Nature as First Custom: Hayek and Popper on the Evolution of Rules and Mind

Gerald J. Postema

"Nature is only first custom, as custom is second nature" -- Pascal

Questions about the nature of informal social rules have become a major focus of attention in legal philosophy and legal theory in recent decades. Since Hart insisted that modern municipal law rests on a fundamental *social* rule, the rule of recognition practiced by law-applying officials (and, perhaps also, rests more solidly on a wide base of network of informal rules or customs), it has been a preoccupation of much of analytic legal philosophy to explain the nature of social rules (or "conventions"). More recently, theorists have come to recognize that an understanding of international law cannot proceed very far without a solid understanding of informal rules ("custom"), for customary law still plays an important and foundational role in that domain. Also, in legal theory there has emerged recently a multi-disciplinary study of what are called "social norms," which arguably fall into another species, along with "customs" and "conventions", of the genus "informal social rules." Similar interest in informal social rules can be found in contemporary moral and political philosophy, often drawing conceptual resources and explanatory frameworks from game theory and socio-biology. In these fields, attention has especially turned to explaining how informal social rules emerge and change. An account of the origin and dynamics of such rules is thought to be fundamental to our understanding of how they function, which, in turn, informs our understanding of law, morality, and political institutions. This focus of philosophical attention is not new, of course. It finds at least one especially perceptive antecedent in the eighteenth century in Hume's attempt to explain the "origins" of justice and allegiance in his *Treatise of Human Nature* [Hume 1998a].

Karl Popper does not directly address this set of theoretical issues, but he offers some surprising insights into the evolution of human rational capacities that may be of use to those seeking an account of the dynamics of social rules that is open to the full range of resources human beings may bring to bear in shaping rules for their lives together. Friedrich Hayek, on the other hand, worked out, over a large number of works, a systematic explanation of the emergence and dynamics of informal social

rules. His theory repays study by those who are interested in contemporary discussions of the evolution of social rules, while its weaknesses invite consideration of Popper's view of the domain of "objective knowledge" and his account of the evolution of mind, or so I shall argue here. Hence, I propose to look at the views of Hayek and Popper on the evolution of mind and social rules, in the hopes that, taking them together, we can gain some insight into issues that have been at the center of attention in much recent moral, legal, and political philosophy.

I begin with an exposition of Hayek's framework for explaining the dynamics of what he calls "grown order", followed by a discussion of problems that threaten to undermine his explanatory scheme. I end with a consideration of ideas that, although not entirely welcome to Hayek, might put us in a better position to solve the problems that prove intractable on Hayek's theory.

I. The Dynamics of Grown Order: Hayek's Explanation of Social Rules and Institutions

Central to Hayek's social and legal theory is his explanation of the nature and dynamics of social rules and institutions, especially the market and law, as what he calls *grown orders*. His explanation is multi-layered, accounting for the social rules and institutions we see in terms of deeper levels which are increasingly less obvious to the casual observer. To understand the typical operation and dynamics of social rules, he argued, we must look more closely at the operation and dynamics of rules of the mind. The superstructure, as it were, of social institutions builds on, but never transcends, a rich and layered substructure of more basic rules of conduct and of thought. Our complex social institutions are anchored in this more basic substructure and are radically dependent on it.

Hayek's systematic explanation relies on three fundamental ideas: the idea of a rule and rule-following, and the "twin ideas" of spontaneous order and evolution, the two components of his idea of grown order. The three ideas are interdependent parts of a single, integrated explanatory scheme, designed to show that key elements of social life are ordered but not the product of some designer, but rather the "unintended consequences" of impersonal and external forces operating on behavior and thought of human beings directed to other ends and purposes. I begin with Hayek's notion of rules, because the other two notions work with an idea of social order regarded as the product of behavior directed by rules.

A. Hayek on Rules

1. Rules as subject-grasped and subject-directing patterns

For his purposes, Hayek deploys a very broad concept of a rule. Rules, as he proposes to use the concept, direct both thought and conduct. Rules of thought or mind concern matters of immediate perception, as well as judgment and higher order structures like mathematical concepts and abstract theories [Hayek 1967, 23-24, 43-46]. "Perception," for Hayek, includes everything from immediate sensory judgments (e.g., a rhythmic pattern of lights going on and off) to organized, albeit very particular, judgments about one's situation (e.g., the entrepreneur's sense that a certain product might succeed in the current market). Rules falling along this wide spectrum, on Hayek's proposed understanding, are (i) recognized or *projected patterns*—configurations of items or elements that are grasped (at least in the most basic forms) as a *Gestalt*—that (ii) generate a determinate *response* in the subject [Hayek 1967, 23, 45, 52; Hayek 1973, 75]. This proposal needs unpacking. Four elements of this proposal call for our attention.

First, rule-patterns exist only as "recognized" [Hayek 1967, 23, 45], although this "recognition" need not be conscious, let alone articulated by the subject. That is, the patterns may be grasped by the mind of the subject—they are "in" or "of" the mind—at a pre-conscious level, entirely "without intellection," or they may a matter of conscious mental construction. Thus, the patterns are not, strictly speaking, detected in the nature we perceive, but rather are responses to encounters with nature and projected onto it. Secondly, rule-patterns are always *abstractions*, patterns resulting from selecting and ordering certain elements and ignoring others of experience [Hayek 1973, 30]. Moreover, rule-patterns are, in Hayek's view, *generic* in the sense of being logically universal, but also in the sense of comprising lots of real (not merely logically possible) circumstances and instances.

Third, the grasped patterns are always accompanied by *dispositions* of *response*, either a disposition to *see*, *feel*, or possibly to *judge* something, or a disposition to act in a certain patterned way [Hayek 1973, 75, 79]. It is not entirely clear whether Hayek's view is that the pattern causally generates the disposition or that what the subject experiences is a *patterned disposition* to perceive or to act. I suspect that he thinks that, at least at the most basic levels, the latter is true, although at higher levels there may be room to distinguish the pattern grasped and the subsequent judgment or action. In any case, Hayek's rules are not inert: they are *determinants* of thoughts and especially behavior [Hayek 1973, 79]. Rules are not merely grasped by the subject, they *direct* the subject.

This supplies the foundation for Hayek's account of rule-following. At its most fundamental level, for Hayek, rule-following is a matter of "know how,"[1] that is, a disposition to act (perceive, judge) in a certain way, arising from a situation that, having grasped its significance, disposes one so to act. Rules, on this view, are never merely regularities or patterns; rather, they are *grasped* patterns that are or give rise to dispositions.[2] Thus, for Hayek, rules are *subject-grasped* and *subject-directing* patterns. Moreover, when we observe rule-generated behavior of things in our environment, including the actions of other agents, we understand each instance of the grasped regularity as having a common cause; not merely presenting us with a pattern, but also manifesting a rule-governed order. Rules manifest themselves in such regularities.

Fourth, since rules of conduct are dispositions to act, rather than actual patterns of actions, it is possible for a subject to be directed to act by a rule and yet not act on the rule. This will occur when the conditions for the realization of the disposition are not met; and, in Hayek's view, among the most important conditions for realization of a rule-disposition is the condition that there is not some other disposition operative in the subject which prevails at the time of action. Not only do rule-dispositions live in subjective environments along-side other rule-dispositions, but, in Hayek's view, rules of thought and conduct always exist together in complex networks. Rules are able to do their work, with subtlety and flexibility, in part because they get their content and meaning from their place in a system of rules, forming "chains" of interconnected meanings [Hayek 1967, 57-8]. This as true of our patterned responses at a very primitive psychological level as it is of sophisticated patterns of reasoning and codes of law.

2. The Primitive Evolution of Rules of the Mind

Rules of thought and action, in Hayek's view, *constitute* the mind of human individuals [Hayek 1973, 18, 30]. They are products of the encounters of human individuals with their natural and social environments [Hayek 1973, 17-18]. Hayek's account of the basic process by which "rules of the mind" are formed is an early form of what is now called "connectionism" or "neural network" theory [Gaus 2006, 248-52]. Roughly, the view is that, at the most primitive of formation, rules of the mind are the effects of external causes on an individual's sensory apparatus. The external world causes certain neural responses with some determinate configuration. The mind "grasps" a pattern when two events trigger the same configuration and that configuration, further, yields some response on the part of the individual, either a

[1] Hayek draws heavily on Ryle's [1949] distinction between "knowing how" and "knowing that."

