The Evolution of Society and Mind: Hayek’s System of Ideas

Gerald F. Gaus

I. INTRODUCTION

As a rule, Hayek has not been treated kindly by scholars. One would expect that a political theorist and economist of his stature would be charitably, if not sympathetically, read by commentators; instead, Hayek often elicits harsh dismissals. This is especially true of his fundamental ideas about the evolution of society and reason. A reader will find influential discussions in which his analysis is described as “dogmatic,” “unsophisticated,” and “crude.” In this chapter I propose to take a fresh start, sketching a sympathetic interpretation of Hayek’s accounts of social evolution and mind as fundamental to his thinking. My basic claim is that Hayek’s views on social evolution and reason are not only intimately bound together, but they also depend on his analyses of complex orders, scientific explanations of such orders, and the place of rules in complex orders. Because so few commentators recognize that his claims about evolution are embedded in a system of ideas, most misunderstand him.

II. THE COMPLEX ORDER OF ACTIONS

Complex Phenomena

Hayek repeatedly refers to “the twin ideas of evolution and spontaneous order.” Although some commentators question whether these ideas are related, Hayek’s insistence on the link between evolutionary analysis and spontaneous orders in writings spanning a number of years indicates that we need to make sense of the “twin ideas thesis” if we are to grasp what he has in mind.

Hayek tells us that attempts to understand human interaction through the ideas of evolution and spontaneous order are the main tools for dealing with complex phenomena. It is, I think, the notion of complex phenomena that is the key. Hayek’s analysis of complexity, especially of social complexity, is built on seven key claims, most of which are part of current analyses of social complexity.

(i) Complex phenomena, according to Hayek, display abstract patterns composed of a large number of variables. Hayek goes so far as to define complexity in terms of a large number of variables. This is perhaps the least adequate feature of his analysis. Although complex systems typically involve a large number of elements, it is the character of the elements’ interactions and the resulting patterns that are fundamental.

(ii) Organized complexity occurs “when the character of the structures showing it depends not only on the properties of the individual elements of which they are composed, and the relative frequency with which they occur, but also on the manner in which the individual elements are connected with each other.” Hayek, then, certainly sees that complex systems arise because of the nature of the interactions of the elements. Importantly, he points to the idea of an emergent property:

The “emergence” of “new” patterns as a result of the increase in the number of elements between which simple relations exist, means that this larger structure as a whole will possess certain general or abstract features which will recur independently of the particular values of the individual data, so long as the general structure (as described, e.g., by an algebraic equation) is preserved. Such “wholes,” defined in terms of certain general properties of their structure, will constitute distinctive objects of explanation for a theory, even though such a theory may be
merely a particular way of fitting together statements about the relation between individual elements.\(^8\)

This is crucial: in analyzing something as a complex phenomenon our concern is the pattern of relations that pertains among the elements. It is the abstract, emergent, pattern that is the crux of complexity. The abstract pattern cannot be predicted from a small sample of the individual elements. Contemporary complexity theorists, for example, see liquidity as an emergent property of a huge number of related water molecules; although liquidity is a property that causally arises out of the interaction of a very large number of individual molecules, the precise properties of waves and ripples could not be predicted from what we know about molecular chemistry, nor does the property appear in a small sample of the molecules.\(^9\) The same emergent property may obtain at times \(t_1\) and \(t_2\) even though none of the individual elements at \(t_1\) persist at \(t_2\). The complex order — the pattern of relations — is this emergent property, not the individual elements.

(iii) Complex systems can be tightly coupled. As Hayek notes, in a complex order the state of the system at any one time depends on a number of factors, and if even one is varied, there may be profound changes throughout the system.\(^10\) The behavior of tightly coupled systems is difficult to predict as they are characterized by error inflation: a small error in predicting one variable can lead to drastic errors in predicting the overall system’s state or, as Hayek would say, its pattern.\(^11\)

(iv) Complex systems are apt to be self-maintaining.\(^12\) By this Hayek meant that such systems have a tendency to persist and to respond to a range of exogenous and endogenous changes.\(^13\) This is fundamental to the idea of a spontaneous order (though not all spontaneous orders are complex).

(v) In many complex systems we cannot measure how close the system is to equilibrium, though we have good grounds to suppose it is never in equilibrium.\(^14\) The most our theories can do is tell us that the system moves towards equilibrium. Thus our theories of equilibrium (say, price theory) will not allow us to reliably predict actual prices.\(^15\) It is important to realize that Hayek accepted the legitimacy of mathematical modeling of the economy; what he dismissed was any claim that we could reliably estimate actual values and so employ our model to generate fine-grained predictions.

(vi) Complex systems such as the economy are characterized by constant novelty.\(^16\) We need to remember that Hayek, like the Austrian School of economics in general, insisted on the importance of dynamic and unknown factors in economic life. “The solution of the economic problem of society ... is always a voyage of exploration into the unknown, an attempt to discover new ways of doing things better than they have been done before.”\(^17\)

(vii) Because of the complexity of the system, there is “no global controller that can exploit all opportunities or interactions....”\(^18\) This brings us to the very heart of Hayek’s economics: global planners could not secure and employ sufficient information to direct individuals to employ capital and labor in an efficient way. Thus Hayek’s position in the socialist calculation debate: “the ‘data’ from which the economic calculus starts are never for the whole society ‘given’ to a single mind which could work out the implications and can never be so given.”\(^19\)

**Explanation of the Principle**
Given the features of complex phenomena, the scientific study of complex orders cannot aim at the prediction of the “specific” future states or values of the individual elements. Hayek realized that the idea of a “specific” prediction is context-dependent; his claim, though, was that in many natural sciences (such as parts of physics), “it will generally be possible to specify all those aspects of the phenomenon in which we are interested with any degree of precision we may need for our purposes.” In contrast, when dealing with complex phenomena we are simply unable to specify the values (in contemporary terms, the system is modeled in nonlinear equations which have no unique solution); we can only predict the “range of phenomena to expect.” We can understand the general principles on which the system operates, and with this knowledge we can predict the parameters within which the system will settle. This is, as Hayek says, an idea of “great importance for the understanding of the theoretical methods of the social sciences.” It is the failure to understand the limits of social prediction that leads to an ill-fated attempt to employ science to engineer society.

