

The Evolution, Evaluation and Reform of Social Morality: A Hayekian Analysis*

Gerald Gaus

1 THE FUNDAMENTAL PHILOSOPHICAL QUESTION ABOUT THE EVOLUTION OF MORALITY

1.1 Two Senses of “Morality”

Like Darwin, I “fully subscribe to the judgment of those writers who maintain that of all the differences between man and the lower animals, the moral sense is by far the most important.”¹ A fundamental project in evolutionary science is to understand how this distinctive capacity, which appears to require that an individual sometimes refrain from the course of action that maximizes his fitness,² could have evolved. We have recently witnessed powerful analyses of the evolution of biological and psychological altruism, reciprocal cooperation, of our ability to follow rules and to socially enforce them, and of the development of conscience. Most, but not all, of this work has focused on biological evolution, employing both natural and social selection models, and increasingly employing some version of multi-level selection. This important work has made great progress in helping us understand the evolution of the building blocks of cooperation and morality. And until we know how our basic moral sense could have evolved, the entire moral enterprise — and by extension, the nature of human social life — remains an evolutionary mystery.

As important as this work is, I shall largely put it aside here, and focus instead on another line of analysis: the evolution of social-moral rules or, we might say, systems of social morality. Suppose, as I think is the case, that not only our basic moral

*My deep thanks to the members of my graduate seminars on diversity and on moral evolution at the University of Arizona for helping me think through some of these difficult issues. I am especially grateful for the comments and suggestions of Chad van Schoelandt, Stephen Stich, John Thrasher and Kevin Vallier. Special thanks to Fred D’Agostino and Ryan Muldoon for spurring me to think about these matters, and helping me do it.

1. Charles Darwin, *The Descent of Man*, second edition (New York: Penguin, [1879] 2004), p. 120.

2. Whether this is *mere* appearance — and, importantly, just what it could mean — is a crucial issue. The work of the last twenty years, I believe, inclines to the view that it is not simply appearance, and true biological altruism may evolve. But this is still a matter of controversy.

The Evolution and Reform of Social Morality/2

capacities have evolved but, through cultural evolution (or gene-culture coevolution), so has the content of our social morality — i.e., the system of moral rules and norms that forms the basis of our normative and empirical expectations, and which are enforced through general social disapproval of violations (see §2).³ The simple yet fundamental philosophical question arises: *so what?* From the perspective of the social scientist it might be of the utmost importance that culture evolves, and that our social-moral rules are a fundamental part of culture. Yet the moral philosopher wants to know what this fact about the history of our accepted social morality tells us about what our social rules *ought* to be. “[T]alk of morality,” Anthony O’Hear points out, “is itself ambiguous. Do we mean morality as that which is done and enforced within a particular group? Or do we mean that which in some absolute sense simply ought to be done, regardless of group norms and loyalties? And what, if anything, is the connection between the two?”⁴ If the first sense of morality does not tell us anything about the second, then the evolutionary theorist appears unable to shed any light on the moral philosopher’s question: what rules ought we to have? Too often the evolutionist has simply dismissed the philosophers’s concern. Ken Binmore, for instance, insists that “orthodox moral philosophy has gotten us nowhere because it asks the wrong questions. If morality evolved along with the human race, asking how we ought to live makes as much sense as asking what animals ought to exist, or which language we ought to speak.”⁵ This, though seems more evasive than helpful. Surely, we can and do step back from our evolved morality and ask whether it is the one we ought to have.

1.2 The Allure of Progress

Moral and political philosophers committed to an evolutionary analysis of culture have often sought to connect O’Hear’s two meanings of “morality” through a claim

3. For general analyses of such rules, see Cristina Bicchieri, *The Grammar of Society: The Nature and Dynamics of Norms* (Cambridge: Cambridge University Press, 2006); Gerald Gaus, *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), chaps. 3 and 4.

4. Anthony O’Hear, *Beyond Evolution: Human Nature and the Limits of Evolutionary Explanation* (Oxford: Oxford University Press, 1997), p. 101.

5. Ken Binmore, *Natural Justice* (Oxford: Oxford University Press, 2005), p. 1.

The Evolution and Reform of Social Morality/3

that evolution is morally progressive: the evolution of *positive morality* as that which is done and enforced within a particular group tends to increasingly approximate *true morality* in the sense of what truly ought to be done (or, let us say, the rules that are normatively justified). Nineteenth century evolutionary accounts of morality invariably associated evolution with development or progress.⁶ L.T. Hobhouse — perhaps the preeminent English political philosopher at the close of the century — applied evolutionary thought to social systems, seeking to show that evolution led to an increase in social integration and harmony, a research program that his student, Morris Ginsberg, continued throughout the first part of the twentieth century.⁷

A remarkably similar account of moral progress has recently been advanced by Philip Kitcher, one of today's leading philosophers of science. On Kitcher's analysis, the “ethical project” has its roots in our altruism failures — failures that lead to social conflict. Morality evolved as an adaptive response to such failures. The function of social morality throughout its evolutionary history has been, and remains, to remedy these failures. Echoing Hobhouse, Kitcher insists that “[e]thics must continue to promote social harmony through remedying altruism failure.”⁸ As Kitcher understands it, ethical progress is characterized by refinements in this function of social morality, leading to enhanced social harmony and, it seems, equality.⁹ To be sure, Kitcher rejects what he calls “crude evolutionary reductionism” according to which, apparently, there is a one-to-one mapping of evolved positive group morality to progress in justified morality.¹⁰ Ethical progress can be “unsteady,” and

6. See, e.g., Herbert Spencer, “Progress: Its Law and Cause,” in *Essays: Scientific, Political, and Speculative. Library Edition, containing Seven Essays not before republished, and various other Additions* (London: Williams and Norgate, 1891), ¶2. This essay, sketching Spencer's evolutionary theory, predates the publication of *The Origins of the Species*.

7. See, for example, L. T. Hobhouse, *Mind in Evolution* (London: Macmillan, 1901); Hobhouse, *Development and Purpose* (London: Macmillan, 1927); *Social Evolution and Political Theory* (New York, Columbia University Press, 1911); Morris Ginsberg, “The Concept of Evolution in Sociology,” *Proceedings of the Aristotelian Society*, ns 31 (1930–1): 201–224.

8. Philip Kitcher, *The Ethical Project* (Cambridge, MA: Harvard University Press, 2011), p. 239.

9. On progress and equality, see especially *ibid.*, chap. 6; on equality see *ibid.*, chap. 8.

10. *Ibid.*, p. 213.

evolutionary adaptations can be morally regressive yet, it seems, the evolution of positive and justified morality are closely tied.

Despite its allure, most careful analyses of social and moral evolution have long refused to associate the evolution of positive morality and moral progress. A fundamental concern has always been that progressive views often misunderstand evolution as a teleological process aiming at a goal, rather than a causal process driven by variation, selection and transmission. Leaving aside this fundamental error, moral philosophers have been especially critical of more orthodox Darwinian accounts of morality via natural selection, pointing out that Darwin stressed the “struggle for existence” as a central organizing idea of his theory.¹¹ Many moral philosophers have been skeptical that this Malthusian selection mechanism systematically maps on to what is highest for humans.¹² Indeed, it was the moral valorizing of the struggle for existence and its outcomes that was so objectionable in the hands of nineteenth-century Darwinian sociologists such as William Graham Sumner.¹³ So even if, say, evolution was “progressive” in some way, such as the development of complexity, many philosophers have insisted that there is no reason to think that this sort of progress constitutes, or implies, moral progress.¹⁴

2 HAYEKIAN ANALYSES OF THE EVOLUTION OF RULES AND ORDER

2.1 Cultural Evolution: variation and transmission

My aim in this essay is to take some preliminary steps towards understanding the conditions under which the evolution of positive morality is informative about true or justified morality, without appeal to vague claims that evolution is morally

11. “A struggle for existence inevitably follows from the high rate at which all organic beings tend to increase.” Charles Darwin, *On the Origins of Species* (Oxford University Press, 2009), chap. 3. Darwin explicitly credits Malthus with the idea. See *The Autobiography of Charles Darwin 1809-1882, with the original omissions restored*, edited by Nora Barlow (London: Collins, 1958), p. 120.

12. See, for example, D. G. Ritchie, *Darwinism and Politics*, second edn. (London: Swan Sonnenschein, 1891), pp. 1ff. Nevertheless, even Ritchie attempted to link evolution and progress in the second essay in the volume. Compare O’Hear’s skepticism, *Beyond Evolution*, pp. 2ff.

13. See, e.g., William Graham Sumner, “Sociology” in *War and Other Essays*, edited by Albert Galloway Keller (New Haven: Yale University Press, 1919), chap. 7.