[2] In his essay, "Rules, Perception and Intelligibility" Hayek traces our rule-following capacity to our nervous system that acts as both "a movement pattern *detector*," recognizing actions conforming to rules, and "a movement pattern *effector*" generating those actions in appropriate circumstances, or at least disposing us to act in those ways [Hayek 1967, 45].

phenomenal experience or behavior. So, for an individual to learn a rule or pattern is just for there to be established in that individual's brain a neural pathway that is triggered by multiple external events.[3] These external, connection-establishing events have their sources in the physical environment or the social environment of individuals. We can expect different individuals to have relevantly similar experiences and responses to the external world to the extent that they interact with a similar environment (and the causal mechanism by which neural networks are established works in a similar manner in those individuals). This potential for overlapping rules of the mind is reinforced by the influence of the social environment on the development of the mind. The root of social "learning" is the inborn capacity of human beings (and other higher animals) to mimic the behavior of those around them [Hayek 1967, 47-48]. Following the lead of eighteenth century Scots,[4] Hayek observes that, even very young infants, without the benefit of a mirror to observe their own movements, are able to reproduce the movements or gestures of those around them. These mimicked movements establish a network which is triggered later by other behavior registered by the individual as similar. This provides the basis for learning of routines of action and of perception and thought. And this learning not only crosses sensory modalities (as in the primitive case across sight and kinesthetic modalities), but also boundaries between persons. Rules are "grasped" simply by being enacted, as it were, in the behavior of the individual learning them, where "enacting" means that a disposition to respond is established by the neural configuration established.[5] Moreover, Hayek observes, in our social environments we not only learn certain common behavioral routines, but also their "meaning". We perceive in the movements of others the agent's mood or attitude that makes the movements intelligible to us [Hayek 1967, 55, 59]; we grasp the behavior "purposive" rather than random, not just behavior with some regularity, but rule-following behavior. This in turn enables a degree of understanding across minds, a basis for a degree of *Verstehen* [Hayek 1967, 58-60].

[3] Note that on this view we cannot infer that the external world comes already patterned, but only that the mind responds to stimuli by organizing them into patterns. "Abstraction," Hayek insists, is not an advanced activity of the mind but rather absolutely the most basic and primitive. The mind's initial response to the external world is a pattern-forming response [Hayek 1973, 30].

[4] Hayek cites Dugald Stewart and Adam Smith, but the locus classicus of this line of psychological observation is Book II of Hume's *Treatise of Human Nature* [Hume 1998a].

[5] Hayek does not explicitly acknowledge the further important fact that it is typical of human learning that the "similarity" of the responses is not simply a matter of the events triggering the same neural configuration, but of its being recognized by others and that recognition being recognized by the learner. In this way, a distinctively social form of learning, not matched by learning in a subject's physical environment, takes place. This, in Hume's view, is one key source of the human capacity for what he calls "sympathy".

3. The Implicit Dimensions of Rules

A core feature of Hayek's theory of rules is his doctrine, repeated with the frequency of a mantra, that subjects directed by rules of thought and action need not be, and indeed predominantly are not, aware of these rules. The rules are implicit, matters of only tacit understanding. They are, as he says, "known by none, but understood by all" [Hayek 1967, 46]. Hayek's doctrine of the implicit dimension of rules comprises several related claims, some based on observation or argument, some asserted but never adequately defended. He begins from the observation that we are able to act on very sophisticated rules without even the slightest awareness of them. His favorite example is that of children who manage to use language with great facility without any awareness of its rules of diction, grammar and syntax, let alone a capacity to articulate those rules [Hayek 1967, 42-44]. He then goes on maintain repeatedly that this is true about the vast bulk of our (patterned, rule-governed) knowledge of our physical and social world. Moreover, not only are these rules (hence, this knowledge) currently unarticulated, but the vast bulk of it cannot be articulated or even brought to our awareness. This is, in part, due to the fact, as Hayek sees it, that most of these rules are highly localized, restricted to certain times, places, and circumstances of individuals and embedded in the particular activities and skills of their ordinary practical lives. These rules are so deeply embedded in their practice that they cannot be brought to consciousness without distilling away most of their content. The problem, it seems, lies in part in the fact that something can be made explicit to consciousness only if it is articulated linguistically, and we lack the resources to articulate the content linguistically. But it is due even more to the fact that, we could not capture it even if our linguistic resources were far more sophisticated, because it is so vast, interconnected, and embedded in practice. Thus, inevitably, a very large part of the whole which gives determinate meaning to any given rule in particular circumstances remains inaccessible to the agent who learns how to follow it. Moreover, since we are unable to make this knowledge explicit, Hayek concludes, we also cannot share it. It is widely dispersed and in very large measure private.

This conclusion rests on a very strong assumption of *subjectivism* (maintained despite the potential for a more modest version represented by his recognition of the possibility of *Verstehen*). This very general and deep assumption takes various forms in Hayek's work. For our purposes it surfaces in two forms. (a) In his epistemology, subjectivism takes the form of the claim that "knowledge exists only as knowledge of [i.e., possessed by] individuals" [Hayek 1960, 24]. He rejects any idea of social, shared, or common knowledge.[6] (b) The second form in which subjectivism surfaces

[6] Although he rejects any notion of shared or common knowledge, he does insist both that widely dispersed and largely private knowledge is nevertheless indirectly *available* to individuals and that this knowledge is *embedded* in rules. The market is Hayek's favorite example of how dispersed and private knowledge is nevertheless socially available. The market is a framework of rules that serves to

is in his understanding of methodology of social explanation.[7] In Hayek's view, our understanding of social rules and institutions must not only start from, but must also always be ultimately reducible to (or in some other suitably strong sense brought back home) to statements about the mental states of individual subjects. Hayek recognizes that it is possible for us to gain some understanding (*Verstehen*) of each other, but this involves grasping what another mind desires, intends, or "means" [Hayek 1967, 58-60]. It is a matter of grasping something in or of the mental state of another person and this grasp will always be subject to very severe limits, because we can only grasp what is conscious to the other person and that, as we have seem, is only a very small portion of the basis of their perceptions, judgments, desires, and purposes.

B. Two Explanatory Models: Spontaneous Order and Evolution

Commentators and critics often treat Hayek's evolutionary explanation of social rules and order as independent of and in competition with explanations drawn from the idea of spontaneous order. This, I think, is a mistake. In Hayek's eyes interdependent explanatory schemes.[8] His account of the evolution of social rules depends heavily on the idea of spontaneous order and the role of rules in producing that order; moreover, the sources of disequilibrium, and hence innovation (mutation) and reproduction (replication), needed for the evolutionary story, occur within the process explained by the spontaneous order scheme. At the same time, the notion of spontaneous order presupposes rules that direct the behavior of individuals from

---

coordinate the actions and interactions of countless numbers of agents, each market player acting on his or her own local knowledge. The price system does not itself "contain" within its mechanism the knowledge of each player, but it enables each to adjust their decisions to the knowledge abstractly represented by the prices offered. Prices are, as it were, content-independent markers of dispersed knowledge that remains essentially unarticulated and inaccessible in any more direct form. In market economies, knowledge drives activities of parties without the need for any central accumulation of that knowledge. Such knowledge remains dispersed, it is never common.

Similarly, Hayek maintains that knowledge is "embedded" in social rules just in the sense that the process by which they have evolved ensures that they are tested against a wide range of circumstances and have proved to be adequate adaptations (adequate for group effectiveness, as we shall see) to those circumstances. Again, there is, strictly speaking, no accumulated wisdom of the ages stored in these rules; rather, the rules are simply the product of an impersonal (and for that matter, content-independent) processes which nevertheless offers promise of a substantial degree of success in day-to-day social interactions.

[7] A classic statement of Hayek's methodological subjectivism can be found in *The Counter-Revolution of Science*: "Not only man's actions towards external objects, but also all the relations between men and all social institutions can be understood only in terms of what men think about them. Society as we know it is, as it were, built up from the concepts and ideas held by the people; and social phenomena can be recognized by us and have meaning to us only as they are reflected in the minds of men." [Hayek 1952, 34-5.]

[8] This interdependence is clearly evident in Hayek 1967, ch. 4. Heath (1992) and Gaus (2006) are rare among readers of Hayek to recognize this interdependence and only Gaus, in my view, comes close to understanding the nature of this interdependence.

which the order emerges, rules that, on Hayek's account, have emerged from an evolutionary process; moreover, spontaneous order explanations are incomplete explanations of the dynamics of social order and social rules, because (except within certain limits) spontaneous order explanations are static such that when internal forces no longer suffice to bring the disorder back into disequilibrium, the evolutionary account must explain how a new order is established. Thus, to understand Hayek's proposal for explaining social institutions, we must relate these two explanatory schemes. Although they are interdependent, it does not distort them too much to view them as working in two stages, following the trajectory of a spiral rather than a vicious circle. The image of a spiral also allows us to capture the idea that the interdependent processes of spontaneous order and evolution build on previous and in some cases more basic stages.

1. Spontaneous Order Explanations

We are tempted to regard all manifestations of order in nature and social life as products of design-governed efforts; however, Hayek argues, many ordered structures must be explained as undesigned, endogenous, and self-generating. The "order" or observed pattern emerges from the interaction of a large number of elements responding to their environment (including the behavior of other elements) according to certain forces or rules that direct that behavior. Hayek's model of spontaneous order applies to both natural phenomena like the formation of crystals or patterns of iron filings and social phenomena like a living language or the market [Hayek 1967, 39-40; Sugden 1998b, 485]. I will focus here on spontaneous social orders.

