in these examples.

We have achieved little comprehensive understanding of the processes that show how either divergence or convergence may exist between the two concepts. In particular, our professional tradition is not geared to modeling ecological processes that can enable us to better understand emergent social systems. How, for example, might stateless groups discover specialization, comparative advantage, exchange mechanisms, and the supporting property rights that enable wealth creation?[3]

To explicate further the two kinds of rationality, consider this description using the perspective of game theory: Our professional approach to any observed problem area is to write down an abstract game model analyzing the phenomenon in a particular situation or institution – such as the free-rider problems in the lighthouse or in the grazing commons – contrasting the equilibrium of the model with optimality. The models of the institution and of an optimal outcome are each exercises in constructivist rationality. But in these exercises we take as given the abstract game situation or observed phenomenon, as well as the social structure and rules of the governing arrangements – for example, contracting lighthouse companies in

[2] For experiments in the voluntary provision of public goods based on incentive rules, see the references on the topic in Smith (1991); also see Ostrom et al. (1994) for a treatment of the theory, institutional analysis, and experimental and field studies of common property resource problems.

[3] A first effort to create an experimental design to examine these most rudimentary of all questions has been reported by Crockett et al. (2006).

a market, or developing a governing commons management institution in the Swiss Alps. The latter are natural or spontaneous examples of ecological rationality in the economy – self-governing institutions that emerged out of human experience. Now go to the laboratory to test, for example, a model of common property resources (such as the institution described by Netting, 1976); laboratory studies in this vein have been reported (see, for example, Cesari and Plott, 2003). The subjects in the experiment interact under rules derived from the observed field situation. The important difference is that because the experimenter assigns all the private values and costs in the economic environment created for the experiment, we can determine the equilibrium predicted to obtain and evaluate its optimality or efficiency. Suppose that the subjects converge to the predicted equilibrium in dynamic interaction over time. This is a laboratory example of ecological rationality showing the capacity of motivated subjects to achieve the efficient static outcome over time by unknown dynamic mental and social processes that are not modeled in these or other studies.

I think that improved understanding of various forms of ecological rationality will be born of a far better appreciation that most of human knowledge of "how," as opposed to knowledge of "that," depends heavily on autonomic functions of the brain. Human sociality leads to much unconscious learning in which the rules and norms of our socioeconomic skills are learned with little specific instructions, much as we learn natural language; think of it as the developing "social brain" at work. This contrasts with explicit learning of a new skill like playing a piano piece or bidding in an auction, which requires attention, emulation, and adaptation resources initially, but then soon becomes as unconscious a practice as any routine mental process that is taken over and guided by the practiced brain. This is what Polanyi (1962) calls tacit knowledge, and it has its own dynamics of acquisition through intuitive (inarticulate able) processes.[4] We learn social exchange without the self-aware application of attention, emulation, and adaptation resources, but the acquired skills enable gains from personal exchange that reward and help perpetuate that learning. Humans are not "thinking machines" in the sense that we *always* rely on self-aware cognitive processes, which is why all such approaches to learning are inherently limited.

[4] This knowledge is an essential reason why the transferability of results between the laboratory and the field, from one set of field observations to another, and between two laboratory experiments is fundamentally an empirical proposition, and not a methodological question to be settled by argument. (See Smith, 1982a, for discussions of "parallelism" between field and laboratory; field versus laboratory experiments is a frequently visited issue most recently explored by Harrison and List, 2004).

This book is prominently motivated by the results and methods stemming from the study of behavior in experimental economics and in field tests and applications, but these results and methods do not inform the whole of the book's message nor are they the only observations that give its message coherence. The laboratory evidence is broadly interpreted as a window on the human career – its development, meaning, and change. Accordingly, throughout I have tried to relate laboratory discoveries to history, anthropology, archeology, ethnology, field empirical studies, psychology (including its important social and evolutionary branches), animal behavior, philosophy, methodology of science, neuroscience, the history of ideas, and, indeed, life experience.