**The Emergent Order of Actions**

Hayek subtitles his “Notes on the Evolution of Systems of Rules of Conduct” “The Interplay between Rules of Individual Conduct and the Social Order of Actions.” Given what we already know, we can grasp Hayek’s distinction between specific rules of conduct that individuals follow, and the “social order of actions” — the emergent property — that arises from a system of rules. Hayek says that (1) it is conceivable that the same social order of actions might be produced by entirely different sets of individual rules; and (2) the same set of individual rules may lead to very different social orders, depending on the environment in which it operates. Both points, of course, are fundamental to the idea of a complex phenomenon: it can be produced by different sets of individual elements and the same set of individual elements may produce very different emergent features.

Throughout his long career — and certainly since the 1950s — Hayek’s overriding concern was the analysis of the emergent property he called “the order of actions”:

> It is the resulting overall order of actions but not the regularity of the actions of the separate individuals as such which is important for the preservation of the group; and a certain kind of overall order may in the same manner contribute to the survival of the members of the group whatever the particular rules of individual conduct that bring it about.

Hayek’s fundamental insight is that the survival of a society depends on the emergent property of orderly cooperation of different individuals which has a complex relation to the rules of conduct individuals follow. Thus it is a serious misunderstanding of Hayek to claim, as does one commentator, that

> Hayek’s distinction between the group (in the sense of an order....) and its institutions (such as rules and paradigms) is pointless. The distinction is redundant because the group neither acts as an individual nor operates as a cohesive form. It is merely, given external circumstance, a mirror image of the rules adopted by the individuals.

Although the order of actions arises out of the set of rules and institutions, it is not just the mirror image of them (as we have seen, the same order might arise from an entirely different set). So far from being pointless, the fundamental interest of Hayek’s account...
is that his analysis of complexity allows him to distinguish the pattern (emergent) property from the set of rules which gives rise to it, and so his analysis focuses on the pattern property. This order does not depend on predictably of individual actions, which is one reason why Hayek can advocate a dynamic social order in which individuals are constantly doing new things in new ways (and this is one reason why he is not a conservative).

III. SOCIAL EVOLUTION

Evolutionary Accounts as “in Principle Explanations” of Dynamic and Open Complex Orders

Recall Hayek’s insistence that the twin ideas of a spontaneous order and evolution are the main tools for understanding complexity. Theories of spontaneous order (e.g. economics) — which explain how the complex order of social actions can be self-organizing and self-maintaining — are essentially what Hayek calls “models” of the complex phenomenon of the social order of actions, providing general accounts of how the elements relate. As Hayek remarks, all of economics can be understood as modeling the complex order. But to understand the workings of the social order of actions as a spontaneously organized and self-regulating complex phenomenon is, still, essentially a static explanation. Moreover, Hayek insists that any such model of a spontaneous order is incomplete because the behavior of the order does not simply depend on the internal relations of the system but exogenous (outside) forces; indeed, it may have been shaped by the specific series of environments it passed through. “Though it is reasonable to believe that structures of the kind will in a definable environment always behave as they do, the existence of such structures may in fact depend not only on that environment, but also on a definite sequence of such environments ....”

Hayek is so attracted to evolutionary accounts of the order of actions because they hold out the promise of providing “in principle explanations” of the alteration and development of complex orders without supposing anyone fully understands the working of the order. As Hayek sees it, evolutionary accounts provide the real alternative to design theories, and they articulate precisely the “explanations of principle” that are appropriate to complexity. In biology, Darwinian theory allows us to understand the principles that regulate the development of species, shows us that some developments are outside the possible range of values (e.g. that horses will suddenly give birth to winged offspring), but it is unable to generate specific predictions about the future of individuals or species. Hayek’s analysis was path-breaking here, showing how evolution is a case of complexity theory. Compare a recent writer on evolution:

Algorithms must always produce the same result if they start from the same point. This seems to suggest that, if evolution follows an algorithm, its results must be predetermined and predictable. This is not the case and chaos theory explains why not. There are many simple processes, like dripping taps or moving gasses, or the path drawn by a swinging pendulum, which are chaotic. They follow simple and mindless algorithms but their end results are complex, chaotic and unpredictable. Beautiful shapes and patterns can emerge, but although the kind of pattern may be repeatable, the detail cannot be predicted without running the procedure right through. And since chaotic systems can be highly sensitive to initial starting
conditions, a tiny difference at the beginning may lead to an entirely different outcome. Evolution is like this.\textsuperscript{35}

\textit{The Core Idea}

Before we get into details, complexities, and problems, it will be helpful to get clear about the outline of Hayek’s account of social evolution. We have seen that the \textit{explanandum} (that which is to be explained) is the rise and development of an emergent property, viz., the social order of actions. As Hayek says, “the selection process will operate on the order as a whole.”\textsuperscript{36} This is the “Great Society”: an overall spontaneous order of adaptations that allows for coordinated action.\textsuperscript{37} The \textit{explanans} (that which does the explaining) is an evolutionary account whereby the rules and institutions that give rise to this order (i.e., this emergent property) are selected via a competition (“in the widest sense”)\textsuperscript{38} among social orders. The emergent property, we have seen, arises out of a system of rules; therefore the competition among these social orders is determined by their constituent rules and institutions as they operate in specific environments. Social orders of actions are typically differentiated by their constituent rules and institutions; variation in the rules and institutions can provide a competitive advantage in the competition between social orders, leading to selection of a social order of actions with certain sets of rules.