14. See O’Hear, *Beyond Evolution*, p. 74. See also John Tyler Bonner, *The Evolution of Complexity by Means of Natural Selection* (Princeton: Princeton University Press, 1988).

progressive. We cannot make headway on this problem, however, without being clearer about the idea of moral evolution, as I hope to show that different accounts have very different resources for answering our fundamental philosophical question. I suppose here an account of moral evolution as a feature of cultural evolution, a type of analysis pioneered by F.A. Hayek, and more recently advanced by, among others, Robert Boyd and Peter J. Richerson.¹⁵ Hayek insisted that social evolution, though it was a path-dependent process¹⁶ that relied on competitive selection, did not rely on Darwinian natural selection.¹⁷ Cultural evolution, says Hayek, “simulates” Lamarckian evolution because acquired characteristics — rules and institutions — are transmitted from earlier to later generations.¹⁸ This is accomplished, he argues, through individual-to-individual transmission of social-moral rules, crucially through imitation.¹⁹ The more recent, and much more sophisticated, work of scholars such as Boyd and Richerson has greatly added to our understanding of cultural transmission, distinguishing conformity bias (doing as most others do), prestige bias (copying high status individuals), unbiased transmission, and various content biased transmissions such those having more vivid content.²⁰ It is important to appreciate that the teaching and preaching of

15. It is appropriate to commence with Hayek not simply because of the present occasion, but far more importantly, because Hayek was developing subtle accounts of moral and social evolution, complex systems and social morality from the 1950s through to the 1980s, when few social theorists would go anywhere near these topics. Well in to the 1980s the application of evolutionary thought to society was commonly associated with eugenics, imperialism and fascism. Recall the abuse heaped on E.O. Wilson for his *Sociobiology: The New Synthesis* (Cambridge MA: Harvard University Press, 1975).

16. See F.A. Hayek, “Notes on the Evolution of Systems of Rules of Conduct” in his *Studies in Philosophy, Politics, and Economics* (Chicago: University of Chicago Press, 1967): 66-81 at p. 75. See also Whitman, “Hayek contra Pangloss on Evolutionary Systems,” *Constitutional Political Economy*, vol 9 (1988): 450-466.

17. Hayek suggests a gene-culture coevolution thesis, but I shall not explore it here. See his “Notes on the Evolution of Systems of Rules of Conduct,” p. 71.

18. F.A. Hayek, *The Fatal Conceit*, edited by W.W. Bartley III (Chicago: University of Chicago Press, 1988), p. 25. For a contemporary discussion of whether cultural evolution is “Lamarckian” (allowing for the inheritance of acquired characteristics), see Alex Mesoudi, *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences* (Chicago: University of Chicago Press, 2011), pp. 43-44.

19. Hayek, “Notes on the Evolution of Systems of Rules of Conduct,” p. 67; F.A. Hayek, *The Political Order of a Free People* (Chicago: University of Chicago Press, 1979), pp. 156-57.

20. For an easily accessible version of their work, see Peter J. Richerson and Robert Boyd, *Not by Genes Alone: How Culture Transformed Human Evolution* (Chicago: University of Chicago Press, 2005), chap. 3.

social-moral rules is an important form of transmission.²¹

An evolutionary analysis requires, in addition to a *transmission* mechanism, sources of *variation* and *selection*. In cultural evolution, variation in social-moral rules can come from random changes, errors in transmission, drift or explicit revision.²² This last is especially important: there is no reason why a theory of cultural evolution cannot appeal to explicit efforts to improve social-moral rules; in this sense cultural evolution is by no means simply “blind.” Some might decide that a current moral rule is objectionable, and so, say start preaching an alternative. Although Hayek's evolutionary account is often criticized as having no room for conscious attempts at innovation, an evolutionary analysis *requires* variation, and Hayek certainly accepts that rules can be consciously altered.

2.2 Selection, Macro and Micro

It would seem that what cannot be consciously determined on Hayek's social evolutionary account is whether a cultural innovation is adaptive (but see §5). That requires competitive *selection* mechanisms. Hayek's account of selection is complicated, indeed more so than he often suggests. Selection occurs at both the macro and micro levels.²³

At the macro level, “the selection process of evolution will operate on the order as a whole;” what is selected, Hayek argues, is an “order of actions” that arises from numerous interacting rules, other elements of the social system and the wider environment.²⁴ At the macro level selection pressures operate directly on “the order

Their ground-breaking modeling of cultural evolution was presented in Robert Boyd and Peter J. Richerson, *Culture and the Evolutionary Process* (Chicago: University of Chicago Press, 1985). For an overview see Mesoudi, *Cultural Evolution*, chap 3. Hayek notes prestige bias; “Notes on the Evolution of Systems of Rules of Conduct,” p. 79.

21. This form of transmission has recently been stressed by Christopher Boehm, *Moral Origins: The Evolution of Virtue, Altruism and Shame* (New York: Basic Books, 2012), chap. 7.

22. Mesoudi, *Cultural Evolution*, chap 3; Richerson and Boyd, *Not by Genes Alone*, pp. 68ff.

23. On the contrast between micro and macro social evolution, see Mesoudi, *Cultural Evolution*, chaps. 3-5.

24. Hayek, “Notes on the Evolution of Systems of Rules of Conduct,” p. 71. On Hayek's notion of the order of actions, see Eric Mack “Hayek on Justice and the Order of Actions” in Edward Feser, *The Cambridge Companion to Hayek* (Cambridge: Cambridge University Press, 2006): 259-86.

of actions of a group.”²⁵ This distinction between a set of rules and the order of actions to which it gives rise is a fundamental insight of Hayek’s, which allows us to distinguish in our analysis the focus of selective pressure and the underlying rules, which are transmitted (§2.1). On Hayek’s analysis, a group of individuals living under a set of social rules R , composed of rules $\{r_1\dots r_n\}$, will give rise to a certain abstract pattern of social interactions, O , on which macro selection operates.²⁶ Hayek advanced a rather strong emergentist relation between R and O , seeing R as a complex system with O as an emergent property.²⁷ We need not follow him quite that far. What is fundamental to the analysis is that a specific order O_x is an abstract pattern of a large number of human interactions, which does not arise from any specific rule r , or the aggregated effects of a set of independent rules, but from a set of interacting rules in an environment E . Hayek described this as a sort of holism: “systems of rules of conduct will develop as a whole.”²⁸ We need, though, to distinguish two aspects of such “holism.” One is simply a restatement of the idea that O as a whole is the focus of macro selection; the other is that every rule in R is dependent on every other rule. This second claim is, once again, overly strong (see §4.3) and in any event not required for a Hayekian analysis; so long as there are considerable interdependencies in R , O will possess the sort of complex, non-aggregative relation to $\{r_1\dots r_n\}$ that characterizes adaptive landscapes, and sets the stage for some of Hayek’s crucial insights (§3).

On Hayek’s analysis macro social evolution is based on a form of group selection. “The rules of conduct have ... evolved because the groups who practiced them were more successful and displaced others.”²⁹ Just what is meant by “group selection” is a

25. Hayek, “Notes on the Evolution of Systems of Rules of Conduct,” p. 72.

26. F.A. Hayek, “The Theory of Complex Phenomena” in his *Studies in Philosophy, Politics, and Economics* (Chicago: University of Chicago Press, 1967): 22-42, at pp. 23-24.

27. I have analyzed this thesis in “Hayek on the Evolution of Society and Mind” in *The Cambridge Companion to Hayek*, pp. 232-258.

28. Hayek, “Notes on the Evolution of Systems of Rules of Conduct,” p. 71.

29. F.A. Hayek, *Rules and Order* (Chicago: University of Chicago Press, 1973), p. 18; Hayek, *The Fatal Conceit*, p. 25. Sewall Wright participated in Hayek’s evolution seminar at Chicago. See Bruce Caldwell, *Hayek’s Challenge* (Chicago: University of Chicago Press, 2004), p. 299. Hayek certainly advances what might be called a genuine multi-level selectionist account, in which the success of a group affects the

vexed issue; models with very different dynamics are often categorized under this rather vague term.³⁰ Leaving nomenclature aside, a crucial claim is that if society S_1 , characterized by order of actions O_1 , is more productive than S_2 based on O_2 , society S_1 will tend to win conflicts with S_2 , a mechanism akin to natural selection.³¹ But perhaps more importantly, the members of S_2 , seeing the better-off participants in S_1 characterized by O_1 , may either immigrate to S_1 , or seek to copy the underlying rules R_1 , thus inducing differential rates of reproduction between the two sets of underlying rules.³²

Although in some statements Hayek seems to suggest that all selection occurs at this macro level, his more nuanced view is that, while the macro level is the primary locus of selection, rule selection also takes place in the form of competition between rules within a society.³³ For a rule r to be selected, it must be contributory to a selected order, O , but it must also attract allegiance within the group of individuals who coordinate via r . Individuals are constantly testing rules to determine whether conformity suits their overall concerns; “it is, in fact, desirable that the rules should

selection of individual traits within it, allowing traits that have an in-group disadvantage to be selected. “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.” Hayek, “Notes on the Evolution of Systems of Rules of Conduct,” p. 72. This sort of group selection hypothesis is not supposed in this paper.