The very association of the word "ecological" with experimental evidence may seem strange to those who think that experimental evidence is somehow artificial, whereas "ecological" is natural. But "ecological" is just another word for the occurrence of a rule-governed, self-organized order, and I want to avoid compartmentalizing observations and labeling them in separate boxes without seeking unifying themes. We seek coherence, and if we are to find meaning we should not reject the idea that all humans in all situations are intuitive, feeling, searching, and acting organisms who do not naturally compartmentalize knowledge – except in formal modeling exercises – when they join the task of deciding and choosing. One should not presume that the actions chosen by a laboratory subject in a market, or in an anonymous interaction with another subject in an extensive form game, yields no insight into the human enterprise – at least not without a larger penetration of the experimental evidence and its use in test bedding in economic design and industry/policy applications, or without a larger examination of social science learning.

There are five fundamental propositions that inform much of the content of this book:

- Wealth creation depends essentially on knowledge and skill specialization. This includes innovation and technological change, because these are central parts of the acquisition of the tacit knowledge of "how."
- Specialization is possible only through the sharing and exchange systems that derive both from personal human sociality and impersonal market institutions. Hence, specialization is not a phenomenon that depends only on markets, although that is most certainly the source of its large-scale success in the modern world of wealth creation – a central theorem in Adam Smith's second book. Specialization and exchange

are far older than markets, which underscores the importance of better understanding human sociality.

- The personal knowledge that underlies specialization and exchange at any time is dispersed, private, and therefore asymmetric in all social systems.
- Neoclassical and information economics enable impersonal social systems to be characterized by an equilibrium given the state of existing knowledge, whereas experimental economics has demonstrated the efficacy with which people operating through extant impersonal market institutions are able to discover equilibrium outcomes through repeated interaction over time.
- Missing or incompletely developed in economics are models of how people are able to discover equilibrium outcomes so effectively given only private information and the message space of surviving market institutions; how the study of personal exchange systems can enable us to better understand their role in early human discoveries of specialization; how specialization and exchange relate to innovation and technological change; and how institutions emerge and survive in human socioeconomic development.

The main themes of the book may be summarized as follows[5]:

- What is generally known as "Das Adam Smith problem" (that there is an inherent contradiction between the *Theory of Moral Sentiments* and the *Wealth of Nations*) is an artificial problem in that "the propensity to truck, barter, and exchange one thing for another" applies to both personal exchange (which, as I see it in retrospect, is central to but certainly not all of the content of Smith's first book, which dealt broadly with human sociality) and impersonal exchange in markets (the theme of his second book). (Also see North 1990, 2005.)
- Both Hayek and Adam Smith (and his contemporaries, including David Hume and Adam Ferguson) well understood the coexistence of the two rational orders: constructivist and ecological. Many contemporary economists do not have such an understanding (but some do; see in particular the treatments by Binmore 1994, 1997; also Nelson and Nelson, 2002). Part of this book provides the conceptual foundation for these two rational orders.

[5] This summary draws directly on that of Andreas Ortmann, who reviewed an earlier version of this manuscript. I am much indebted and grateful to him for a thorough and inspiring review that enabled me to make many valuable revisions and additions to the text, the references, and the style.

- Traditional economic theory has long chased the fiction that purposive human action requires deliberate calculation based on constructivist rationality. Hence, for over eighty years following W. S. Jevons, who wrote in 1871, theory failed to anticipate that individuals do not require complete information to achieve equilibrium market outcomes in repeat interaction, a finding long replicated experimentally across many different, even quite complex, economic environments.
- In the midst of our constructivist adventures, and separate from them, institutions have emerged that are "ecologically rational" and that economists would be hard put to improve on, even if such institutions had always been an integral part of economists' perceived task. Underlying this second "rational order" are, roughly, Darwinian selection arguments. In the same way that natural systems such as an ecosystem or the human body (itself a cellular ecosystem) can regulate itself, so can social institutions (such as villages, cities, markets, associations, and scientific communities, which are all supported by endogenous property rights systems that sometimes become externally codified). Generally, human institutions and decision making are not guided only or primarily by constructivism, which is much more important in generating variations – social and economic innovations – than in selecting which ones shall survive.
- In achieving efficient cooperative outcomes in market exchange experiments, individuals are observed to maximize their payoffs, based on the use of monetary rewards to induce value (cost) on outcome states. The underlying classical model of behavior, *homo economicus*, appears thereby to be strongly supported in these impersonal exchange environments. But these exchange institutions in the laboratory are supported by externally enforced (property) rights to act that prohibit taking without paying, and giving without being compensated. Hence, action in the strict self-interest does not conflict with joint social betterment. But it would be quite wrong to conclude from this observation that across the extended range of experimental studies we always observe *homo economicus* taking action for his immediate interest alone.
- Thus, in Part III of the book I turn to an examination of the world of personal socioeconomic exchange, mostly as it emerges in two-person extensive form game trees. Exchange in these economic environments cannot lead to cooperative joint maximization without an individual being exposed to defection by her paired counterpart, who will