There is a rough and ready analogy here with one understanding of Darwinian evolution. On the face of it (remember, complications will come later) in Hayek’s account rules play a role analogous to genes in biological evolution; whereas individual organisms are constituted by following the instructions of genes, a Great Society is constituted by following the instructions of rules.\textsuperscript{39} And just as genetic variation can give competitive advantage to an individual organism in its competition with others, rule variation can perform the same role in competition between social orders of actions.

Like evolutionary explanations in biology, it is easy to confuse Hayek’s evolutionary theory with a functionalist account.\textsuperscript{40} In a functionalist biological explanation, the presence of, say, the heart, is explained by its function in the overall system — the role of the heart in keeping animals alive. A functionalist may claim that the heart exists because it is needed to pump blood. Thus it may seem that when Hayek says that practices have been maintained “because they enabled the group in which they have arisen to prevail over others,”\textsuperscript{41} he is saying that the practices are there because they are needed for the group to prevail. Consequently, many suppose that Hayek’s \textit{explanandum} is why we have the specific rules we have, and his \textit{explanans} is that they are needed for the complex order. This, I think, gets things precisely backwards. Evolutionary theories are not functionalist: they are causal explanations of the development of an organism. However, once the evolutionary account has been given, over a wide range of traits, we will be able to see that the trait performs a function. Hearts do indeed perform a function. But in the evolutionary theory the existence of the heart is not explained in terms of its function, but in terms of the history of the organism and the competitive pressures that selected those with certain genes that developed into hearts. Similarly for Hayek. Once he has provided an evolutionary account of the rise of the social order of actions, he then has grounds for saying that the rules and institutions we find in this order generally have served a function that has given the order an advantage in the environment in which evolutionary selection occurred.
Units and Mechanisms of Selection
If this core idea is to be filled out in a convincing way, Hayek must specify and explain two crucial mechanisms: the selection and replicator mechanisms. Let us first consider selection mechanism. It would seem that any evolutionary explanation of the rise of X_α (entity X with feature α) must be able to identify at least one mechanism according to which in some past environment E, mechanism M selected X_α over some X_β. Hayek is often criticized for not identifying “the” selection mechanism, but an adequate evolutionary theory need not rely on a single mechanism. Darwin himself proposed both natural and sexual selection as mechanisms, and these apparently can work in opposite directions. Biologists are still debating whether sexual selection can produce traits such as male bird feathers that do not render the male fitter to survive in its environment and may actually result in traits that make the male less fit to survive. As Hayek often reminds us, though, we ought not to take Darwinianism as the template for all evolutionary accounts. For Darwin, one selection mechanism involves differential survival rates: those Xs with α have a higher probability of survival than those with β. But a non-Darwinian evolutionary account also could be based on, say, differential growth rates. Suppose we begin in E with a limited amount of space, and five entities (X_1-X_5) that cannot reproduce or die, but can grow to some maximum or shrink down to some minimum. We could still have an evolutionary explanation of why X_1 takes over the space, even though there is no difference in reproduction or survival rates — evolutionary accounts thus need not be based on natural selection.

I stress this because even careful commentators are apt to suppose that either Hayek’s evolutionary account must closely parallel Darwin’s, or else it is simply metaphorical or confused. We must be careful to consider Hayek’s selection mechanisms on their own terms. Four possible selection mechanisms can be identified.

1. Group survival. Although Hayek generally insisted on the distance between his theory of “cultural” evolution and Darwin’s account of biological evolution, he also stressed that both rely on competition for survival. Of course the difference — and why Hayek is not a Social Darwinist — is that the competition is not between individuals, in which “fitter” individuals survive, but between social orders of actions. Unfortunately, Hayek often employs the more accessible notion of competition between “groups,” in which one group prevails over another. He writes:

   The rules of conduct have … evolved because the groups who practiced them were more successful and displaced others. They were rules which, given the environment in which men lived, secured that a greater number of the groups or individuals who practiced them would survive.

This passage presents problems. By simply talking of “groups,” and then adding at the close that “greater number of the groups or individuals” survive, Hayek appears to be directly falling prey to problems of group selection and collective action. It looks as if Hayek is claiming that an individual’s chance of survival is maximized if she belongs to a group that maximizes its own chance of survival. But this raises familiar problems of collective action and the rationality of free-riding. If individuals are confronting prisoner’s dilemmas, it may be rational for each person not to do that which is good for the group, even though this leads to a situation which is disadvantageous for all. One of the lessons of game theory is that what is good for the group may not be rational for anyone.
Space does not allow us to go deeply into these issues; even simply in terms of rational choice narrowly understood, they are much more complex than the above suggests. In relation to Hayek, however, we must keep two points in mind. First, Hayek’s concern is not simply groups, but “the order of actions of a group.” This means that our units are systems of cooperation — arising out of a system of rules — and for Hayek this means that the rules actually regulate people’s actions. Secondly, as we shall see in section III, Hayek repeatedly insists that our reason is itself evolved, and is itself shaped by the order of actions. This does not mean simply that the degree to which people are moved by reason is shaped by the order of actions. When Hayek tells us that “individual reason is a product of inter-individual relationships,” he is arguing that our conception of rationality is produced by social life. Mind is a product of the social order in which it has evolved. Successful orders of action evolve conceptions of reason that induce general rule following.

It is not clear in what sense Hayek advocated “group selection.” The idea of group selection is not pellucid, but we probably should not understand an evolutionary theory to raise the problems of “group selection” just because the theory selects among groups. Group selection is controversial in biology because it refers to selection among group traits that leads to the selection among individual traits. A typical group selection story would be that because groups of altruists do better than groups of selfish people, individuals in those groups do better than individuals in selfish groups, and this explains why these surviving individuals have an altruistic trait. The idea is that individual characteristics are explained by group membership. Sometimes Hayek seems to uphold group selection in this sense:

Although the existence and preservation of the order of actions of a group can be accounted for only from the rules of conduct which individuals obey, these rules of conduct have developed because the individuals have been living in groups whose structures have gradually changed. In other words, the properties of the individuals which are significant for the existence and preservation of the group, and through this also for the existence and preservation of the individuals themselves, have been shaped by the selection of those individuals from the individuals living in groups which at each stage of evolution of the group tended to act according to such rules as made the group more efficient.