30. While the importance of forms of multi-level selection in biological evolution is still hotly disputed, I think there is conclusive reason to view multi-level selection as fundamental in cultural evolution. For a very helpful discussion, see Samir Okasha, *Evolution and the Levels of Selection* (Oxford: Oxford University Press, 2006).

31. On modeling group conflict as fundamental to social evolution, see Samuel Bowles and Herbert Gintis, *A Cooperative Species: Human Reciprocity and its Evolution* (Princeton: Princeton University Press, 2011).

32. F.A. Hayek, *Law, Legislation, and Liberty*, vol. 3: *The Political Order of a Free People* (Chicago: University of Chicago Press, 1979), pp. 26, 159; Hayek, *Rules and Order*, p. 3, 17-18; Hayek, *The Fatal Conceit*, pp. 6, 25, 43. For a general analyses of different forms of cultural group selection and the plausible time spans under which they might operate, see Robert Boyd, Peter J. Richerson and Joseph Soltis, “Can Group-Functional Behaviors Evolve by Cultural Group Selection?” in Robert Boyd and Peter J. J. Richerson's edited collection, *The Origin and Evolution of Culture* (Oxford: Oxford University Press, 2005): 204-226.

33. “[C]ultural evolution operates largely through group selection.” *The Fatal Conceit*, p. 23, emphasis added.

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.”³⁴ Although Hayek himself disparaged rule selection based on how well a rule conformed to one's social or moral ideals,³⁵ any plausible account of the selection of moral rules within a group must accord weight to how well those rules conform with the moral sense and judgment of different individuals.³⁶ One of the factors that determine within-group fitness of a moral rule is its ability to secure allegiance and be taught to the next generation. This is a case of content bias: rules that accord with people's moral sensibilities are more apt to be learned and transmitted. Hayek was certainly right to model micro-evolution into his account, but he was needlessly restrictive of the factors that affect cultural success and transmission.

3 TWO LANDSCAPES

3.1 Adaptive Landscapes in Social Evolution

Having some idea of the outlines of a Hayekian account of social evolution, we can model it a bit more rigorously as a type of “rugged landscape.” Consider a simplified version of Hayek's analysis, a set R with only two rules, r_1 and r_2 . As Hayek stressed, the rules interact, such that their effects on O are not simply aggregative (§2.2). The fitness of O is a non-additive function of r_1 and r_2 , thus

34. Hayek, *The Constitution of Liberty* (Chicago: University of Chicago Press, 1960), p. 63.

35. On page 161 of the Epilogue to *The Political Order of a Free Society* (the last volume of *Law, Legislation and Liberty*) Hayek 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.” For a general analysis of the role of conscious deliberation and choice of rules in Hayek, see Sandra J. Pert and David M. Levy, “Discussion, Construction and Evolution: Mill, Buchanan and Hayek on Constitutional Order,” *Constitutional Political Economy*, vol. 19 (2008): 3-18.

36. It does no good for Hayek to claim that this moral sense is atavistic, as it was formed under conditions of living in small bands, and so should be ignored. We must remember that this is an account of the mechanisms of cultural evolution; we cannot rule out processes on evaluative or moral grounds. See *The Political Order of a Free Society*, pp. 153-59.

giving rise to an NK optimization problem, which characterize many evolutionary adaptive landscapes. Figure 1 illustrates a highly simplified version of such a landscape.

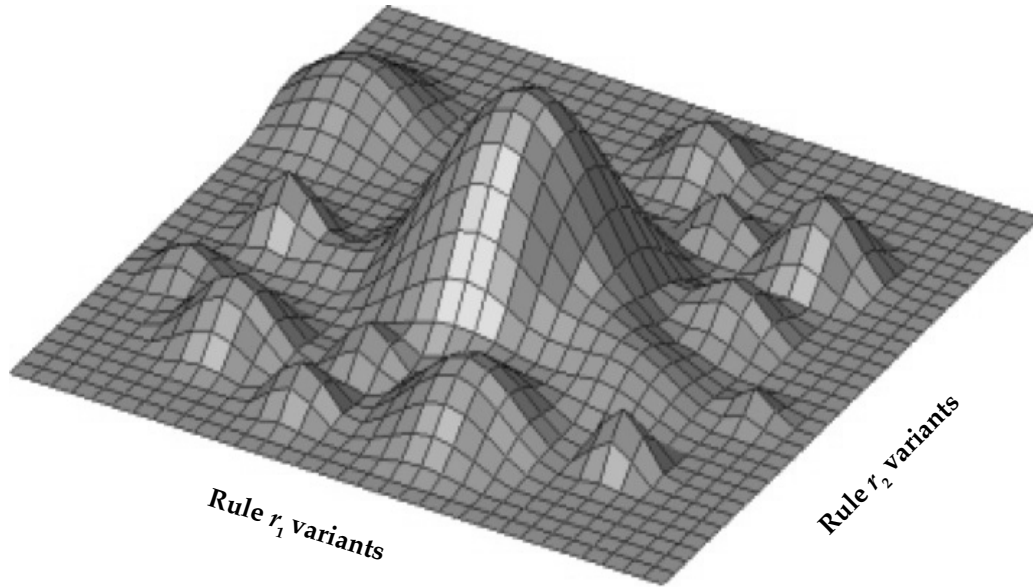


FIGURE 1: AN HAYEKIAN ADAPTIVE LANDSCAPE³⁷

I suppose now an environment that is sufficiently stable such that the landscape does not significantly change (but see below, §4.2). I suppose also that r_1 and r_2 are each classes of rule variants regulating some area of social life that can be arrayed along a dimension of similarity.³⁸ The topography of the landscape represents the fitness of the resulting O .³⁹ The ruggedness of the landscape results from the adaptive optimization problem possessing multiple N dimensions with K

37. From <http://cairnarvon.rotahall.org/2007/01/02/on-fitness-landscapes/> Used with permission.

38. I consider the idea of rule variants in *The Order of Public Reason* (Cambridge: Cambridge University Press, 2011), pp. 270-71.

39. In many models of adaptive landscapes the points represent population average fitness; in others the points represent individual fitness, and populations are groups of points. In evolutionary modeling these different versions can lead to very different insights and problems. In the present social evolutionary context O is an individual order, not a population average. See further Sergey Gavrilets, "Evolution and Speciation on Holey Adaptive Landscapes," *Trends in Ecology & Evolution*, vol. 12 (August 1997): 307-312.

interdependencies. When $K=0$, i.e., when there are no interdependencies between the dimensions, the landscape will not be rugged. Hayek's claim that selection operates on O , which emerges from *multiple interdependent rules*, however, commits him to the type of rugged adaptive landscape depicted in Figure 1.⁴⁰

We can now see why a Hayekian cannot be a simple believer in social evolution as always leading up, up, and onward.⁴¹ To take a simple case: if evolutionary selection focuses on two traits that are not interrelated we can model the fitness landscape as having a single peak, as in Figure 2.

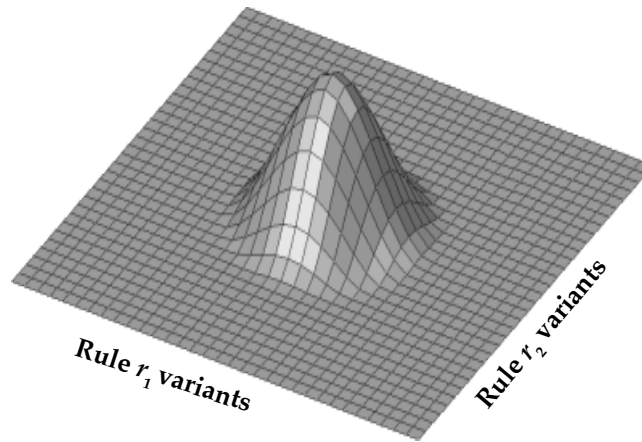


FIGURE 2: A MOUNT FUJI “PROGRESSIVE” LANDSCAPE⁴²

In a Mount Fuji landscape, the global optimum will be reached by a series of selected variations, no matter where on the landscape society is. In Figure 2, we suppose that there is one trait that regulates the functional capacity of the rules to resolve altruism failure, and that this is the only feature that is selected for. This results in constant gradients from all points to the global optimum. In contrast, in a rugged landscape it

40. For an important pioneering analysis, see Stuart A. Kaufman, *The Origins of Order* (New York: Oxford University Press, 1993), especially chap. 2. See also Sergey Gavrillets, “High-Dimensional Fitness Landscapes and Speciation” in *Evolution—the Extended Synthesis*, edited by Massimo Pigliucci and Gerd B. Müller (Cambridge, MA: MIT Press, 2010): 45-80.