To begin, Hayek distinguishes between rules of conduct and the social order that they (indirectly) generate [Hayek 1967, 66-69]. "Order" as Hayek understands it, is that "state of affairs in which a multiplicity of elements of various kinds are so related . . . that we may learn from our acquaintance with some . . . part of the whole to form correct expectations with regard to the rest" [Hayek 1973, 36]. Social order is an emergent property of the actions and interactions of a large number of agents, that is, it is an abstract pattern manifest in the interactions of particular individuals which may persist even if all the individuals are replaced by others [Gaus 2006, 233-4]. This pattern is the product of (i) the actions of large numbers of individuals (ii) in an environment of a determinate nature which (iii) includes the actions of others, all of whom (iv) respond to local knowledge of that environment (v) from a potentially wide variety of motives (vi) within the limits defined by the system of rules in force in the group. This order is "spontaneous" because it is the result of individuals arranging themselves according to "forces" (in the social context: motives within the framework defined by rules) in a specific environment. The order is the resultant of the balance of these forces [Sugden 1998, 487]. Because of the interaction among the individuals and the feedback from this interaction, the properties of the order are not simply the aggregate of the properties of individual elements, but rather are emergent from them.

The relationship between the rules operative in a given social context and the order that emerges is indirect and complex, because the order emerges from the combined influence of the rules and the environment on the choices and consequent interactions of the agents. Thus, it is not the case that every set of rules can be expected to produce a corresponding order; indeed, some rules may prevent any order from forming or may produce an order that is dysfunctional from the point of view of the group or of its individual members [Hayek 1973, 43]. Also, it is possible that the same social order may be produced by different sets of rules and that the same set of rules may yield different social orders [Hayek 1967, 67-68; Hayek 1973, 43-44]. A change in the rules may not result in a change in the social order and rules may produce very different orders in environments that have significantly different properties. The environments in which individuals interact and the way and the extent to which the rules influence the actions of individuals, greatly affect the relationship between rules and the resulting social order or disorder. Finally, a social order may emerge even if there is no universal regularity in the behavior of individual members of the group. This lack of uniformity can be of two broad kinds. (i) The rules may call for quite different routines of conduct from different people in different roles, stations, or circumstances. What is important for social order is not uniformity of behavior across the membership of the group, but its coordination. Coordination requires only that the rules be broadly compatible, making possible coordinated behavior of those directed by them.[9] (ii) There may be some, perhaps even a substantial, degree of irregularity of behavior (i.e., deviations from the rules) in the group. Just how must irregularity or deviance a social order can tolerate is determined by a wide variety of factors, some environmental, some psychological, some having to do with the internal relations among the rules. Moreover, Hayek realizes, that must be some degree of this kind of irregularity if the social order is to have the flexibility needed to cope with exogenous shocks, and to permit endogenous changes, that cause adaptive changes in the order.[10]

Sugden and Gaus have identified several salient features of spontaneous orders as Hayek conceives of them. First, they are *path-dependent* in the sense that the properties of the order at any point in time depend on its history [Sugden 1998b, 488; Gaus 2006, 233]. Second, they *approximate*, but never strictly achieve, *equilibrium*

[9] This is a consequence of Hayek's view that rules come linked together in integrated packages, rather than merely aggregated in sets. But the idea is in tension with his official view that rules of spontaneous orders are "abstract" both with respect to the ends served by them and with respect to the agents and circumstances to which they apply [Hayek 1967, 56; Hayek 1973, 50].

[10] This flexibility, as we shall see, is essential to his scheme of evolutionary explanation, but again this feature is in tension with his frequent insistence that rules must be followed *rigidly* [Hayek 1967, 90-91]. The latter dogma, like so many other statements of Hayek's, is overly broad and incautiously formulated for his primary purpose, which at the point in the text was to counter the idea that rules are mere rules of thumb for agents who are act-utility maximizers (governed by "expediency" rather than "principle" [Hayek 1973, ch. 3]).

[Gaus 2006 234], with the result that there is always some greater or smaller degree of disequilibrium in the system. Nevertheless, third, spontaneous orders are, within limits, *self-maintaining* [Gaus 2006, 234]; that is, it can survive exogenous and endogenous shocks, restoring its (approximation to) equilibrium. Finally, Hayek recognizes that the spontaneity of a social order is a *matter of degree*.[11] As Sugden points out, spontaneity is a function of at least two properties: dispersion of power and redundancy [Sugden 1998b, 487]. If we understand "power" to be the extent to which an individual can influence the properties of the social order, then we can see that the more widely power is dispersed over a population, the less power each individual will have; and, thus, the greater the dispersion, the greater will be the spontaneity of the order. Similarly, spontaneity is in part a function of the density (the number and overlapping nature) of the relations among the members of a group and the interchangeabilty of the parts. Together, these yield redundancy in a system: the greater the redundancy in a system the likely is it to be less affected by deviations of small numbers of the members. Dispersion of power and redundancy admit of degrees, and, as a consequence, so will the spontaneity of an order.[12]

With these general features of spontaneous orders in we can get a sense of the nature and limits of the dynamic movement within spontaneous orders as Hayek understands them. First, changes in the environment in which members interact (exogenous shocks) may lead to adjustments of behavior of the members within the parameters defined by the existing rules, thereby re-establishing the order. We might call this a case of simple self-maintenance of the order. It is also possible that exogenous shocks (or endogenous challenges to the rules) result in change of the rules. This change may not produce an overall change in the order, in which case we have a more complex form of self-maintenance, or the change of the rules may be substantial and influence the integrity of the social order. Changes of some rules may bring about shifts in other rules of the system and these adjustments may restore the (near) equilibrium of the order. Other changes in the environment or changes of the rules may yield require substantial adjustments in the behavior of members of the group thereby altering the nature of their interactions. In that case, the emergent social order will also change resulting over time either in disorder or the emergence of a new order with different properties.

In each of these cases, individual members may be affected, as may the felicity and fortunes of the group as a whole. Hayek makes clear that the fact that an order emerges spontaneously from the interactions of a group does not guarantee that the

[11] Hayek 1973, 41-42, 45-46. His penchant for sharp dichotomies, especially between made/imposed order (*taxis*) and grown/self-generating order (*kosmos*), often obscures this fact, perhaps not entirely unintentionally.

[12] The important conclusion we must draw, although Hayek obscures it in many ways, is that we may have to ask what degree of spontaneity of a social is desirable, and what reasons ought to guide that choice.

order is beneficial, let alone optimal, either to individual members or to the group as a whole [Hayek 1967, 67; Hayek 1973, 43-44]. Indeed, it is possible that a set of rules may even prevent order from emerging, or bring about damaging and socially dysfunctional disorder. Hayek's notion of order is entirely value neutral and the fact that an order has arisen spontaneously implies no special value and offers no guarantee of its being in any way beneficial.[13] There is nothing in the idea of spontaneous order to ensure that a coordinated order will be achieved through interactions of individuals (even if they are directed by or act within the limits of rules), neither will insure that order, once achieved, will be maintained. If social interaction achieves and maintains order, this will be due in part to forces outside those operative within spontaneous orders.

Thus, the idea of spontaneous order provides a model for explaining the emergence and alteration of social rules. On this model, new rules emerge and are altered in response to changing environmental conditions or in response to changes in rules that result from the irregular behavior of some individual members. The balance of forces within the order brings about these changes, without the intervention of any designers who have a view of the whole system of rules and the order it tends to produce. But we are left with several major questions. One question is: where do the rules that initially structure the spontaneous order originate? Another is: what more precisely is the process by which rules are adjusted in response to exogenous and endogenous shocks? How are rules changed? What role does the judgment or practical reasoning of the individual members play in responding to the shocks, or in creating those (endogenous) shocks?[14] And how are rules that actually prevail selected, if, by hypothesis, this is not done by individuals taking account of the impact of changes of the rules on the social order as a whole? What reason have we to think that some order will be achieved and that it will be in some sense beneficial? For answers to these questions Hayek directs our attention to his account of social evolution [Hayek 1973, 44], the necessary complement to the explanatory structure provided by the idea of spontaneous order.

2. Evolution of Social Orders and Social Rules

Hayek's account of the evolution of social rules and institutions is a generalization of the Darwinian account [Hayek 1967, 32], but the precise components and mechanisms of Hayek's account are difficult to pin down. This is

[13] Of course, later Hayek seeks to link spontaneous orders with individual freedom, but that is not part of his initial construction of the idea of spontaneous order as an explanatory device. And this extension of the concept of spontaneous order depends on principles or evaluative premises that are not at the core of the notion itself.