So it looks as if Hayek’s claim is that group selection leads to the selection of individual traits — a full-fledged group selection account. I think, though, that this is misleading. Group selection accounts in biology claim that, because it is better for an individual to be a member of an altruistic group, an individual who has the altruistic trait will do better, and that is why the individual has the trait. And this is what raises the prisoner’s dilemma problem: it is even better to be a non-altruistic member of an altruistic group.

Hayek’s claim, though, is different. As I have mentioned — and more on this anon — Hayek has a general theory that reason is significantly shaped by culture. Because of this, if it is advantageous for the group to be characterized by a way of reasoning, this will be instilled into the individual by the culture: the option of reasoning in an individually advantageous but uncooperative way is undermined, at least to some extent, by the account of the social roots of reason.

2. Group growth. Hayek tells us that comparative increase of wealth and population are means of evolutionary selection. An evolutionary account of customs, he holds, must
show the “distinct advantages by those groups that kept to such customs, thereby enabling them to expand more rapidly than others and ultimately to supersede (or absorb) those not possessing similar customs.” However, this is not really a different mechanism than survival — these are specific traits that tend toward survival. There is certainly a good case for wealth and population growth being understood as traits that are conducive to the survival of social orders (just as good nutrition is a trait that is conducive to the survival of an organism). Those that are wealthier will be able to expand, fend off external threats, and their culture may come to dominate others. We also would expect them to attract immigration: Hayek suggests that this attractor trait is probably more important than internal population growth.

These claims have not been well received (many take them to be an unfortunate feature of the *Fatal Conceit*, a book completed during Hayek’s final illness, and some of which seems to reflect the views of his editor, Bartley, who finished the manuscript). The analysis, though, is not really implausible, whereas many of the critiques are. For example, it is sometimes offered as a criticism of Hayek that the liberal western world has much lower birth rates than many less developed countries, and China, an authoritarian state, is the most populous country. But these criticisms miss the mark.

Selection mechanisms select by giving an advantage to some trait over a time span. Evolutionary selection is not subject to particular counter-examples. No doubt there were cases in which Neanderthals beat out *Homo Sapiens* for survival: this would not show that we were not more fit in the evolutionary environment in which we competed. So too we may find that because of drastically reduced death rates combined with traditional birth rates, some countries may experience a population explosion. The question is whether, over the long term — and remember, Hayek thought that the most important achievements of cultural evolution occurred before recorded history — certain sorts of cooperative orders are better able to sustain larger populations, and whether this gives them an advantage that leads to their greater survival. Given the collapse in recent times of Soviet-style command economies, which were manifestly unable to produce sufficient wealth to support their populations, Hayek’s idea should not be lightly dismissed.

3. *An endogenous mechanism.* Reflect again on the rough analogy according to which rules are to orders of actions as genes are to organisms. Keeping that in mind, we need to be clear whether our aim is to explain how (a) a specific gene evolved or (b) how a trait of the organism evolved. It might seem that these are just different sides of the same explanation: if we explain how $X_\alpha$ (an organism with a specific trait) was selected by an evolutionary mechanism, we must explain how the $\alpha$-gene was selected. The evolutionary account of how $X_\alpha$ arose must *ipso facto* be the evolutionary story of how $\alpha$-gene arose. Interestingly, things are not so simple. We might have some story about the selection of $\alpha$-genes that did not entirely depend on a story about, say, the competitive advantage of $X_\alpha$. Suppose that there was a competition between genes that was not determined by the competition of $X_\alpha$ with its rival $X_\beta$. Imagine that we are biological organisms of a somewhat different type. We still have genes that determine our traits, and natural selection still works in the familiar way. But we also have homunculi floating around in us. These homunculi have their own aims and, importantly, to survive a gene must receive a general endorsement from the homunculi.
population. Unless the homunculi endorse the gene, the gene dies out, and so the organism (us) loses a trait. Now we have genes that are pressured in two directions. They are pressured in natural selection: they will be maintained in the gene pool only so long as no competitor pops us that confers a survival advantage to the organism. On the other hand, they will only be maintained if they garner endorsement from the homunculi. Of course what the homunculi want, and what causes them to endorse a gene, will be crucial in filling out the story (we would expect that they are not too prone to reject genes that give rise to traits that have a competitive advantage in the competition between organisms).

Such an evolutionary account would be complicated, but it is by no means confused, and it is certainly not crude. True, sometimes the “unit of selection” will be the organism’s traits and sometimes the genes. Now there is reason to think that Hayek’s account of social evolution is similar to this more complex picture. Hayek writes that a person’s “thinking and acting are governed by rules which have by a process of selection been evolved in the society in which he lives.” Understood thus, it looks now as if his project is to explain how each rule (not the order of actions itself) evolved within the society. That this project may rely not only to the rule’s ability to produce a competitive order of actions, but the rule attractiveness to individuals, is suggested by Hayek’s remark that “[t]he competition on which the process of selection rests must be understood in the widest sense. It involves competition between organized and unorganized groups no less than competition between individuals.” This stress on individual competition and the evolution of rules suggests that, instead of a competition between social orders, Hayek has in mind a competition between individuals within a social order that leads to the selection and evolution of rules. So we seem to have two evolutionary competitions, pressuring rules from two different directions. (i) A competition exists between social orders which are, as it were, carriers of rules (as individual organisms are carriers of genes). The rules give a social order a certain competitive advantage, but the rules are only selected insofar as they are part of the evolved social order of actions. (b) There is also a competition between individuals and groups, and this competition selects certain rules as conducive to individual/group success.