41. See Whitman, “Hayek contra Pangloss on Evolutionary Systems.”

42. From <http://cairnarvon.rotahall.org/2007/01/02/on-fitness-landscapes/> Used with permission.

is possible for selection to get stuck at a local optimum that is far short of the global optimum (see further §5.2). Occupying a local optimum — in Figure 1, the top of a small hill — any near variation in either r_1 or r_2 results in a less fit order of actions.

None of this is to say that orders cannot evolve to the global optimum in a rugged landscape. One important consideration (to which we shall return, §4.3) is the basin of attraction of the global optimum. In Figure 1 the basin is large: i.e., from a number of different initial pairs of variations, there is a steady gradient to the global optimum. In other landscapes, as we shall see, the basin of attraction of the global optimum is much smaller. More generally, if different orders are spread over the landscape, some will climb to the top of higher local optima; under some conditions an order that climbs to a higher optimum (i.e., is more fit) could displace one stuck at a less fit one; if the less fit order observes the more fit, it may copy its $\{r_1, r_2\}$ variants and, perhaps (see §5.3), “jump” to the higher optimum. But observe also that even if O , stuck at a poor local optimum, can jump to a better local optimum, this could move O further away from the global optimum. The conditions under which O will find the global optimum is a fundamental issue in evolutionary modeling; we shall consider some of these matters in section 4.

3.2 Evaluative Landscapes

A striking development in the last twenty years has been the application of NK optimization analysis to a wide variety of evaluative problems. Rugged landscape models have been employed in management and other organizational contexts as an analysis of complex problem solving,⁴³ in epistemology and the philosophy of science as a way of understanding how diverse communities can best maximize a complex objective function,⁴⁴ and in the theory of collective deliberation and

43. See e.g., Daniel A. Levinthal, “Adaptation on Rugged Landscapes,” *Management Science*, vol. 43 (July 1997): 934-50 and, more generally, Scott E. Page, *The Difference* (Princeton: Princeton University Press, 2007). Especially important in management science has been the work of James G. March; see e.g., “Exploration and Exploitation in Organizational Learning,” *Organization Science* (Special Issue: Organizational Learning: Papers in Honor of (and by) James G. March), vol. 2 (1991): 71-87.

44. Fred D’Agostino, *Naturalizing Epistemology* (New York: Palgrave Macmillan, 2010), esp. chap. 7; Fred D’Agostino, “From the Organization to the Division of Cognitive Labor,” *Politics, Philosophy and Economics*, vol. 8 (2009): 101-129; Michael Weisberg and Ryan Muldoon, “Epistemic Landscapes and the Division of Cognitive Labor,” *Philosophy of Science*, vol. 76 (April 2009): 225-252.

democracy.⁴⁵ I have explored how it can be applied in more explicitly social-moral theoretical contexts.⁴⁶

Most large-scale moral evaluative problems in social and political philosophy readily fit the requirements of *NK* analysis. Consider, for example, questions of just distribution. Amartya Sen tells a parable of three children quarreling over a flute.⁴⁷ If, says Sen, we consider only claims based on who can best use the flute, it goes to Anne who alone can play it; if we consider only claims of need, it goes to Bob who is so impoverished that he has no other toys; if we consider only claims to desert and self-ownership, it goes to Carla, who made the flute. All three standards, Sen argues, qualify as impartial principles of justice. “At the heart of the particular problem of a unique impartial resolution of the perfectly just society is the possible sustainability of plural and competing reasons of justice, all of which have claims to impartiality and which nevertheless differ from — and rival — each other.”⁴⁸

If we follow Sen (and on this point I certainly think we should), justice is *multidimensional*. It also seems uncontroversial that these dimensions display *interdependencies*. For example, the “justice value” of a social state in which people generally deserve what they need — a wonderful condition — may be much higher than simply an aggregation of its need satisfaction and desert satisfaction scores; alternatively, a condition in which need and desert are balanced may be much more just than one which scores extraordinarily highly on one dimension but very low on the other. If optimizing justice is this sort of complex task, we are again confronted with an *NK* optimization problem, and so we are again confronted with a rugged landscape. Thinking back to our simple case of Figure 1, with simply two rule

45. Hélène Landemore, *Democratic Reason* (Princeton: Princeton University Press, 2013), esp. chap. 4.

46. “Social Contract and Social Choice,” *Rutgers Law Journal*, vol. 43 (Spring/Summer 2012): 243-76; “Between Discovery and Choice: The General Will in a Diverse Society,” *Contemporary Readings in Law and Social Justice*, vol. 3 (2011): 70-95. In what follows I will call the landscapes, “evaluative,” “moral” and “justice” landscapes. Although in some contexts it would be important to distinguish these, the analysis presented here is sufficiently general that we can treat these terms as synonymous.

47. Amartya Sen, *The Idea of Justice* (Cambridge, MA: Harvard University Press, 2009), pp. 12-15.

48. *Ibid.*, p. 12.

variants $\{r_1, r_2\}$, the topography is now determined not by Hayekian fitness, but by the justice score of each combination of variants.⁴⁹

I thus model “justified morality” as a function that generates a “justified morality (or justice) score” for every order of actions that emerges from various rule combinations. Some philosophers may resist this, insisting that “true” or “justified morality” is a single unique point (on the modeling employed here, the global optimum). So some philosophers might maintain that the global optimum is all that counts: if we are not there, we are nowhere. Attractive as this may be to many philosophers, I set aside this rather austere depiction of our moral thinking. My concern is with those philosophers willing to adopt Sen's imagery of a moral mountain range, where we are trying to climb upwards to increasingly just arrangements, while also recognizing that in some case moving to a “more just” state of affairs may in some sense lead us further from the “most just” order.⁵⁰ It is this conception of “true” morality that is captured by the evaluative landscape model.

4 CONDITIONS FOR THE CORRELATION OF THE TWO LANDSCAPES

4.1 The Correlation View

We now see that for the theorist of social-moral evolution to reply to the philosopher's “so what?” challenge she need not appeal to some vague idea that evolution is a mechanism of moral progress. The question is: are the adaptive and evaluative landscapes correlated? To the extent to which we have reason to suppose that the topographies of the two landscapes are positively correlated, we then have grounds for supposing that our evolved positive morality is indicative of true morality or justice. This is not to say that the fact of evolution is in itself a basic justificatory grounds for claims about true morality or justice, but it is to say that the fact that our positive morality has evolved can ground a confidence that, to some extent, we also have made some headway towards true morality or justice.

49. I assume here that each rule varies along one evaluative dimension.

50. See Sen, *The Idea of Justice*, 102; A. John Simmons, “Ideal and Nonideal Theory,” *Philosophy & Public Affairs*, vol. 38 (2010): 34-35.

Suppose we are at some point a on the adaptive landscape. If the two landscapes are significantly positively correlated, we know that how “well we have done” from the evolutionary perspective is indicative of how “well we have done” morally — our current position m on the moral landscape. This does not entail that either a or m is a local, much less a global, optimum, though evolutionary selection drives us toward local optima, and in some cases the global optimum; neither does correlation imply that, if a is a local or global optimum on the adaptive landscape, m will be one on the moral landscape.⁵¹ And most importantly, the correlated landscapes view does not suppose that evolutionary and moral progress march hand-in-hand (the correlation need not be 1), much less that we are climbing perfectly correlated Mount Fuji landscapes. Because of this, the correlation view avoids one of the most dubious aspects of (some statements of) Hayek's analysis, which elsewhere I have called the “sufficiency claim” — that order O has evolved is sufficient reason for us to endorse it as moral, or at least not to question it.⁵² As we shall see (§5), moral evaluation, criticism and reform make perfect sense on the correlation view. And yet it captures Hayek's important insight that we do not know the full conditions for a perfectly just social order and could not plan one out from scratch (§5.2). We approach justice through an evolutionary process.