[14] That is, how do changes in the group's rules occur and to what extent are these changes the product of deliberate choices and actions of members of the group? These questions lie at the core of both Hayek's account of spontaneous order and of the companion theory of evolution of social rules.

largely because never in his many discussions of social evolution does he offer a careful, systematic statement of his theory and it is difficult to reconcile the many different partial accounts one finds scattered in his work. This is not the place to attempt to reconcile all these passages; rather, I will offer a reconstruction that seeks to remain faithful to the central motivation of Hayek's account in the hope that is results in a plausible version of that account.

Every evolutionary explanatory scheme must provide (i) a mechanism of selection, including specification of (a) the unit of selection and (b) the basis for selection, and (ii) a mechanism of change, including accounts of (a) the source of innovations or variation ("mutations") and (b) their reproduction ("replication") in the population. Let us look at these elements in order, beginning with the mechanism of selection.

Unit of Selection. On Hayek's account, the evolutionary selection of social rules is indirect, the result of evolutionary forces operating on the social order as a whole; that is, the selection social rules is the result of *competition among social orders* [Hayek 1967, 71]. However, this does not yet determine the unit of selection; it does not tell us whether we should look for the effects of changes in the social order on the felicity and functioning of individual members or on that of the group as a whole.[15] Hayek is often criticized for being inconsistent about whether he favors a group or an individual adaptiveness criterion [e.g., Heath 31-33]. Gaus maintains that Hayek embraces both independent accounts and holds that social orders are subject to two competing forces of evolutionary selection [Gaus 2006, 240-46]. I believe, however, that Hayek thought that the two elements are closely integrated in a single explanatory account. Impacts on and competition among groups and among individuals are both important for his unified account and inseparable.[16] For, on his view, groups have no aims and enjoy no benefits of their own apart from the aggregate good of individual members. At the same time, individuals benefit—they "succeed" in carrying out their ends and aims, as he likes to put it—only when there is an effective social order that coordinates their efforts and interactions with other members of the group. Moreover, it is judgments of relative "success" of individuals acting within the framework of a given system of rules that is an important part of Hayek's account of the mechanism of change in his evolutionary story. We have no guarantee, of course, that a given social order that functions well for the group as a whole will prove optimal or even beneficial for each individual member; so there is room for familiar problems of collective action to complicate the individual/group relationship and critics are quick to point out that the evolutionary process that depends on group selection can be

[15] It also leaves unspecified how the relevant group is to be determined, but I will ignore this indeterminacy.

[16] See Hayek 1973, 18 and 80 where the two are closely linked—or, as some critics would have it, confused.

undone by such collective action problems. However, Hayek tends to downplay the potential conflict over the distribution of the benefits of this group success and the possibility of substantial opportunities for individuals to ride free on the cooperation of others. We will explore his reason for doing so in Part II. B.

Putting this issue aside, it is possible to say with reasonable confidence that, although he thought individuals play an important role in the evolutionary process, Hayek took the unit of selection to be the social group [Hayek 1967, 67-68, 71-72; Hayek 1973, 9, 17-19, 74, 99, *passim.*]. The effects on individual felicity and functioning play an important role in the process, as we shall see, but evolutionary forces of selection work group level in Hayek's model.

Basis of Selection. Social orders are selected by evolutionary forces, according to Hayek. The basis for selection is primarily and ultimately the "success" or "effectiveness" (sometimes he says "efficiency") of the group the interactions of whose members tend to manifest a certain order relative to and in competition with the relative success of other groups in the vicinity [Hayek 1967, 67, 71-2; Hayek 1973, 11, 17, 80, 99]. Hayek does not specify this criterion of "success." We are told that the more successful groups "prevail over" or "displace" their competitors [Hayek 1967, 70; Hayek 1973, 9, 18, 99], or perhaps simply grow and, thereby, are better able to produce wealth and conditions for decent life for their members [Hayek 1973, 80]. Some readers contrast group survival with group growth [Heath 32-33; Gaus 2006 240-43], but Hayek seems to think these are closely related. Groups "prevail" over other groups, he maintains, not necessarily through a clash of forces and the literal destruction of rivals, but rather through doing a better job of enabling individuals to achieve their goals. They, thus, tend to attract members of other groups, leading eventually to the demise of the rival or its assimilation into the more successful group [Hayek 1973, 169 n7]. Hayek's basic thought seems to be that it is through doing a better job of guiding expectations and coordinating interactions of individual members than their rivals groups grow stronger, wealthier, and more powerful and are then able, either through conquest or through assimilation to win in competition with rival groups. Hayek does not rule out evolution that is red in tooth and claw, but he seems to think that it more typically proceeds in a more pacific manner.

Groups prevail in virtue of the properties of their social orders, properties which are products of interaction structured by the groups' rules of conduct. Thus, social rules are selected for their contribution to the "success" of groups (consisting of the "success" of their members). Notice two key features of this account of the selection of rules. First, it is moves entirely without design at the social level (although it may involve a vast number of locally oriented, goal-directed decisions and choices by individuals). Selection operates at the group level, but no agent or collectivity of agents decides or acts at that level. Second, rule selection is relative to several conditions. Selection is relative (i) to the system of rules of which it is a part, and

hence to the history of the development of those rules, (ii) to the environmental conditions in which the group which practices the rules must function, and (iii) to the groups that happen to be in the vicinity at the time and in the environment and compete with that group. Thus, there is no basis for concluding from the stable existence of a system of rules in a group at a given time that those rules are optimal, or optimal for that group, or even for that group in that environment. For the only rules tested are those that in fact developed historically in the prevailing group and in its rivals. We cannot even conclude for evolutionary success that the rules operative in a group at a given time are superior to those of its past. Past rules might actually be better for the group, but they may no longer available, given the evolutionary history of the group [Gaus 2007, 163-4].[17]

Mechanism of Change—the process of innovation and replication. Evolution is a dynamic process, so in addition to the mechanism for selection, we need an account of the forces that introduce and replicate changes which then may bring about changes in the social order on which the forces of selection operate. We can gain a sense of what is needed at this point in the explanatory structure by looking at the analog of species evolution. Biological evolutionary forces operate on traits of individuals of a given genotype and changes of the genotype result from mutations of genes in individuals, which are then passed on to other individuals through reproduction. Changes producing individual traits that enable the species better to meet the challenges of its environment are selected; those that do not, die with the individuals and their offspring that carry the mutations. Evolutionary adaptations of a species depend on just enough flexibility of the genetic structure to allow for mutations combined with sufficient rigidity to insure that mutations are transmitted with fidelity to other individuals of the species. The mutations come from random, exogenous influences on the genes as new individuals are produced.

In Hayek's model of social evolution social order is the analog of individual traits and rules play the role of genes. Thus, Hayek's model needs (i) rules of some degree of flexibility, (ii) a process by which variations in rules can arise, and (iii) a process by which variations are transmitted to other individuals in sufficiently large numbers that they can have some impact on the social order as a whole. Hayek has something to offer on each of these points.

First, Hayek maintains that social rules, although they call for strict adherence, are "voluntary" in the sense that deviations are possible and are not so severely sanctioned that individuals never have an incentive to consider deviation [Hayek 1960, 63]. Or perhaps we should say that social rules are adaptable to the extent that

[17] Thus, again, if Hayek wishes to draw conclusions about the rationality or merits of evolved rules, he must do so on the basis of premises not included in this explanatory scheme.

they enjoy this flexibility. Second, changes in the rules, on Hayek's model, are the results of decisions and actions of individuals seeking to realize their goals with a view only to local circumstances and local effects of their actions within the framework of the established rules. Changes arise from individuals engaging in "trial and error" testing, which can have its roots either in *mistakes* or intentional *experimentation*. Actions that deviate from the rules are assessed in terms of their relative success in furthering the goals of the agent. Deviations have their causes in changes in the environment or in individual's imagining new ways of adjusting to the existing environment. While these changes may influence the social order as a whole, individuals respond only to local conditions without appreciation for such systemic effects.