If we allow ourselves to forget biology for a moment, the basic idea seems sound. Orders of actions are emergent properties constituted by a system of rules. What rules exist in an order of action is determined both by how well that order of action fares in competition with other orders and the ability of the rules to garner support from the individuals who follow them and are competing with other individuals within the society. So rules will have selection pressures from two very different directions. I think this is probably Hayek’s settled view. Admittedly, he typically insists on the competition between social orders account, though in the Fatal Conceit he does say that “cultural evolution operates largely through group selection,” suggesting a non-group mechanism. In The Constitution of Liberty he writes:

[II]t is, in fact, desirable that the rules should be observed only in most instances and that the individual should be able to transgress them when it seems to him worthwhile to incur the odium this will cause. ... It is this flexibility of voluntary rules which in the field of morals makes gradual evolution and spontaneous growth possible, which allows further modifications and improvements.
And in the Epilogue to *Law, Legislation and Liberty*, he argues that the steps in cultural evolution toward large scale coordination “were made possible by some individuals breaking some traditional rules and practising new forms of conduct — not because they understood them to be better, but because the groups which acted on them prospered more and grew.” There are three ways this individual-initiated change might be understood:

1. It might be that the development of new practices by individuals is aimed at a better order of actions. This is the interpretation Hayek explicitly rejects, and given his analysis of complexity, he is right to do so.

2. New rules might be conceived, a la Darwinian evolution, as random mutations — simply a source of variation and not involving another selection mechanism at all. Hayek’s insistence towards the end of his career on group selection suggests this. But this looks somewhat implausible. We would expect, for example, that there will be some mechanisms that make some rules more likely to catch on in a group than others (ones that are easier to learn, ones that allow the individual more room to satisfy his own interests, ones that serve individual interests in their competition with others). It is not simply that groups with rules such as “don’t satisfy your self-interest” lose out to groups that have competing rules, but that such rules have a hard time getting the support of enough homunculi.

3. So the account might claim that though individuals do not innovate, or imitate the innovations of others, because they have a grasp about how to improve the overall order of actions, the conditions under which rules loose or gain the support of the constituent individuals are non-random and relate to local perceived individual benefits. Rules that allow individuals to locally satisfy their interests are apt to catch on, and those that are parts of badly-performing orders are apt to be violated. This would add a useful non-Darwinian element to Hayek’s account: the occurrence of “mutations” (i.e., new rules) would be non-random: they occur more often when they are needed. (When things are going badly for people, they break the old rules more often and try new arrangements.) This makes the wealth criterion more plausible: Hayek does not have to say that the less wealthy societies are simply overwhelmed by the wealthier, but that as societies lose their ability to produce wealth, defections from existing rules increase, leading either to swift changes into other orders or collapse.

**Rules as Replicating Instructions**

Let us return again to Darwinism to orient ourselves. We have seen that an adequate evolutionary account must identify units of selection and selection mechanisms. In addition, a Darwinian account must include a high fidelity replicator — in contemporary biology, the gene. To say the gene is a replicator is to say that it copies itself; it has high fidelity because its copying is highly accurate. However, it is not perfect: its imperfection is a source of variation in genes — the faulty copies are what we call “mutations.” Darwinian evolution is based on differential reproduction, and so copying is crucial. This is a part of Hayek’s account of social evolution: he details how norms of trade spread in the ancient world, and this included colonization: a sort of copying of rules from one society to another. This, though, is not at all central. What is really important is that the rules have high fidelity replication insofar as they must be accurately “copied” by each individual subject to them. Cultural evolution, says Hayek,
“simulates” Lamarckian evolution because acquired characteristics — rules and institutions — are transmitted from earlier to later generations. This is done, he indicates, through the ability to imitate.

The worry is that the imitation of rule following behavior does not possess high fidelity. Imitating the results of another’s actions — his output, as it were — is a low fidelity way of copying. If each generation simply observes the previous generation’s behavior, and infers from this what the rule is, we would expect large drifts in the interpretation of rule over many generations. Imagine a game of “do what I do” in which there is a long line of people. The first person performs some relatively complex action, and the second person copies her, the third person copies the second, and so on. By the end of a long line we would expect a very different performance, because of the low fidelity of the replicating method.

It is hard to underestimate the importance of this to an evolutionary account. If from generation to generation there is a great deal of random drift in how the rule is interpreted, then even if we grant everything Hayek tells us about the selection mechanism, it will be of no avail, for the advantages of adopting the rule will be lost by random changes. It is no good to adopt a good rule that quickly mutates into a bad rule. Now language greatly improves fidelity: instead of simply copying the results, we copy a set of instructions about to do. (Susan Blackmore compares trying to copy the soup of a good cook with copying down his recipe: we expect much higher fidelity from the latter.) The apparent problem for Hayek is that he insists that the rules which guide the members of a society are largely unconscious. A rule, Hayek says, is a “propensity to act.” Customs and habits, which Hayek repeatedly sees as central to an order of actions, are “unconscious rules.”

That rules in this sense exist and operate without being explicitly known to those who obey them applies...to many of the rules which govern the actions men and thereby determine a spontaneous social order. Man certainly does not know all the rules which guide his actions in the sense that he is able to state them in words. Hayek, then, cannot avail himself of the fidelity of language, since so many rules are not conscious. How, then, can he ensure high fidelity?

IV. MIND AND EVOLUTION

Neural Network Models

To understand Hayek’s sophisticated solution to this problem, we need to reflect on his theory of the mind. Hayek is recognized today as an early neural network modeler. Neural network theory, or connectionism, can be seen as a development of the associationist psychologies of Locke, Hume, and James Mill according to which thoughts are connected by the laws of association (such as similarity, continuity, and so on). The crux of associationism has been summed up thus:

Events that occur in space or time become connected in the mind. Events that shared meaning or physical similarity become associated in the mind. Activation of one unit activates others to which it is linked, the degree of activation depending on the strength of association. This approach held great intuitive appeal for investigators of the mind because it seems to capture the flavor of cognitive behaviors: When thinking, reasoning, or musing, one thought reminds us of others.