51. We shall see in section 5.2 that this point has important implications.

52. I consider a slightly different version of this claim in *The Order of Public Reason*, pp. 420ff. As John Thrasher has pointed out to me, James Buchanan advanced similar criticisms of Hayek; see, for example, Geoffrey Brennan and James Buchanan, *The Reason of Rules* (Cambridge: Cambridge University Press, 1985), pp. 9-10. See also Pert and Levy, “Discussion, Construction and Evolution: Mill, Buchanan and Hayek on Constitutional Order.” At various places, rather than endorsing the claim that the evolution of an order is sufficient for its moral endorsement, Hayek appears to suggest that the question of moral justification is rationalistic and so inappropriate: morality requires “following of moral traditions that are not justifiable in terms of the canons of traditional theories of rationality....The process of selection that shaped customs and morality could take account of more factual circumstances that individuals could perceive, and in consequence tradition is in some respects superior to, or 'wiser,' than, human reason....” (*The Fatal Conceit*, p. 75) In one sense, this is consistent with the correlation analysis presented here (that we can use evolution as indicative of the adequacy of our morality, and so “in some sense” it is a source of justificatory wisdom), but in a more radical interpretation it seems to suggest that we cannot rationally morally evaluate our currently evolved order. O'Hear rightly criticizes this radical view, pointing out that Hayek himself engages in overall evaluative judgments; *Beyond Evolution*, p. 148. As Caldwell notes, caution must be exercised when relying on *The Fatal Conceit* in interpreting Hayek, as some of the final text seems to reflect the views of Bartley, who finished the manuscript because of Hayek's failing health. See *Hayek's Challenge*, pp. 316ff.

4.2 The Substantive Condition for Correlation: Dancing in the Right Way

THE ENDOGENEITY OF TRUE MORALITY. The correlation view is compelling when two conditions hold. The first, which I explore in this section, is substantive: that we have available some plausible account such that one landscape causally affects the other, producing the positive correlation. Under this condition, an observed correlation would not be spurious, and we could have confidence that one landscape is in some way truly tracking the other. An important insight of Hayek is that the evaluative landscape tracks the adaptive landscape because, at any given time t , what we believe about true justice or morality (the evaluative landscape) is a function of our current cultural evolution (the adaptive landscape).

From the perspective of social and cultural evolution, we currently possess an evolved system of social morality. Now, Hayek argues, we certainly can criticize and reform this morality, but our resources for thinking this criticism through will crucially depend on our presently evolved morality. Hayek's rejection of proposed revolutions based on radical moral views is not that they criticize the current evolved morality, but that they fail to consider the entire evolved system, building on only some bits of it. "I wish neither to deny reason the power to improve norms and institutions nor even to insist that it is incapable of recasting the whole of our moral system in the direction now commonly conceived as 'social justice.' We can do so, however, only by probing every part of the system of morals."⁵³ On this account, then, our views of true morality (i) must commence with our currently evolved positive morality and (ii) an adequate criticism and reform must have some grounds for thinking the entire system is, from the moral point of view, better. Recall that Hayek conceives of moral rules as producing an order of actions, which is a type of emergent property (§2.2) on the system of moral rules; it is ultimately the pattern of human interactions, O , that we are interested in from the normative perspective (§5.3). A proposed global moral reform that was only based on a fragment of positive morality (say, our devotion to equality), would have no grounds for predicting what order of actions could possibly result — except, perhaps, a reasonably certain

53. Hayek, *The Fatal Conceit*, p. 8.

knowledge that no order of actions at all would result, because the moral system had become too simplified.

This analysis of moral evaluation and reform has much in common with coherence analyses of justified belief. On a standard coherence view, one starts with a system of belief $\{b_1...b_n\}$, which forms a system of mutual inferences with a degree of coherence, C .⁵⁴ The aim of epistemic improvement — a body of better justified beliefs — is to improve C by alternations (basically, deletions and additions to) $\{b_1...b_n\}$, producing $\{b^*_1...b^*_n\}$. Now we can see that $\{b_1...b_n\}$ and $\{b^*_1...b^*_n\}$ will be strongly correlated. This is not to say that every case of belief change will result in marginal changes, such that the two sets always will be very similar; in some cases $\{b^*_1...b^*_n\}$ may be a rather large jump from $\{b_1...b_n\}$, for example where some central organizing belief of the latter had to be dropped. But over the run of revisions, the two will be strongly correlated because $\{b^*_1...b^*_n\}$ is largely endogenous to $\{b_1...b_n\}$; the former is, to a large extent, the outgrowth of the latter. Not too surprisingly, just as Hayek's view has been deemed “conservative”⁵⁵ so have coherence theories; both see the current body of rules/beliefs as assets to be conserved while improving upon them. For both, the core insight is that we can only start our thinking from where we are, so an improvement typically will be a type of conservative modification of where we are.

To be sure, many moral theories deny (i) and (ii). Many assert that moral knowledge is entirely independent of current positive morality (e.g., some follow Plato in claiming to possess direct intuition of an objective, mind-independent, moral truth), or that moral improvement need not take into account the entire current system, but can select simply one part of our positive morality and employ it as the basis of a revolution in morals (as Godwin and Bentham sought to do with the greatest happiness principle). Our concern here is not to show that all proposed

54. I consider coherence theories in more detail in *Justificatory Liberalism* (New York: Oxford University Press, 1996), chap. 6.

55. Cf. F. A. Hayek, “Why I am not a Conservative” in the *The Constitution of Liberty*, pp. 397-411. See further section 5.1.

moral theories endorse correlation (that would be absurd) but to gain a better understanding of the sort of accounts of morality that do so.

HOW TRUE MORALITY AFFECTS EVOLVED MORALITY. We have seen that on a Hayekian account, at any given time t , what we believe about true justice or morality is a function of our current cultural evolution. Causal influence also runs in the other direction, for the positive morality evolved at t is to some extent the result of the population's views on true morality at $t-1$. As we saw in section 2.2, on the Hayekian analysis social evolution occurs at the micro and macro levels; especially important here is the former. At the micro level, whether any specific rule variant r survives as part in the system of moral rules partly depends on r 's tendency to be internalized and followed by the population. Again, this depends on two key claims (i) the population as a whole tends to have a reasonable grasp of justified or true morality and (ii) the tendency of individuals to internalize and act on rule variant r is significantly influenced by their understanding of true or justified morality. This second claim does not presuppose that this understanding is the sole factor influencing internalization and conformity. It is nearly universally acknowledged that normative guidance evolved, as it were, on top of an essentially egoistic motivational system;⁵⁶ the factors influencing rule conformity are certainly complex. We need only suppose here that reflective normative deliberation is an important factor in determining rule conformity: rules that strike many as mistaken, unjust or unjustifiable, will tend to fail to attract sufficient conformity to survive in the moral system.⁵⁷ A striking example are moral rules against homosexuality and homosexual parenting; they have been abandoned with breathtaking rapidity.

Many moral philosophers are apt to resist the first claim, that the population generally has a pretty good understanding of justified or true morality. From Plato

56. Kitcher makes much of this in *The Ethical Project*, Part I. For a fascinating narrative of this evolution, see Christopher Boehm, *Moral Origins*. For more rigorous modeling, see Bowles and Gintis, *A Cooperative Species*.

57. This supposes, *pace* the well-known work of Jonathan Haidt, that reflective moral consciousness is an important variable in action. See his "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment," *Psychological Review*, vol. 108 (2001): 814-34. My understanding of Haidt's work owes much to Piper Bringham.

to the present, moral philosophers have often insisted that inquiry into “true morality” is a specialized and technical endeavor which only some (i.e., they) have adequate competence. Importantly, endorsing claim (ii) does not require denying that philosophers have expert knowledge.⁵⁸ Scott E. Page's analysis of optimization on rugged evaluative landscapes indicates that while a group of homogenous experts (those conceiving of the problem in similar ways and employing similar approaches) will seldom get stuck at poor local optima, they do tend to get caught on “high” but suboptimal peaks. In contrast, in a diverse population of less (but not in-) competent searchers, some in the group will tend to (under certain assumptions, they definitely will) find the global optimum.⁵⁹ More generally, competent diversity rather than expertise is generally the best way to explore rugged evaluative landscapes.

This has fundamental implications for understanding Hayekian morality. As Hayek stressed, his main focus was on the evolution of morality in what he called “the Great Society,”⁶⁰ a large-scale diverse society. It is such diverse societies that provide the basic dynamics for efficient searching of rugged evaluative landscapes if one person's discoveries are taken up by others. To exploit the exploration of others⁶¹ we need to learn from them or, more generally, copy them.⁶² Conformity and prestige bias (§2.1) are no doubt important ways in which this occurs. In any event, we have reason to suppose that in certain sorts of diverse societies with significant copying of the successful, claim (ii) is warranted: the general population will have a good grasp

58. Of course neither does it imply that this claim of philosophers should be accepted!

59. Page, *The Difference*, Parts II and III. Page's book is based on formal theorems developed with Lu Hong. See Hong and Page, “Problem Solving by Heterogeneous Agents,” *Journal of Economic Theory*, vol. 97 (2001): 123-63; Hong and Page, “Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101 (Nov. 16, 2004): 16385-16389.