This account of the initial causes of variations makes several assumptions about individual agents. First, although they are not equipped with a view of the operation of the social order as a whole, they must be to some degree both self-aware and situation-aware. Moreover, they do not act on established rules entirely uncritically. They appreciate that the rules make a demand on them, which they ignore only at some cost, but they are sometimes willing to risk paying those costs in order to better realize their goals. This judgment of there being a better chance of realizing their goals may be limited to a single rule that seems to stand in their way, but, since rules come is complex packages, it may also involve a more complex assessment that includes awareness of the way that rules interact in particular circumstances to limit or expand opportunities for successful realization of goals. That is, these characters are norm-appreciating, norm-following, self-aware and situation-aware local optimizers (or satisficers), who may also be aware to some extent of how the system of rules to which they are subject work together to structure the situations and options they face. Hayek's language of "trial and error" is vague, but he must have something along these lines in mind.[18]

How then are these modifications established as rules for the group? We learn from each other by example and imitation, Hayek argues, although neither the teacher nor the pupil may be able to articulate the rule they observe [Hayek 1960, 28-9; Hayek 1973, 19; 1977, 166]. Although he insists that learning by experience is "not

[18] I am here drawing out implications of Hayek's vague language of trial-and-error-generated changes in the rules. Hayek frequently claims that this process is unintentional and blind relative to larger purposes and aims. If he means by this that changes in individual rule-following behavior *happen* but the changes are *not made* by the individual, then we must conclude that Hayek has no account of the mechanism of change and his model of social evolution is fatally incomplete. On Hayek's model, mere changes in behavior are not directly selected by evolutionary forces, because those forces operate directly on social orders. The changes in behavior become evolution-relevant only when they congeal into rules which are replicated in the decisions and actions of a sufficient number of members of the group. Thus, to save Hayek's account, we must take his talk of the unintentional and blind character of the process to refer to individual's lack of awareness of systemic effects and purposes at the level of the social order, leaving him space to develop an account of micro-level intentional activities of individuals along the lines suggested above.

primarily by reasoning, but rather by observance, spreading, transmission, and development of practices" which prove successful [Hayek 1973, 18], nevertheless, it is not in mere (unquestioning) *observance*, but in *observation* of the rule-directed behavior and its local success, that imitation is rooted [Hayek 1960, 28]. Imitation starts small, but through the accumulation of large numbers of such small deviations yielding individual rules ("practices"), which then catch on with others, the rules eventually spread through the group to a point sufficient for them to be established as a group practice and have some influence on the social order as a whole. Moreover, since the rules in question have been taken up and used by lots of people they prove to be serviceable in a wide variety of circumstances [Hayek 1976, 21; Heath 1992, 42]. Not all individual rules are imitated, not all imitated rules spread, and not all that spread get established in the group as a whole, but some do and among those that do some will introduce changes in the social order that better equip the group to meet its challenges. Of course, some rule-changes may make the group less effective, and in that case rules that catch fire may die with the group that practices them.

II. Problems of Identification and Normativity: the Possibility of Common Social Rules

Hayek's explanatory theory, integrating two complementary explanatory schemata, is impressive in its scope and ambition, if disappointing in its lack of rigorously articulated detail. One may wish to challenge the theory at several points, but I propose to inspect just one aspect of process of emergence of rules on which both schemata depend: the mechanism of change in the evolutionary story which is also the pivot of the equilibrating mechanism of the spontaneous order story. Hayek's discussion at this point is critical for the success of his explanatory account as a whole; it is also the point at which his theory joins issue with recent attempts to explain the nature and dynamics of social rules, customs, and conventions. Let us then take a closer look at Hayek's account of the process of rule-change.

A. The Tasks

To begin, it is useful to note an important difference between biological and social evolution. Biological evolution is *endosomatic*, as Popper would put it; that is, it proceeds by selecting species physical traits that are expressions of genes. As we have seen, in Hayek's account of social evolution, rules play the role of genes. Social evolution is quasi-endosomatic (or, if we tolerate neologisms, *endopsychic*). I say "quasi," because the rules are rooted in subjective dispositions of thought and action (which, of course, supervene on a somatic base). Hayek's story of social evolution is a story of rule-formation, rule-transformation, rule-transmission, and group rule-

adoption. This story introduces a level of complexity and a set of problems not encountered in biological evolution, for what must be explained is the emergence and establishment of rules in the behavior of a group. There are at least four distinguishable tasks: to explain (i) how it is that *rules* emerge, which (ii) are *social* rules and (iii) the *same* rules across individuals, which (iv) then *spread* through the group as a whole. Let us look at each of these tasks.

First, Hayek's must explain how it is that *rules* emerge and change through the activities of an individual's "mistakes" or "experimentation." Rules that allow for change are flexible because are "voluntary." But this flexibility must be of a certain kind. The pattern-consistent behavior now in view is *called for*, not merely produced by, the rules, and off-pattern behavior must be understood to be a *violation* of the rule, not merely a deviation from the pattern. That is to say, the rules now in view have an essential normative dimension. Thus, for an individual to grasp the rule, it is not enough that she behave in rule-consistent way; she must also grasp it *as a rule*. This does not require, of course, that she be able to articulate this recognition, let alone be able to explain its rationale, but it does require that in her practice of the rule she understands both that it is possible to act off-pattern (that her compliance is to that extent voluntary) and that off-pattern behavior is not merely *deviation* but *deviance*—that it is not merely different from what the rule would lead one to expect, but that it *fails* to conform to the rule. For this, the "flexibility" required of the rules is two-fold: deviation must be possible and the individual must have some degree of distance from her disposition, such that the option of deviance is open to her. This distance is even more important if she is to relate consequences of her deviance from the rule to the rule itself in his "trial and error" experimentation. Indeed, on Hayek's model, the individual must be aware of the rule and its suitability for the situation he faces—understanding that situation in terms sufficiently general for his relating consequences back to the pattern which also applies to other situations—and the place of the rule in the complex of rules that give each rule its meaning.

In addition, individuals must be able to recognize the difference between mere *non-conformity* with a rule and *conformity* to a *different* rule. This generates two problems or tasks. First, if new rules are to be introduced by the behavior of the individual, if only on a trial basis, the individual must have a motive to look to an alternative rule, rather then merely to exploit opportunities for improvement of his condition within the regime of existing rules. Hume's "sensible knave" experiments with various forms of conformity and non-conformity with existing rules, but has no interest in introducing new rules into the regime. He aims only to take advantage of the convenient cooperation of others. Hayek's account needs some answer to the challenge posed by the sensible knave. Second, the rule-tester must have some means of "enacting" or putting in place an alternative rule by means of his deviant behavior. To put the problem in terms familiar to students of customary international law, the

question is how is it that *ex iniuria oritur lex*?—how can deviance create new rules? In Hayek's framework, the answer to this must start with the actions or attitudes of the individual rule-innovator, they must choose to "enact" a rule for themselves, and it continues with an account of how this innovation is taken up by others. Unfortunately, Hayek is silent on the initiation of the process, but assuming some account of this process we must consider the problem of up-take, which has two dimensions: how to establish rules that are rules *for* the group (social rules) and rules *in and of* the group (taken up by the group).

Thus, Hayek faces the task of giving account of the rules as *social rules*; that is, he must explain how rules emerge, not just *personal* rules for the individual who is "testing" the rules, but rules *for* the group. This is necessary because the rules he seeks to explain are rules coordinating complex interactions among agents whose actions are interdependent—the outcomes of the actions of each are the vector sum of the actions of all in a context of limited space, time, and resources—and who must be aware of this fact and of the fact that others are aware of this. That is to say, they must be, at a minimum, strategically rational. This awareness, we must assume, is available to individual rule-innovators because they are aware of their local situations and these features are impressive, salient features of those situations. So, in circumstances characterized by a high degree of interdependence and the persistence of coordination and cooperation problems, the individual must not view the world as exogenously determined parameters for his own decision, but rather he must look at the whole system of interactions and consider a rule, or a number of subtly interconnected rules, which the all those involved in the concrete problem of interaction can jointly follow. In Hume's vivid image, the task for rule-innovating individuals is not that of replacing one brick in a wall with another, but rather replacing a stone in an arch or vault, each stone of which depends on the all the others for the stability and integrity of the vault [Hume 1998b].[19]

The task for Hayek's evolutionary account at this point is to explain how individual rule-testers identify rules of the kind that can perform this complex social function. Moreover, no adequate "test" of such a rule can be performed unilaterally. The rule must be taken up to some degree by others in the group facing the common interaction problem. As Lon Fuller pointed out years ago [Fuller 1969, 4], getting customs or conventions started in conditions of complex social interaction is not like blazing a path through the undergrowth, each successive party treading the path

[19] In game theoretic terms, what the individual seeks is some device that yields a "correlated equilibrium" [see Vanderschraaf 1995; Postema 1998]. The rule, "go on green, stop on red" at intersections that have red-green traffic signals is such a rule, as is the rule "yield to traffic approaching the intersection from the right, otherwise proceed." Note that these rules call for different behavior from, i.e., they assign different "roles" to, the parties approaching an intersection depending on their relationship to some external feature of their common situation (the traffic signal, e.g., or the spatial relations of the parties).

making it more distinct and less formidable. Establishing the rule itself requires coordination.