defect indeed if he is *homo economicus* and always chooses according to payoff dominance. But such defection is not the norm even in single-play games between anonymous players. If property rights emerge in these two-person exchanges, they must do so by mutual consent in the form of reciprocity and sharing customs that eschew either party taking without giving. In these situations, we observe other-regarding behavior that supports more cooperation than the standard model predicts.

- Many if not most scholars, in the belief that our experiments fully control for everything except preferences, have modeled these other-regarding behaviors as due to other-regarding preferences (utility) in the tradition of static equilibrium theory. This model confounds reputation-based reciprocal motives to cooperate through exchange with the notion that cooperation requires preferences to be altruistic. Consequently, I prefer not to refer to the *homo economicus* model but rather more generally to the standard social science model (SSSM), which may appeal to social preferences or other formalisms to explain the prediction failures of the static selfishness model (Barkow et al., 1992).
- There is an interesting parallelism in the way that our brains and the socioeconomic world evolve and function. Both the world and our brains have evolved problem solutions, essentially via forms of selection that are not a significant part of our formal reasoning efforts. Whereas in the world our social brains have evolved institutions to solve problems, the brain has evolved internal off-line parallel processing capacities that enable us to function in daily life without continuous monitoring and conscious control, an important adaptation to the emergent mind as a scarce resource. Our unawareness of these processes, and our egocentric tendency to believe that we are in control, lead naturally to what Hayek (1988) called the "'fatal conceit;' the idea that the ability to acquire skills stems from reason" (p. 21).
- Hayek's research program identified three complex emergent forms of order in the biological and cultural coevolution of the human career; all are prominent in the concept of ecological rationality: (1) the internal order of the mind; (2) the external order of social exchange; and (3) the extended order of markets. The first, an inquiry into the neuropsychology of perception, began in the 1920s and was completed and published by Hayek (1952). The second form of order concerns human sociality in small group interaction. In this book, this form is

particularly relevant to the study of exchange behavior in two-person extensive form games. The third, concerned with market order and welfare, is Hayek's best-known legacy. All these themes will be retrospectively evident in this book, although they were not part of the motivation and development of experimental economics or my own learning from experiment.

- Experimental economics allows us to study and better understand ecological rationality and the manner in which constructivist and ecological rationality can inform each other. Specifically, experimental economics allows us to test propositions derived from rational reconstructions of processes driven by ecological rationality to test the validity of those reconstructed interpretations. For example, in Chapter 12 we test the proposition that the cooperation observed in two-person trust games arises from reciprocity (favors are rewarded with favors) rather than altruistic preferences. Experiments provide a relatively low-cost methodology for studying that which is not or might be. But to do that successfully and comprehensively, we also have to look beyond the lab to related field studies and applications.
- This development, along with the many new computer communication technologies, has led to the important new subfield of Economic Systems Design (ESD), which combines constructivist tools and learning from experience (ecological processes) to fashion new group decision-making institutions, testing them in the laboratory and in the field, and modifying them in the light of experience. Testing is crucial because our constructions may err by failing to model the correct elements, by building on inappropriate assumptions, by being infeasible to implement or impractical for the participants, and so on.

Both the constructivist and ecological themes in this book apply also to method in science and experiment including economics. I explore that development (see Part IV) and use it to explain why the falsificationist thinking of scientists defines neither what scientists do or exclusively what they should do, although it explains much of what they say about what they do. I will also treat the logical incompleteness of the methods of science (and mathematics) and why the failure of all attempts to construct a rational methodology of science is not cause for alarm or postmodern cynicism. What saves the day is human sociality, as it operates in our scientific communities and enables us to muddle through in spite of the rhetoric of falsification tests. In this respect, human success in science is not so dissimilar from human success