60. Hayek, *Rules and Order*, pp. 1-8, and 146ff where he talks of the “open society.”

61. See March, “Exploration and Exploitation in Organizational Learning;” Scott E. Page, *Diversity and Complexity* (Princeton: Princeton University Press, 2011), 122-24.

62. On the importance of imitative learning to the evolution of morality, see Hayek, *The Fatal Conceit*, p. 21.

of improvements on the evaluative landscape, and this will affect the fitness of rule variants.

COUPLED LANDSCAPES. The analysis of this section leads to the conclusion that on certain Hayekian views the adaptive and evaluative landscapes are coupled. Coupling of two adaptive landscapes is by no means unusual. If two species are interdependent, their adaptive landscapes will be coupled: a change in one species' adaptive landscape will produce changes in the landscape of the other. When this type of coevolution occurs we have dancing landscapes.⁶³ So too with the evolutionary and evaluative landscapes. Social evolution affects our views of what is morally justifiable and so changes the evaluative landscape, and these understandings of moral justifiability feed back into the adaptive landscape, changing the fitness of orders of actions. It is this dance that is the foundation of the moral relevance of social evolution.

Once two systems are coupled in this way, optimization dynamics become much more complex; neither system can be strictly modeled as having local optima that serve as steady attractors. Indeed, it can be difficult to model such systems in terms of optimization at all.⁶⁴ Modeling depends here on the nature of interactions and, as we shall presently see, characteristics of each landscape. Under some conditions couplings of rugged landscapes can lead to chaotic fluctuations in each landscape.⁶⁵ Crucial to whether the coupled systems are chaotic or display order is the complexity of each maximization problem, a question to which we will now turn.

4.3 The Formal Condition for Correlation: Modest Interdependence (No Holism)

Critical to whether a rugged landscape has significant optima to which the system gravitates or, instead, displays chaotic features, is the complexity of the optimization problem. Recall that rugged landscapes are created by *NK* optimization problems:

63. Page, *Diversity and Complexity*, p. 94.

64. Kaufmann, *The Origins of Order*, p. 238.

65. *Ibid.*

we are seeking to optimize over N dimensions with K interdependencies between the dimensions (§3.1). Recall also that if $K=0$, the N dimensions are independent, and we are faced with a simple aggregation problem: as we increase our success on any dimension we move higher on the landscape. However, as Kaufmann stressed in his analysis, if we have a number of dimensions and interdependencies are very high, the landscape will be fully random.⁶⁶ Let us call a *high dimensional landscape* one in which many dimensions display a large number of interdependencies, at the limit each dimension is affected by all others. In terms of social evolution, in a high dimensional landscape there is no systematic relation between the fitness of O and its one-rule variant mutation. Such landscapes have a very large number of poor local optima.⁶⁷

The crux of such high dimensional landscapes is that the fitness (or, more generally, value) of any one rule is a function of all others, producing what Kaufmann called “a complexity catastrophe.”⁶⁸ Now as we saw earlier (§2.2), Hayek is attracted to a type of holism, as are many philosophers. “A sensible contractualism,” writes T.M. Scanlon, “like most other plausible views, will involve holism about moral justification.”⁶⁹ According to such holist views, the justification of every element of a system of values or beliefs is dependent on many others — such systems are often depicted as “webs,” indicating a very high degree of interdependence among many variables. It is precisely such systems that give rise to complexity catastrophes; a small variation in one value can jump the system to a radically different state. Coupling two such systems produces a highly chaotic dance.

66. *Ibid.*, pp. 45ff.

67. *Ibid.*, p. 47. In deliberative searches groups will tend to get caught at a local optima since they abound (but see note 88). Crucial to Hong and Page's fundamental theorem is that only the global optimum is shared by all perspectives; if a local optimum is shared, collective search may end there. In high-dimensional landscapes poor local optima abound, and everyone is likely to share one, thus halting the collective search. Hong and Page, “Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers,” pp., 16387ff

68. Kaufmann, *The Origins of Order*, p. 52.

69. T. M. Scanlon, *What we Owe Each Other* (Cambridge, MA: Harvard University Press, 1998), p. 214.

As K decreases (i) the number of local peaks decreases, (ii) the slopes lessen, so that the basin of attraction of the optima are wider (the same optimum is reached from a wider array of starting points), and (iii) the peaks are higher.⁷⁰ Additionally, in low dimensional landscapes (iv) the highest optima tend to be near each other⁷¹ and (v) the highest optima tend to have the largest basins of attraction.⁷² As K decreases the landscape becomes correlated within itself. In the case of the adaptive social evolutionary landscape, one-rule variants will yield orders of action O^* that have values that are correlated with that of the current order O . In a low dimensional evaluative landscape, one-rule variants produce overall moral changes that are incremental *vis a vis* the current order.

Correlation of the adaptive and evaluative landscapes thus requires social evolutionary and moral evaluative theories that possess multiple dimensions with a modest degree of interdependence. In such correlated landscapes adaptive heights are associated with evaluative heights, and such heights have large basins of attraction; small changes do not drive either system to radically different values. What is required is that the landscapes waltz, not jitterbug.

All of this constitutes something like a possibility proof; under some conditions the two landscapes are correlated, and when this occurs, knowing that our social morality is adaptive tells us something important about true morality. The conditions for the correlated landscape answer to the philosopher's fundamental question are certainly non-trivial — but not, I think, implausible. I have argued elsewhere that evaluative judgments have modest interdependencies.⁷³ Whether the social evolutionary adaptive landscape is low/modest or high-dimensional is a critical issue to be investigated. On the one hand, it certainly is plausible to suppose that cultural traits are highly interconnected, forming complex systems with

70. Kaufmann, *The Origins of Order*, p. 243. D'Agostino notes these features in *Naturalizing Epistemology*, pp. 118-19.

71. Kaufmann, *The Origins of Order*, p. 60.

72. *Ibid.*, pp. 62-63.

73. See my *The Order of Public Reason*, pp. 495-7.

extremely rugged landscapes.⁷⁴ Yet relatively simple models and analyses that focus on the adaptive value of single traits have been extremely enlightening. And, indeed, insofar as cultural evolution postulates selection through copying useful traits, it must be possible to identify a trait (or small number of allied traits) as useful independently of the adaptiveness of the entire order, *O*, which results from their inclusion. While theorists such as Boyd and Richerson rightly insist that simple models can be useful steps in a cumulative understanding of more complex phenomena, this has its limits:⁷⁵ if the adaptive value of a very small variation in *O* is uncorrelated with *O*'s present value, models of the adaptive value of single traits become less enlightening.

5 SOCIAL EVOLUTION AND MORAL REFORM

5.1 Does Social Evolutionary Analysis Condemn Intentional Moral Change?

Focusing on the formal features of the landscapes also helps us make progress on a second question that is repeatedly raised about evolutionary accounts such as Hayek's: do they provide a basis for deliberative criticism and intentional reform of our morality? Hayek depicts an order of actions *O* as an emergent property that arises out of a complex and self-organizing system of moral rules, *R*. He repeatedly stresses that because *R* is a complex system, its micro behavior cannot be predicted or controlled.⁷⁶ It is because of this complexity, and our resultant inability to engineer the system, that Hayek so stresses self-organization and evolution.⁷⁷ That we can

74. As did Hayek. Hayek, "The Theory of Complex Phenomena." I explore this claim in some depth in my "Social Complexity and Evolved Moral Principles" in *Liberalism, Conservatism, and Hayek's Idea of Spontaneous Order*, edited by Louis Hunt and Peter McNamara (London: Palgrave Macmillan, 2007): 149-176.

75. Robert Boyd and Peter J. Richerson, *The Origin and Evolution of Cultures* (Oxford: Oxford University Press, 2005), chap. 19.

76. Hayek, "The Theory of Complex Phenomena." I explore this claim in some depth in my "Social Complexity and Evolved Moral Principles."