This brings us to the third task such an account of the emergence of social rules faces. The rules in question must not only be rules *for* a group, they must be rules *in*, practiced by, the group (or some sub-group large enough to provide a good test of the rule). That is to say, the rules must be passed on, transmitted to others. Hayek's proposal here is that rules are transmitted by "imitation." The suggestion is that an individual enacts a rule in his own behavior (however that is accomplished) and this rule is observed, and its example is followed, by others. For this to happen, the observer must recognize the rule-following behavior of the innovator. Only some behavior responsive to observing the rule-innovative behavior of another agent will result in transmitting the new rule. This problem has two dimensions. First, what must be observed and imitated is not merely deviation from the established rules—the sensible knave's exploitation of the cooperation of others—but rather behavior conforming to an alternative rule. "Do as I do" in cases of deviation from established rules is crucially ambiguous between these two modes of imitation. This is not a problem of motivation, but rather a problem of interpretation of the "example" set by the observed behavior. Second, if the example is regarded as an example of alternative-rule-following, the pressing problem for the observer is determining what that rule is. What is needed for rule-transmission, on analogy with reproduction of new individuals with mutated genes in the biological case, is that *the same* rule is passed on. This requires that the rule in question be identified. This *problem of identification* bedevils almost all current game theoretic accounts of the evolution of social rules, although the problem is systematically masked by theorists,[20] and it poses a key task for Hayek's theory as well. Imitation may be involved, but we need some reason to think that imitation is a reasonably faithful reproducer of the rules from the rule-introducing member of a group to members.

Finally, since on Hayek's theory rule-innovation begins in small local contexts but, given the nature of spontaneous order, new rules can only influence the social order as a whole if they are practiced widely in the group, he needs an account of how such rule-innovations *spread* through a population. What must be explained is both how the new rules spread and how the changed rules that spread are relatively faithfully reproduced across the larger group population. This magnifies the problems mentioned in the previous paragraph. The mechanism of change must be capable of producing and reproducing social rules in the group with a substantial degree of fidelity, otherwise rule-innovation will only be a cause of noise, disequilibrium, and eventual deterioration of the established rules. Hayek must explain why we can hope that out of the process of individual rule-testing, new regimes of rules *of the group* can emerge.

[20] For a discussion of this problem see Gopal and Janssen 1996 and Sugden 1998a.

B. Does Hayek's Theory Permit Successful Performance of these Tasks?

Hayek must solve the above four problems for both his theory of spontaneous order and his evolutionary theory to succeed. They must be solved in order to make plausible his claim that a spontaneous order is, to some degree, self-maintaining and his claim that sometimes disequilibrating forces establish new "orders of action" rather than merely set off a spiral of disorder, orders of action which can then be tested on the field of group competition where evolutionary forces work. Of course, for his account to succeed, it is not necessary that all rule-innovative activities of individuals result in establishing new group-wide regimes of rules and corresponding social orders. It will surely be the case that some individual "experiments" amount only to knavish deviance without establishing new rules, and some individual rule-innovations will not be taken up by others or not taken up by a sufficiently large sub-population that the social order is materially affected, and some such group-wide rule-innovations may produce overall disorder rather than new forms of social order. Nevertheless, Hayek must be able to show how each of the above problems can be solved and give us some reason to think that the solutions achieved at each level are frequent enough to provide the mechanism of change that the spontaneous order and evolution explanations rely on. If "solutions" are only random and rare, the proposed explanatory schemata fail. They will not be able to account for the actual emergence, existence, and operation of social rules as it promises.

Hayek's understanding of rules and the agents directed by them to some degree promotes solutions to these problems, but it also puts substantial obstacles in the way of solving these problems. First, his understanding of rules as *dispositions* of thought and action is too limited to permit room for the normative dimension of the rules he seeks to explain. Glass has the disposition to shatter when struck sharply, but should a pane of glass fail to shatter upon being struck sharply, no *violation* of a rule has occurred, only a deviation from the expected pattern of glass-shattering behavior. The deviation calls for a revision of our understanding of the disposition, not for change in the behavior of the glass. This familiar point needs no further elaboration here. The consequence, however, is that the notion of dispositions alone cannot provide the conceptual resources for needed to explain normative rules. Equipped only with the notion of dispositions we cannot distinguish deviance from deviations, let alone distinguish conformity to a new rule from mere non-conformity. That is, Hayek's very broad and undiscriminating understanding of rules seems to preclude recognition of their dimension of normativity.

This creates an obstacle to successful performance of the tasks outlined above, however, let us assume Hayek's individuals have the capacity to grasp rules as norms

for their behavior and ask whether Hayek has an answer to the knave's challenge. Here the prospects are a bit brighter. The knave's challenge can be seen as two-fold. (i) Given the distinction between mere non-conformity and conformity to an alternative rules, the individual rule-tester must have some motivation to test alternative rules, rather than merely seek opportunities for advantageous non-conformity; and (ii) this alternative behavior must be recognized by others as alternative-rule-following rather than mere advantage-seeking non-conformity. Hayek's response to the first of these challenges rests on his rejection of the conception of human rational agents from which the knave's challenge seems to emerge. Human beings are, first of all, not rational expected utility maximizers, but rather *rule-following animals,* he insists [Hayek 1973, 11]. Of course, they seek to satisfy their desires and realize their aims, but they always do so within a framework of rules. This ordered structure is important to them because the rules provide a source of *intelligibility* in their social lives, which, presumably, Hayek takes to be of more fundamental to them than marginal gains from exploiting free-rider opportunities.[21] Thus, rules, for Hayek, must not be conceived as obstacles to achieving maximal utility (more or less useful under some circumstances), but preconditions for effective or "successful" pursuit of ends, and, it must be said, preconditions of having or forming meaningful ends in the first place. In view of this importance to the individual of being able to see his own behavior as rule-governed as well as to see his environment as intelligibly ordered, the individual will not be willing to adopt strategies that significantly risk disorder either at a personal or social level. Thus, while within limits the individual will have an incentive to explore adjustments of the rules to enable him better to realize his aims, he will do so with a keen sense of the need for an ordered structure for this pursuit, both for himself and for others.

Hayek also seems to have resources for answering the second part of the knave's challenge. Again, because of the importance of intelligibility and rule-governed order, individuals, we can assume, will be primed to recognize such behavior in others; moreover, Hayek argues, our foundational mimicking skills make is possible for us to recognize such behavior in others. So, we can conclude that if he has resources for explaining norm-grasping capacities of human beings, Hayek can solve problems posed by the knave.

However, we cannot be as sanguine about his ability to solve the problems posed by the need for common social rules. The device of imitation alone cannot be sufficient to establish *social* rules needed to coordinate behavior in a systemic way, because it is not possible unilaterally to manifest such rules which can then be imitated by others. The problem at this point is that Hayek's story is at least

[21] Hayek 1967, 90-91. Hume also seems to have advanced this sort of argument in his reply to the sensible knave [Hume 1998b]; or so, at least, I have argued Postema, 1988.

incomplete. We need a richer account of the capabilities and resources on which individuals can draw to grasp and seek to solve problems of complex social interaction. Hayek, however, is reluctant to do so, because that might seem to require of the agent more awareness of the larger systemic situation, and especially of the complex substratum on which rules depend, than he is willing to allow.

His doctrine of the inaccessibility of the substratum of rules of theory and action puts a far more serious obstacle in the way of successful explanation of the emergence of common social rules. This doctrine creates difficulties at two points. First, it makes it very difficult to see how individual rule-innovators can achieve the distance on the rules that direct their behavior needed to assess them. Second, because this substratum is not only inaccessible to each individual, but also *private*, Hayek is unable to provide a solution to the problem of identification. Hayek allows that there may be overlap of basic rules of thought and conduct among individuals in a group, since they will have encountered largely the same natural environment and since they will have picked up many of the same behavioral routines through primitive mimicking of the behavior of others in their group. These similarities may fund, to some degree, reliable expectations regarding the regular behavior of other. But Hayek insists that these commonalities are very limited, restricted to very commonly occurring circumstances. But novel circumstances and problems of interaction arise constantly and, he believes, we lack the capacities and resources to resolve them spontaneously. This is due, in large part, to the inaccessibility to ourselves, and hence to others, of the vast substratum of experience and knowledge on which our rules are based. Hayek was not forced to this conclusion; indeed, I think he had the resources to explain how a relatively rich "commons of the mind"[22] might develop from the exercise of innate capacities for mimicking and sympathy (as Hume called it) in the thick social environments in which human beings develop, but Hayek refuses to take this route, emphasizing, rather, the inaccessibility and privacy of experience.

It seems, then, Hayek's explanatory project runs aground as a result of two key problems: he cannot account of the normativity of social rules and he cannot solve the problem of identification. Because he cannot explain how those observing the behavior of rule-innovators hit upon *the same* rule, he cannot how rule changes introduced by individuals are reproduced and spread in the group. The forces of change which drive both spontaneous order and the evolutionary process either grind to a halt or offer no hope that the result of individual rule-innovative activity will not lead predominantly to undermining of social order.