Unlike traditional associationism, connectionism abjures any appeal to primitive qualitative differences between sensations or thoughts. Connectionist systems are
composed purely of neurons that simply have on/off states: qualitative differences (e.g., between thoughts) are the results of a complex pattern of neural activation. As one contemporary cognitive psychologist describes them:

Each unit [“neuron”] receives “activity,” both excitatory and inhibitory, as input, and then transmits activity to other units according to some function (usually nonlinear) of the inputs. The behavior of the network as a whole is determined by the initial state of activation and the connections between the units. The inputs of the network also gradually change the “weights” of the connections between the units according to some learning rule. The units have no memory in themselves, but earlier inputs are represented indirectly via the changes in the weights they have caused.

The Relevance of Hayek’s Neural Network Model

Our concern here is not the adequacy of neural network models in relation to their competitors, but the way that Hayek’s early (and, I think, path-breaking) neural network account meshes with his overall social theory. Four features of his connectionist theory of mind are relevant here.

1. Mind, classification and rules. The fundamental aim of Hayek’s neural network theory is to explain “the kind of process by which a given physical situation is transformed into a certain phenomenal picture.” A certain state of the external world $W$ exists at time $t$: how is $W_t$ transformed into a sensory experience $S$ of $W_t$, and how does $S[W_t]$ relate to sensory experiences of other states of the world, and when will these be perceived as the same, and when will the sensation be different? The key to Hayek’s analysis is “classification” via of neuronal connections, “a process of channeling, or switching, or ‘gating,’ of the nervous impulses so as to produce a particular disposition or set.”

By “classification” we shall mean a process in which on each occasion on which a certain recurring event happens it produces the same specific effect, and where the effects produced by any one kind of such events may be either the same or different from those which any other kind of event produces in a similar manner. All the different events which whenever they occur produce the same effect will be said to be events in the same class, and the fact that every one of them produces the same effect will be the sole criterion which makes them members of the same class.

Thus two events are the same just in case they trigger the same neuronal configuration. The central nervous system, then, takes what we might think of as an undifferentiated world and, via the connections in the neuronal network, creates a formal structure of classes of sensations. But this is to say that the mind is rule governed: the neuronal connections constitute perceptions of patterns. In a fundamental sense the mind is a set of rules that takes sensory inputs and yields perceptions. It is important to understand here that Hayek refuses to identify mind and consciousness. In fact — and this gives a contemporary flavor to his theory — Hayek advances a somewhat deflationary theory of consciousness. Consciousness — “higher mental processes” — operates on the same connectionist principles as unconscious or pre-conscious mentality; both are defined by the relevant neural networks.

Some might wonder how this analysis of the formation of sensory experience (inputs) could be relevant to action (outputs). (Our concern, it will be remembered, is learning action-guiding rules.) Hayek has no real problems here as his theory denies any
fundamental distinction between the neural basis of sensation and output or “motor behavior.” Action also follows from neural connections and so human action is inherently rule-based: it is regulated by a network that is based on classification of types of sensation and how they relate to types of responses. Practical rules identify “patterns of actions” that are classified as having the same meaning; the activation of such rules disposes the agent to act. And, again, because so much mental life is unconscious, we can see how Hayek is led to his famous claim of the fundamental importance of unconscious rule following (sec. III above): rule following involves dispositions to act. It is mistaken, however, to see this as evidence of behaviorism. As Hayek points out, although his theory concurs with the behaviorist claim that psychology must not focus (solely) on the conscious, unlike behaviorism, he makes no effort to avoid the mental: quite the contrary — his entire theory of rule following is based on an analysis of the mental, albeit a conception of the mental that has a large role for the unconscious. To learn a social rule, then, is to form a neural network uniting a pattern of inputs with a pattern of output behaviors.

2. Learning. Neural network theories, including Hayek’s, place great stress on learning: the connections and attendant weights that form the network are shaped and reinforced by the environment— the stimuli that the agent encounters. The mind can be understood as a map of the world. The particular environment in which the mind has been formed shapes the neural connections that form the map. Thus we can say that “the apparatus by which we learn about the external world is itself the product of a kind of experience.” This sheds light on Hayek’s often-quoted remark that reason not only shapes culture, but is shaped by it. To be a reasoning creature is to have neural networks of a certain complexity; our environment is crucial in shaping the networks, forging pathways through repeated experiences that give rise to a pattern of connections. Given this, we can see how social rules have high, but not perfect fidelity. They have high fidelity because we are not simply imitating the behavior of others, trying to copy their products (recall the discussion from sec. III). Rather, our similar environment impresses on us similar maps of the world, which include similar perceptions of instances as “the same” and similar types of actions of as “the same” response. In a fundamental sense, we have similar sets of instructions — neural networks. However, these are by no means identical:

The different maps which will thus be formed in different brains will be determined by factors which are sufficiently similar to make those maps also similar to each other. But they will not be identical. Complete identity of maps would presuppose not only an identical history of the different individuals but also complete identity in their anatomical structure. The mere fact that for each individual the map will be subject to constant changes practically precludes the possibility that at any moment the maps of any two individual will be completely identical.

We can see how, on such an account, the evolution of the order of actions, which can yield new rules, can lead to corresponding changes in mind. Thus “cultural selection…creates reason.” The evolution of society is the driving force behind the evolution of mind. This is not to say that the mind is passive: we have seen how individuals are sources of variation and changes in rules. We might say that the development of the mind “feeds back” into the further evolution of culture. It is hard to see in Hayek’s writings, however, a case for the mind as evolving independently from
culture (i.e., via independent selection mechanisms). Thus “the brain is an organ enabling us to absorb...culture.”