77. Hayek repeatedly refers to "the twin ideas of evolution and spontaneous order." See Hayek: "Notes on the Evolution of Systems of Rules of Conduct," p. 77; "Dr. Bernard Mandeville" in his *New Studies in Politics, Economics and the History of Ideas* (London: Routledge, 1987), pp. 249-266 at 250; *Rules and Order*, pp. 23, 158; *The Fatal Conceit*, p. 146. It should be noted that Kaufmann's aim in *The Origins of Order* was to account for self-organization within an evolutionary framework.

rationally design a moral system, he famously argued, is a fatal conceit.⁷⁸ But all this would appear to undermine the efficacy of conscious criticism and reform; if we cannot predict the consequences of changes in our moral system, all moral reform would seem to be a shot in the dark. Thus many believe that Hayek is committed to a Burkean reverence for traditional moral rules, which have critical functions in our order of actions that we cannot fully understand.⁷⁹ Indeed Hayek associates himself with the Whig tradition, which includes Burke,⁸⁰ but Burke was suspicious of moral change; it would seem that an evolutionary view such as Hayek's, despite its protestations, must follow suit.

The analysis of the previous section helps us sort out this problem. If R is a highly complex system characterized by a very high dimensional landscape, then indeed a small variation in one rule r can result in an order O with a drastically different height (value). This may lead to caution about changing our existing moral rules; any small change in R and we could fall far down.⁸¹ However, under these circumstances, the observation that our current morality has evolved (i.e., is adaptive) would not itself provide much support for “moral reverence” towards our existing order. Unless this chaotic moral landscape is tied extraordinarily closely to the adaptive landscape — unless, essentially, every point a on the adaptive landscape is linked to a unique m on the moral landscape — the same level a of evolutionary adaptation will display significant degrees of freedom in relation to a set of points $\{m_1...m_n\}$ with which it is correlated. But if the evaluative system suffers from complexity catastrophe, the points within $\{m_1...m_n\}$ are not correlated with each

78. Hayek, *The Fatal Conceit*, p. 7.

79. Hayek himself commends “the reverence of the traditional,” invoking Hume's doctrine that the rules of morality are not the conclusions of our reason. *The Constitution of Liberty*, p. 63. See further my “Social Complexity and Evolved Moral Principles,” pp. 166ff. O’Hear complains about this; *Beyond Evolution*, pp. 154ff.

80. Hayek, “Why I am not a Conservative.” But see p. 400 where Hayek distinguishes the liberal and conservative attitudes toward change.

81. Or, then again, also way up. On this model we need to add risk aversion to get reverence. Of course other, more specific, models could lead to Burkean reverence. Consider a landscape that has few high peaks that are far apart, each with very sharp slopes, with the rest of the terrain being low and flat. If one thought we are currently at a local optima we would not budge, for fear of falling off a cliff.

other; thus any given a is linked to a set of uncorrelated evaluative points. If so, we cannot infer anything about the goodness of the current rules from the adaptiveness of a , even if we are currently high on the adaptive landscape.

Assume instead that the evaluative landscape models a system with low/modest interactions that is correlated with a low/modest dimensional adaptive landscape. In such an evaluative landscape the basins of attraction of good optima are large, the slopes gentle. Here we need not be worried that moral change will have disastrous results. We could get it wrong and descend a bit on the evaluative landscape, or we could get it right and ascend. But the fact that our current morality has evolved will not lead us to be terrified of the moral results of moral reform. And when the adaptive landscape is similarly low dimensional, we will not fear the social fitness consequences of moral change. It is certainly possible that moral change could move us down an adaptive gradient,⁸² but if the adaptive landscape is not highly rugged (i.e., highly dimensional), even if the evaluative and adaptive landscape are highly correlated, we would not expect sudden losses in social fitness.

5.2 Reform v. Revolution: The Importance of Neighborhoods

Given this, however, we might be driven to ask: what is left of the Hayekian project? While in my view his work on social and moral evolution is independently valuable, in his larger project these elements were marshaled in support of his master claim in social and political philosophy that planning is infeasible. A standard criticism of the infeasibility claims is that his support of deliberate moral criticism and change shows that we can, after all, predict the consequences of social change, and so planning is, after all, possible.

I have argued that, if Hayek's analysis is to avoid a complexity catastrophe, the landscapes must be low/modest dimensional, which does indeed allow moral reform, and so "reverence" for current positive morality is indeed inappropriate. This, though, does not lead to "planning" if, as did Karl Popper, we mean by it a synoptic social engineering that seeks to build the entire system according to some

82. This is a way of modeling O'Hear's point that the best norms may run "counter to survival." *Beyond Evolution*, p. 3. This clearly has limits.

blueprint. Popper contrasted such utopian planning to “piecemeal” reform, seeing Hayek as an opponent only of the former.⁸³ Our question, then, is: does Hayek’s evolutionary analysis provide grounds for this distinction, allowing “piecemeal” reform of our morality, but not utopian reconstruction from the ground up? It should be stressed that while utopian economic plans have faded since Hayek wrote *The Road to Serfdom*,⁸⁴ contemporary moral and political philosophy is to a large, and perhaps increasing, extent committed to the production of utopian moral schemes.⁸⁵ So if his analysis speaks against them, this is of some importance.

And it does. I have stressed that in low dimensional evaluative landscapes where slopes are modest and basins of attractions of high peaks are large, moral reform is well-grounded. We only have to make small variant decisions, and our concern is usually to climb a gradient (but see below). We do not need really powerful predictive tools about the overall effects of moral changes on the order of actions (see §5.3); as Popper stressed, if we get a piecemeal change wrong, we can climb back to where we were, and try again to move in an ascending direction. The costs are apt to be modest. However even in low dimensional landscapes correlation obtains only in some neighborhood; locations outside the correlation neighborhood are fully uncorrelated with respect to one’s present location.⁸⁶ Thus in any move outside one’s present neighborhood to position u , the value of one’s current order of actions (on the evaluative landscape) provides no information about the moral value of u . Thus accuracy in determining the value of u depends entirely on the power of the predictive models employed.

This is the heart of the Hayekian critique of utopian moral change; it is purely rationalistic in the sense that its success entirely depends on models for predicting the order of actions O that will emerge from a set of rules R . Although emergentist

83. Karl R. Popper, *The Poverty of Historicism*, 2nd edn. (London: Routledge and Kegan Paul, 1961), chaps. 20-25. Popper insightfully analyzes the way that holism undermines social experimentation.

84. F.A. Hayek, *The Road to Serfdom* (Chicago: University of Chicago Press, 1944).

85. For just two of many examples, see David Estlund, *Democratic Authority* (Princeton: Princeton University Press, 2008), chap. 14; G.A. Cohen, *Rescuing Justice and Equality* (Cambridge, MA: Harvard University Press, 2008), Part II.

86. Kaufmann, *The Origins of Order*, p. 70.

language tends to suggest that it is well-nigh impossible to predict the order to which a given R produces, we certainly need not go that far. It enough to observe that the predictive models have large variances, and different perspectives employ different predictive models (see further §5.3).⁸⁷ To rely entirely on such models can lead us to unexpectedly awful parts of the adaptive and evaluative landscapes. Thus, because as we move further from our current location our present value (on both the adaptive and moral landscapes) is decreasingly informative of the value of the proposed change, and so we are thrown back entirely on predictive models of at best uncertain accuracy, utopian schemes of moral revolution are much too like shots in the dark.

A caveat is order. I said above that moral reform will typically be a matter of gradient climbing, and in low dimensional landscapes with large basins of attraction for high peaks, this will indeed usually be the case. But a society that only climbs gradients is also one that is apt to get caught at local optima.⁸⁸ In some cases we may occupy a poor local optimum and have a not-so-distant better moral optimum that is sufficiently approximate in its social and institutional character that we can have reasonable confidence in our predictive models. From the *perspective of simply the moral landscape*, risking more radical innovations by seeking to jump to a new optimum may be the only way out, hoping that the predictive models are up to the task, and we do not land at an evaluative gully. However, *from the perspective of the adaptive landscape*, we should recall that the order of actions is also subject to adaptive pressures, which may induce changes in O . If the current poor local optimum on the evaluative landscape is such that O is ill-adapted, given that the adaptive and evaluative orders are coupled, movement in the adaptive landscape will produce changes in the moral landscape. This can result in a new evaluative

87. Hong and Page (“Problem Solving by Heterogeneous Agents”) include in their theorem the way that a group employing diverse predictive models will outperform a single, even very good, model.

88. I have stressed that low dimensional landscapes have many attractive features given the type of problems we have been analyzing. But as has been recognized since Sewall Wright, they pose this problem of getting caught on suboptimal peaks. Suboptimal traps can be avoided in some high-dimensional landscapes. If we consider landscapes in which points are individuals and species are groups of points spread over a N -dimensional area (see footnote 39 above), high dimensional landscapes can display fitness ridges that provide paths from one optima to another. See Gavrillets, “High-Dimensional Fitness Landscapes and Speciation.”

landscape with a gradient to a morally better location. Indeed, a great attraction of an account that employs coupled adaptive and moral landscapes — in addition, of course, to the fundamental one that it both distinguishes and relates fitness and moral evaluation — is that dancing landscapes are less liable to get caught at local optima, as movement in one changes the terrain of the other.