But this conclusion may be too hasty. In fact, Hayek offers an explanation of normativity of social rules which may also enable him to explain how the problem of identification is solved in social groups. Normativity, he maintains, is a dimension of

[22]This is Annette Baier's [1997] felicitous phrase.

only some rules of thought and action [Hayek 1973, 43, 74-75]. Normative rules emerge when individual intellects begin to differ in their perceptions or conduct and there is a felt need to reconcile the differences and to teach and enforce the rules. On this view, appreciation of normativity emerges when individuals observe the possibility of deviations and come to appreciate the need to treat them as violations to be corrected. This much, while rough, is not implausible, but then Hayek's thought takes a surprising turn. Because with the emergence of normative rules comes a felt need for reconciliation of differences regarding the rules, Hayek maintains that this task must be assigned to some agent who can resolve the difference (since members of the group cannot do so on their own). They are, he maintains, assigned to "chiefs," judges, and other authorities who *articulate* the rules and *impose* them on the group [Hayek 1973, 43, 45, 77-78]. They express the rules in a form that can then be communicated and explicitly taught and they call upon members of the community to comply with them, backing up their demands with appropriate sanctions. Authorities are not empowered to make any rules they please, he argues, but only to fill gaps in the body of rules already established (in the most primitive instances, presumably, by natural processes). Expectations are shaped and naturally coordinated by implicit common rules and it is the job of authorities to maintain as best they can this structure of coordinated expectations [Hayek 1973, 99-100]. Thus, they are called upon to fit their newly articulated rules into the framework of rules already in place, with a view to the system of rules and the resulting order of actions as a whole it makes possible. Their aim, Hayek insists, is to maintain the proper functioning of this order of actions, and the measure of its proper functioning is that satisfaction of legitimate expectations is optimized [Hayek 1973, 86-87, 99-103, 116].

However plausible this story may be as an account of (a certain form of) common law reasoning on which he models this part of his account, it surely cannot help him solve the problems threatening to undermine his spontaneous-order-cum-evolution account of social rules. For at the point of introducing authorities empowered to manage the system of rules and maintain the order of action, we have left behind all efforts at explaining the emergence of social rules out of spontaneous, impersonal, and unintentional processes. Authorities, as Hayek describes them, impose explicitly articulated rules where the naturally generated rules run into a swamp of diversity and they do so with the proper functioning of the whole system of rules and the social order they produce fully and explicitly in mind. The order, of course, is thought to have no specific goal other than that of coordinating the expectations of members of the group, but that makes their perspective no less systemic and comprehensive, and their efforts to maintain it no less intentional and "planned." Thus, it is fair to say that Hayek's only developed reply to the problems of normativity and identification is one that does not rescue his favored scheme of explanation, but abandons it. Or at the very least, he must concede that *no* social order is entirely spontaneous and the evolution

of social rules is from the beginning assisted by intentional, system-aware and group-oriented agents of innovation.

How might we try to solve the problems of normativity and identification still looming for Hayek? At this point I propose we turn, finally, to Hayek's intellectual friend, Karl Popper, who may have some resources to offer towards a solution to these problems, although for Hayek they may come at a rather high philosophical price.

III. Objectivity, Discursive Capacities, and the Evolution of Social Rules

A. The Objective World and Discursive Reason

Despite his fundamental subjectivism, Hayek seeks to assure us of the objectivity of social rules and the judgments we make on the basis of them. His basic idea is that social rules are objective in the sense that they are reliably connected to the world outside the subject [Feser 2006, 304-6]. They are *connected to* (but not, as far as subjects can tell, *reflective* or *true of* that world) by virtue of being *adapted to* that external world. He is also, from time to time, inclined to infer further that we have good reason to rely on them even if the rationale for them is unavailable to us; we have reason to accept them "uncritically," as Hayek often puts it. We have seen already that he is not entitled to these conclusions, without substantial additional normative premises.

But we might think that objectivity is not a core concern for Hayek. His spontaneous-order-cum-evolution schema is meant to be explanatory. The explanation must be illuminating, but it need not for that purpose assure us of the objectivity of social rules. For Karl Popper, in contrast, the idea of objectivity is central to his explanatory project. Indeed, it is only with the emergence of what he calls the objective world (or "world three") that social norms, and human reason itself, become possible. To understand this strange idea we need to survey briefly the nature and evolution of the province of objectivity.

In addition to the inner world of subjects—the domain of mental states and dispositions ("world two")—and the world of physical objects ("world one"), there is, according to Popper, a domain of intelligibles, of logical and cultural objects that he calls "world three." Its denizens include the contents of thoughts (the objects of thinkings), numbers, theories, conjectures and hypotheses, arguments, problematics and unsolved problems, as well as cultural things with physical (world one) dimensions like tools, buildings, sculptures, plays, symphonies, and, most importantly, language [Popper 1972, 106-7; 1994, 5-6]. This world is *objective* in two

respects. First, it is objective in the sense that, although is its denizens are products of human activity, they do not depend for their continued existence on their makers [Popper 1972, 112]. These objects are not mind-dependent. Popper rejects any form of psychologism or subjectivism that seeks to reduce world three items to items in the consciousness of subjects. World three is a separate domain. It is brought into being and continually added to by human subjects, but beyond that it is not dependent on them. This domain is also objective in the sense that it is *autonomous.* New items in this domain emerge *on their own* from other items in the domain. Arguments may have consequences yet unrecognized by those who entertain them; new problems, new possibilities for argument, and yet unexplored implications and presuppositions of thoughts are regularly generated in this domain [Popper 1972, 117, 159-61; 1994, 19-20, 24-46]. The natural sequence of numbers, Popper maintains in a favorite example, was a human creation, but once taking its place in the objective domain it generated its own problems, which then were available to be discovered by other thinking subjects [Popper 1972, 117]. Thus, this domain is full of unintended and as yet unappreciated consequences of human invention [Popper 1972, 159-60; 1994, 26]. The autonomy of items in this domain and their impact on subjects, for Popper, strongly argues for their independence from subjective mind.

As mentioned, the objects of this world are, directly or indirectly, natural and unintended products of the human activity [Popper 1972, 112]. Other animals also participate in populating this world—beavers construct dams, spiders spin webs, many species use a form of language to express interior states or communicate information [Popper 1972, 115; 1994, 82-3]—but the world three village became a mighty nation after the evolution of human language. Not all the items in this domain are linguistic or admit of linguistic articulation (music, for example), of course, but Popper maintains that language lies at the foundation of the domain and it is the primary medium in which it grows [Popper 1994, 34, 38, 81]. The special and especially fecund genius of human language lies in its two "higher" functions. In addition to expressive and communicative or signaling dimensions shared with animal languages, there evolved with the human species language capacities that enabled individuals describe the world around them. The *descriptive* resources of language funded the possibility of offering descriptions that did not match the world experienced. This spurred two key developments: (i) the development of imaginative capacities of the human mind—subjects could frame thoughts of counterfactual situations and invent stories—and (ii) the development of criteria for assessing descriptions as true or false and other logical operations on descriptions. These function as regulative ideas governing the activity of describing [Popper 1972, 119-20; 1994, 81, 86-7]. And this, in turn, made possible the development of the *critical* or *argumentative* use of language. For with tools for assessing descriptions came critical assessment of proposed descriptions and explanations, and with them came resources for evaluating these criticisms, regulative ideas of validity and soundness of

arguments, relevance of evidence, and the like [Popper 1972, 119-20; 1994, 86-92]. Rational criticism, working on the offerings of imagination, then became the main instrument of growth of knowledge (of world three) and evolution of human species.

The emergence of these dimensions of language, natural products of human activity, gave rise to new human capacities for imaginative and critical use of language and so a practice of conceiving and critically assessing proposals for making their environment and experience intelligible. The three worlds are distinct, but they interact in important ways, with the second world mediating relations between the first and third worlds [Popper 1972, 112, 117, 147; 1994 7, 20-21] Human minds think up solutions to problems created by interaction with the physical world and create physical objects—e.g., tools, buildings, ways of producing food—to solve those problems; in turn developments in world three act on the subjective conditions (mental capacities and dispositions) of its creators. "The human mind evolved together with world three," Popper insists [Popper 1994, 10]. Rationality and the self, our rational capacities and practices, and each individual's sense of self, are developed by engaging in rational criticism. Through participation in rational critical activities made possible by the objective domain self-transcendence is possible [Popper 1972, 147-8; 1994, 130-40]. We come equipped, genetically or through social learning, with dispositions, routines of perception and action, and expectations about the world around us, but by bringing those dispositions, routines, and expectations into the light of the objective world, we can subject them to critical assessment. Thus, we are never wholly prisoners of our local environments or instinctive routines, dispositions or expectations [Popper 1994, 139]. As he is fond of saying, we are able to throw a rope into the air and scramble up it, provided it gets a hold in the world of critical discussion [Popper 1972, 148].

Moreover, because the domain in which rational criticism takes place is not private or subjective, but rather objective and in that sense public, and it always available to rational minds, rational criticism is always capable of being *intersubjective*. Critical argument is, typically and in its primary form, *discoursive*, to use the now obsolete but extremely useful term); that is, it is a matter of offering reasons and arguments to engaged interlocutors. Self-criticism is possible, of course, but it is tutored in a practice of mutual criticism [Popper 1966 vol. 2, 225-7, 238]. World three is the domain where rational subjects meet, subjects whose rationality and "full consciousness" emerge and are nurtured in these interactions, where arguments are explored, new problems discovered, and propositions and proposals ("conjectures") are posted, and then are subjected to rational discoursive criticism.