3. Decentralization. It should be obvious that neural network theories are decentralized accounts of the mind: the mind is the entirety of neural networks — the complex of relations. It is important that Hayek conceives of intelligence itself as decentralized in a complex network. Consequently, his famous analysis of the rules of society as possessing “much more ‘intelligence’” than do people’s thoughts about their world, is entirely at one with his analysis of reason. Reasoned mental life itself is not subject to a central controller but is dispersed throughout a system of rules.

4. Complexity again. The mind itself is another example of a complex phenomenon, about which we can only give explanations in principle. Again, we see the importance of the theory of complex phenomena, which informs not only his account of society, but of the mind itself. Hayek advances an additional argument that the mind can never fully understand itself. Any modeling of an X, Hayek argues, must employ neural networks that are not a part of X: the model operates on X by employing additional mental processes that model X. But this means that our mind can never form a complete model of itself, since any model must employ mental processes outside the thing modeled, which we cannot do in respect to our own mind. Moreover, because Hayek advocates a mental holism, in which mental states depend on the total activated neural networks, we cannot even form an adequate model of part of the mind. Thus he concludes that “the whole idea of the mind explaining itself is a logical contradiction — nonsense in the literal meaning of the word.” However, this limit on self-understanding would not preclude us from building artificial intelligence systems as complex, or even more complex, than our own mind, since there would be no self-subsuming modeling involved.

V. CONCLUSION: MISUNDERSTANDING AND UNDERSTANDING HAYEK
I have tried to show how Hayek’s work is a sophisticated (and was often path-breaking) system of ideas involving complexity, prediction, evolution, and the nature of the mind. In light of these, I think we can see how the standard criticisms of Hayek’s theory of social evolution are misconceived.

(i) It is sometimes claimed that Hayek’s thought is, at bottom, contradictory: he insists on our ignorance of social processes but, out of his analysis of ignorance, he generates prescriptions about what we should do. If we know enough to say why socialism won’t work, then we must have good enough insights into the economic order to intervene to promote social goals. This, though, is wrong. As we have seen, Hayek believes that we can know quite a lot about the principles on which complex orders operate, and this theoretical knowledge allows us to say that some system states cannot be achieved and that some ways of organizing social cooperation are more efficient than others; we also know that we are unable to predict the course of, or control, the complex order itself. Thus the analysis of complexity provides sound reasons against planning — seeking to control the emergent order. There is nothing contradictory about a mix of knowledge of principle and ignorance of what is a good plan or optimal policy.

(ii) Others believe that Hayek must be a rigid traditionalist, since given his evolutionary account no one has grounds for objecting to the rules of one’s own
We have seen that Hayek certainly provides grounds for taking the existing norms of society seriously. This is not only because they are been selected in the competition between orders of actions, but because they form a network of actions producing an order that we cannot fully understand. Hayek’s main claims, though, are (a) that we cannot devise a reasonable plan to reform our rules in order to develop a better order of actions and, (b) because we cannot fully understand our system of rules, we will be largely in the dark about the overall effects of any change, so we often act in ignorance of the most important consequences. Complexity theory teaches us that we do not have good grounds for fine-grained predictions. The analysis of the economy as a complex phenomenon generates a strong presumption against claims of the expediency of change. None of this, however, shows that individuals must display rigid adherence to existing rules. Hayek’s endogenous selection mechanism requires that individuals sometimes abandon rules and do things in new ways. Although he is much more suspicious of centralized coercive changes of rules (e.g. legislation), there is no theoretical reason why changes in response to proximate (i.e. non-systemic) concerns should be barred.

(iii) More subtly, others question whether there is any good reason to take the outcome of social evolution as having any normative importance for us. Even if it is a fact that our society evolved such rules in competition with other social orders, it does not follow that they are morally good rules. In evaluating this important criticism we must remember that on Hayek’s view cultural evolution shapes our understanding of reason, morality, and our values. It is not as if we have access to an Archimedean perspective from which we can evaluate in toto our evolved morality. Mind cannot stand outside of the order of actions. It is in this sense that we cannot justify our entire morality. We can employ part of our morality to criticize other elements, but this is not to stand back and criticize the outcome of evolution — it is to employ the outcome. Moreover, given Hayek’s holist account of the mind and the order of actions, any proposal to give one element of our traditional morality absolute supremacy and so remake the rest of society on the basis of it alone must involve misunderstanding the rule, whose importance and meaning depend on being embedded in an overall pattern of rules and actions.

My aim has not been to correct all the misunderstandings of Hayek’s account of social evolution — that could become one’s life work. Rather, I have tried to show in this chapter how Hayek offers a system of sophisticated and complex analyses; because the theories of complexity, spontaneous order, evolution, mind, and rule following form their own complex pattern, commentators are apt to focus on just one or two elements which, not too surprisingly, they find inadequate. It is only when we appreciate the genius of Hayek’s linking of complexity theory, spontaneous ordering, social evolution, and neural networks into an overall account of mind and human society that we will be, finally, in a position to see the true difficulties of his system of ideas, and move beyond, by building on, his great work.

Endnotes

1 John Gray (1984) recognizes the systematic character of Hayek’s thought.

3 For doubts about their connection see Paul, 1988; Hodgson, 1993, 177ff. Vanberg (1994, 78) agrees that the two ideas are intimately related. See also Kukathas, 1990, 88ff.

4 Hayek, 1978, 146.

5 Caldwell (2004, 363) correctly argues that Hayek never developed a full-fledged theory of complexity. Hayek’s writings did, however, display many of the ideas that were later crystallized into complexity theory. J. Barkley Rosser, Jr.’s (1995, 185n) observation seems more accurate: “Hayek...was an early and independent developer of complexity theory in something resembling its current form, albeit without computers.”