B. Popper and Hayek on the Evolution of Social Rules

With this brief sketch of Popper's views in hand, let us return to the question of the emergence and dynamics of social rules and the problems Hayek's theory seemed unable to solve. The first thing to note is that both philosophers regard the question of the explanation of social rules from an evolutionary point of view. Also, both maintain that rationality and the human mind evolve with the evolution of the natural and especially social world. Popper's rejection of subjectivism, and championing of an objective and autonomous domain in which rational subjects interact, leads him to a very different view of the nature of the evolution of social rules and institutions. Popper's views stand in stark opposition to some fundamental features of Hayek's conceptual framework. While Popper accepts that much of our knowledge is tacit and dispositional, in much the same sense that Hayek gives this notion, he argues that the resources available to us in the objective domain give us access to this implicit, inborn or socially inbred ("traditional") knowledge and make critical assessment of it possible [Popper 1994, 134-9]. Hayek's mistake, viewed from Popper's point of view, is two-fold. First, Hayek reasons that, since it is impossible to put within any one person's purview (or even that of all of us together) all of what we know implicitly, it is impossible to bring any of it (save a very limited part) to explicit consciousness and the light of critical assessment. Since all cannot be accessed, it cannot be accessed at all, and if not accessed then not assessed. This, Popper argues, is just a mistake. He admits that most of our individual ("subjective") knowledge is implicit, and adds that a very large part of objective knowledge (that which resides in world three) is also not in the command of anyone, and he accepts with Hayek that no one can get command of *all* of either sort of rules or knowledge. Nevertheless, he insists, it is not true that much of it must remain inaccessible to us. He writes,

> While our criticism cannot tackle more than one or two problems or theories at a time . . . there is no problem or theory or prejudice or element in our background knowledge that is immune to being made the object of our critical consideration [Popper 1994, 136].

The view that rational argument must always proceed within a framework of assumptions and thus that there will always be a set of assumptions beyond rational assessment, is, he insists, just a myth ("the myth of the framework"). It is not possible to access all assumptions at once, but no assumption is invulnerable to rational criticism. Second, and following on the first problem, Hayek fails to see that this process of rational criticism is not a lonely or private affair. It is, rather, a matter of bringing bits and pieces of our background assumptions, which define in part the horizons of our expectations, to the objective domain for public inspection and critical assessment. In this way, the private and ineffable is given public form and made available for intersubjective assessment.

Hayek's view, of course, is that it just is not possible to articulate rules that are embedded in practice and dispositions. But Popper does not have to deny this to insist against Hayek that critical assessment of that which can be articulated, in a tentative and piecemeal fashion to be sure, is possible and plays an important role in individual development and social evolution. The feedback loop from critical assessment of the contents of thoughts and theories to the practices and dispositions of the subjects who think them enables us to have a degree of critical control of even that which is not articulable. Moreover, Popper insists, through such "error elimination" human evolution has largely proceeded, since the emergence of language. The possibility of exosomatic testing of solutions to problems has proved to have enormous evolutionary advantages for the human species, he argues, because it "allow theories to die in our stead" [Popper 1994, 12]. Thus, in Popper's view, Hayek ignores one of the most important engines of rule-testing. Critical assessment by *discoursive* interaction in a public domain is a tool of enormous power, a tool Hayek's theory refuses to deploy (except to put it in the hands of authorities who do the work of rule-fashioning for us). But this amounts to a kind of myopia, for there are resources available to human individuals facing serious problems of coordination for solving these problems, including problems of identifying common rules, *jointly* and *discursively*.

Popper offers, in the place of intersubjectivity based on a conception of common knowledge understood as nested subjective or private knowledge (of the sort: I believe that he believes that I believe that…) iterated ad indefinitum, a model of logical space, a public domain or commons, where individuals can meet, engage in deliberation, come to joint solutions to common problems, drawing on common resources. This model puts at the center of the process of rule-formation and rule-transformation the discursive and critical capacities of members of a group, capacities which are systematically left out of typical evolutionary game theoretic models and ignored by Hayek. On Popper's model, the objective domain is a public place in which we can meet in the hope of working out, discursively and critically, the rules we need for cooperation, a place that does not presuppose already shared values, but rather a place of common argument and deliberation, structured, to be sure, by criteria of validity, soundness, and weight of evidence, but a place where these criteria are also subject to critical assessment. This offers also an attractive notion or model of objectivity as well. We might call it "objectivity as publicity."[23]

Gerald J. Postema

Cary C. Boshamer Professor of Philosophy and Professor of Law

University of North Carolina at Chapel Hill

[23] As I have argued in Postema 2000.

REFERENCES

Baier, Annette. 1997. *The Commons of the Mind.* Chicago: Open Court.

Feser, Edward. 2006. "Hayek the Philosopher of Mind". In *The Cambridge Companion to Hayek,* edited by Edward Feser. Cambridge: Cambridge University Press.

Fuller, Lon L. 1969. "Human Interaction and the Law." *American Journal of Jurisprudence* 14: 1-36.

Gaus, Gerald F. 2006. "Hayek on the Evolution of Society and Mind." In *The Cambridge Companion to Hayek,* edited by Edward Feser. Cambridge: Cambridge University Press, 232-58.

Gaus, Gerald F. 2007. "Social Complexity and Evolved Moral Principles." In *Liberalism, Conservatism, and Hayek's Idea of Spontaneous Order*. New York: Palgrave Macmillan, 149-76.

Gopal, Sanjeev and Janssen, Maarten. 1996. "Can We Rationally Learn to Coordinate?" *Theory and Decision* 40: 29-49.

Hayek, Friedrich A. 1960. *The Constitution of Liberty*. London: Routledge.

Hayek, Friedrich A. 1952. *The Counter-Revolution of Science*: *Studies on the Abuse of Reason*. Glencoe, IL: Free Press.

Hayek, Friedrich A. 1967. *Studies in Philosophy, Politics, and Economics*. Chicago: University of Chicago Press.

Hayek, Friedrich A. 1973. *Law, Legislation and Liberty, Volume 1: Rules and Order*. Chicago: University of Chicago Press.

Hayek, Friedrich A. 1976. *Law, Legislation and Liberty, Volume 2: The Mirage of Social Justice*. Chicago: University of Chicago Press.

Heath, Eugene. 1992. "Rules, Function, and the Invisible Hand: An Interpretation of Hayek's Social Theory." *Philosophy of the Social Sciences* 22: 28-45.

Hume, David. 1998a. *A Treatise of Human Nature*. Edited by David Fate Norton and Mary J. Norton. Oxford: Oxford University Press.

Hume, David. 1998b. *An Equiry concerning the Principles of Morals*. Edited by Tom L. Beauchamp, Oxford: Oxford University Press.

Popper, Karl. 1966 *The Open Society and it Enemies*. Princeton: Princeton University Press.

Popper, Karl. 1972. *The Logic of Scientific Discovery*. London: Hutchinson.

Popper, Karl.1979. *Objective Knowledge: An Evolutionary Approach.* Oxford: Oxford University Press.

Popper, Karl. 1989. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Basic Books.

Popper, Karl. 1994. *Knowledge and the Mind-Body Problem*. Edited by M.A. Notturno. London: Routledge.

Postema, Gerald J. 1988. "Hume's Answer to the Sensible Knave." *History of Philosophy Quarterly* 5: 23-40.

Postema, Gerald J. 1998. “Conventions at the Foundations of Law.” In *The New Palgrave Dictionary of Economics and Law*. Edited by Peter Newman. London: Macmillan, vol. 1,465-72.

Postema, Gerald J. 2000. “Objectivity Fit for Law.” In *Objectivity in Morality and Law*. Edited by Brian Leiter.Cambridge: Cambridge University Press, 99-143.

Postema, Gerald J. 2008. “Salience Reasoning,” *Topoi*, 27: 41-55.

Ryle, Gilbert. 1949. *The Concept of Mind*. Chicago: University of Chicago Press.

Skyrms, Brian. 1996. *The Evolution of the Social Contract*. Cambridge: Cambridge University Press.

Skyrms, Brian. 2004. *The Stag Hunt and the Evolution of Social Structure* Cambridge: Cambridge University Press.

Sugden, Robert. 1998a. “The Role of Inductive Reasoning in the Evolution of Conventions.” *Law and Philosophy* 17: 377-410.

Sugden, Robert. 1998b. “Spontaneous Order.” In *The New Palgrave Dictionary of Economics and the Law*. Edited by Peter Newman. London: Macmillan Palgrave, vol. 3?, 485-95.

Vanderschraaf, Peter. 1995. “Convention as Correlated Equilibrium”. *Erkenntnis* 42: 65-87.