7 Hayek, 1978a, 26-27.


9 Waldrop, 1992, 81-83.

10 Hayek, 1967b, 24ff.

11 This is the so-called “butterfly” effect. See Smith, 1998, 16.

12 Hayek, 1967b, 27.

13 Hayek, 1973, 63.

14 Hayek, 1991a, 258.

15 Hayek, 1978a, 27. This is identified as an element of contemporary economic analyses of complexity by Rosser, 1999, 176.


17 Hayek, 1991b, 274.

18 Rosser, 1999, 176.

19 Hayek, 1991a, 247-48. As Kukathas (1990, 57) points out, in comparison to von Mises, Hayek’s critique of socialism is not so much that in the absence of a market prices could not be calculated but that in the absence of a market the necessary information to determine prices could not be collected.

20 Hayek, 1967a, 8.


23 Hayek, 1955, 43.

24 Hayek, 1967c.

25 Hayek, 1967c, 68.

26 Ibid., 68.

27 Khalil, 1996, 194.

28 I put aside here the debate among Hayek scholars whether this is consistent with Hayek’s methodological individualism. See Vanberg, 1994, ch. 5; Hodgson, 1993, 156ff; Witt, 1994, 185; Gray, 1984, 52-53; Khalil, 1996, 191ff. I think Caldwell (2004, 281ff) is right that Hayek’s version of methodological individualism is complicated: given his theory of complexity, the properties of wholes cannot be reduced to the properties of individuals, though those properties result from individuals in relations.

29 Hayek, 1967a, 14ff.
I think it is a mistake to distinguish the market order from the social order of actions; as Hayek suggests, all cooperation is economic. Hayek, 1960, 35. Cf. Hodgson, 1994, 176.

31 Hayek, 1967c, 74-75.
33 Hayek, 1960, 59.
34 Hayek, 1967b, 31ff.
35 Blackmore, 1999, 12.
36 Hayek, 1967c, 71.
37 Hayek, 1973, 2ff.
38 Hayek, 1960, 37.
39 For doubts, see Hodgson, 1994, 164ff. I believe that Hodgson’s doubts are based on his misunderstanding of Hayek’s account of rationality, which I take up in section IV below.
40 On Hayek’s supposed functionalism, see Vanberg, 1994, 84; Hodgson, 1994, 168, 171; see also Gray, 1984, 44ff, 137ff. Hayek’s analysis of complex phenomena, which employs systems theory, may strike some as functionalist. See Hayek 1967a, 20; Hayek, 1988, 28.
44 Hayek, 1988, 23ff. For interpretations of Hayek as a Social Darwinist, see Paul, 1988; Miller, 1989; those who reject this view include Whitman, 1988; Hodgson, 1994, 161-162.
48 Hayek, 1967c, 72.

A crucial question is whether the prisoner’s dilemma is really the proper model for understanding cooperation. There is also a question about whether we should model individuals in terms of their best (self-interested) move in response to others or, as Hayek himself suggests, whether we should suppose that individuals imitate the successful. The analysis of the evolution of cooperation is dependent on these issues. See Skyrms, 1996; Skyrms, 2004.

50 Hayek, 1976c, 72.
51 See Hayek, 1967b. See section IV below.
52 Hayek, 1955, 91.
54 Hayek, 1967c, 72. Thus Hayek (1960, 63) argues that morality is a presupposition of reason.
55 Hayek, 1967c, 72.
56 Kurt Baier (1995) has recently provided a philosophical account of the social roots of reasoning that leads to some similar conclusions.
58 Hayek, 1988, 43.
59 Hayek, 1979, 159.
60 Caldwell, 2004, 361ff.
61 Hayek, 1979, 156.
62 For simplicity’s sake I assume, counterfactually, that each trait of the organism is produced by a distinct gene.
64 Hayek, 1960, 37. Emphasis added.
66 Hayek, 1960, 63. This seems to conflict with this insistence on inviolable rules in *Rules and Order* (1973, 57). But in the passage quoted in the text Hayek is concerned with individual violation; collective and coercive violation by the state, he goes on to say, is not permissible.
67 Hayek, 1979, 161.
68 For a useful account, see Blackmore, 1999, ch. 2.
69 Hayek, 1988, ch. 3.
70 Hayek, 1988, 25.
71 Hayek, 1979, 157.
73 Blackmore, 1999, 61.
74 Hayek, 1973, 74.
75 Hayek, 1967b, 56.
76 Hayek, 1973, 43.
77 Pinker, 2002, 292.
79 D. Dellarosa quoted in Gärdenfors, 2004, 41.
80 Hayek, 1952, 57.
82 On this see Pinker, 1997, 112ff; Gärdenfors, 2004, 40ff.
83 Hayek insisted on the link to his social theory. See Hayek, 1952, v.
84 Hayek, 1952, 7.
85 Hayek, 1967b, 51.
86 Hayek, 1952, 48.
87 Hayek, 1952, 51.
89 Hayek, 1967b, 57.
90 As does Hodgson, 1993, 165.
91 Hayek, 1952, 23ff.
92 Hayek, 1952, 64.
93 Hayek, 1952, 108.
94 Hayek, 1952, 165.
95 Hayek, 1979, 155.
96 Hayek, 1952, 110.
97 Hayek, 1979, 166.
99 Hayek, 1952, 35.
100 Hayek, 1979, 157.
101 Hayek, 1952, 19, 43ff, 185ff.
102 Hayek, 1952, 190.
103 Hayek, 1952, 192.
104 Hayek, 1952, 189.
106 See Paul, 1988, 258.
108 Although one should rigidly adhere to the system as a whole. Hayek, 1967d, 91.
110 Hayek, 1988, 8.
111 Hayek, 1988, 60ff.
112 Hayek, 1967c, 71; Hayek, 1960, 63.

Bibliography