To see this more clearly, consider Figure 3. In panel I, O is presently at a moral local optimum (m), with a higher optimum “near by;” the adaptive point is a , not a local optimum. In panel II a has climbed its local gradient, but because the moral landscape is coupled, the moral landscape has shifted, putting m on an upward gradient.⁸⁹ A nice example of this type of analysis is John Stuart Mill's discussion of capitalism and socialism.⁹⁰ To Mill, from the perspective of justice Victorian capitalism was far below the moral optimum; but, in his eyes, even if it climbed its local gradient to the top, and so became something like the best it could be, it would still be suboptimal from the perspective of justice. On the other hand, forms of socialism looked as if they could be moral attractive alternatives from the perspective of justice, but Mill adamantly opposed a revolutionary leap to such orders. This would require a complete restructuring of the society and economy; not only would the results be highly uncertain, but the transition period may be worse than either alternative. However, Mill insisted the evolution of new forms of partnerships and corporations (that render capitalism more efficient) would also allow competitive market processes within capitalism to test the viability of socialist experiments and provide a smooth moral gradient from capitalism to what he saw as a more just, cooperative, order.

89. Of course the effect can work in the other direction as well, with movement on the moral landscape changing the adaptive landscape.

90. I draw here on Mill, *Principles of Political Economy*, Books II and IV, in *The Collected Works of John Stuart Mill*, edited by John M. Robson (Indianapolis: Liberty Fund, 2006), vols. 2 and 3 and his “Chapters on Socialism” in John Stuart Mill, *Principles of Political Economy with Chapters on Socialism*, edited by Jonathan Riley (Oxford: Oxford University Press, 1994). I am not concerned here with whether Mill was correct about either his moral or social analysis.

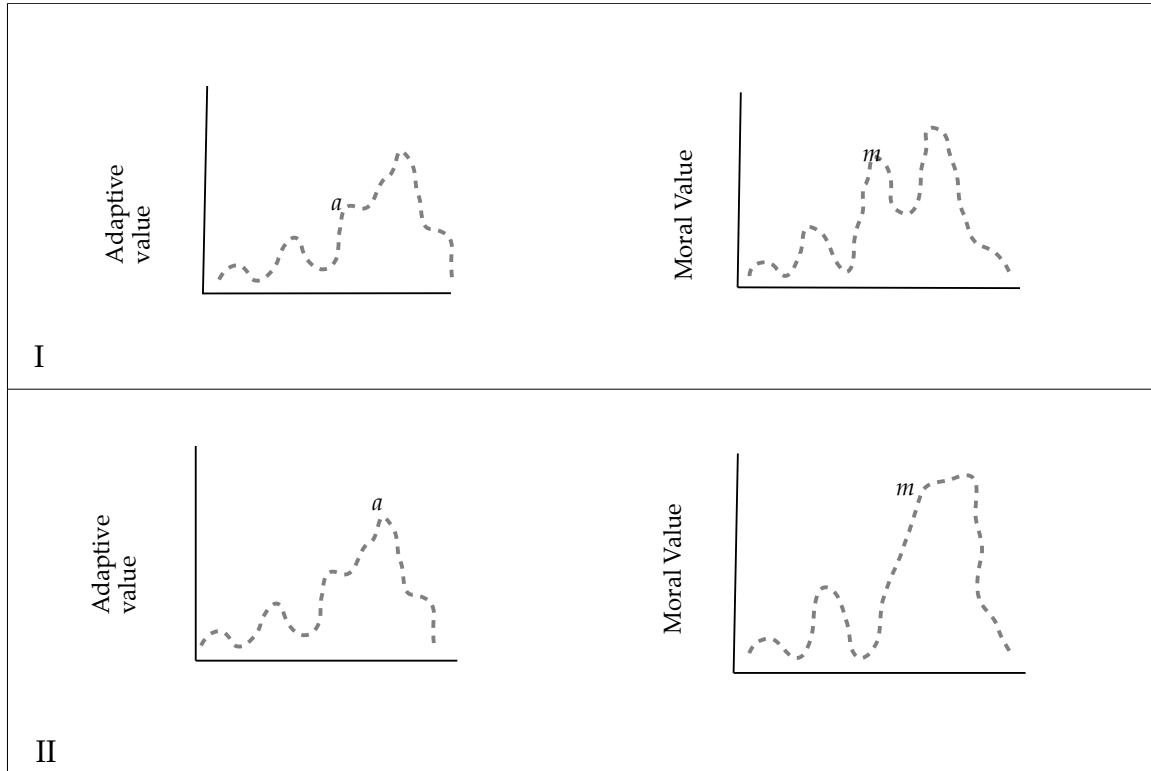


FIGURE 3: DANCING LANDSCAPES

5.3 Moral Deliberation and Predictive Modeling

Fundamental to the analysis of the previous section is the claim that all moral reform relies on predictive models, and utopian reform entirely on them (i.e., they cannot suppose correlation with present values). This may seem odd on a rule-based account of morality. On what we might call a strict “morality-as-rules” view,⁹¹ in reforming a moral order we simply examine some constituent rule r , and ask whether we approve of it, whether it is justified, or whether it is part of true

91. Sen has called this “institutional fundamentalism,” which he mistakenly believes is Rawls’s view. *The Idea of Justice*, pp. 82-3. I supposed some such view in “Social Complexity and Evolved Moral Principles.”

morality. If we decide that it is not, we look for some variant r^* that passes some test of moral justifiability, and replace r with r^* . We could, in principle, perform this deliberation in relation to any order, regardless of how different from ours. We could simply reflect on the rules and give our judgment (an all-too-common method in moral philosophy). For Hayek, of course, this is not sufficient; our ultimate concern is the order of actions that emerges from the system of rules. And that means that our decision whether to replace r with r^* depends on the application of a predictive model, which tells us what an order O^* with r^* would look like.

This is not really an unusual form of rule-based moral evaluation. John Rawls proposes a similar procedure in his interpretation of Kantian ethics.⁹² As is well-known, one of Kant's formulations of his famous categorical imperative requires that a person tests the maxim of her act by seeing whether she could will it as a law of nature. Rawls reinterprets this in terms of a four-step CI [Categorical Imperative] procedure. The first three steps on the CI procedure are fairly straightforward. One commences by adopting a maxim:

(1) I am to do X in circumstances C in order to bring about Y . (Here X is an action and Y a state of affairs).

The second step generalizes the maxim at the first to get:

(2) Everyone is to do X in circumstances C in order to bring about Y .

At a third step we are to transform the general precept at (2) into a law of nature to obtain:

(3) Everyone always does X in circumstances C in order to bring about Y (as if by a law of nature).⁹³

It is the fourth step that leads us to predictive modeling. We are to consider the "perturbed social world" that would result from the addition of this new law of nature and seek to understand the new "equilibrium state" on which this perturbed social world would settle. This social world emerges upon the new law of nature in

92. And, like Hayek, he believes it should be restricted to the neighborhood of the present order. See Rawls, *Justice as Fairness, A Restatement*, edited by Erin Kelly (Cambridge MA: Harvard University Press, 2001), p. 70.

93. John Rawls, "Themes in Kant's Moral Philosophy," in *John Rawls: Collected Papers*, edited by Samuel Freeman (Cambridge, MA: Harvard University Press, 1999), p. 499.

the context of the rest of the present order. We are then to ask ourselves whether, when we regard ourselves as a member of this new social world, we can “will this perturbed social world [order] itself and affirm it should we belong to it.”⁹⁴ The ultimate concern here is the social world — the order of actions — that would result under the assumptions of the model.

6 CONCLUSION

The uncompromising advocate of evolutionary ethics insists that our evolved positive morality *is* our current best approximation of true morality; seeking to depart from it by appeal to critical moral ideals can only lead to disaster. The uncompromising moral philosopher seeks to repel this assault on her domain by drawing a clear line in the sand — that between *is* and *ought*. Evolutionary analysis can explain what our positive morality *is* and how it came about, but this tells us nothing useful about what *ought* to be done. Indeed, leading philosophers have recently proclaimed that true morality and true justice *could* be permanently outside the motivational horizon of humans. Darwin's fundamental insight — that morality is a distinctive evolutionary achievement of humans that made possible our form of eusocial life — has, the uncompromising moral philosopher tells us, no real normative significance. Both are simple views: for both there is only one landscape to be traversed.

Thoughtful philosophers such as Kitcher struggle to capture the normative significance of the fact that our evolved morality makes human social life possible while preserving the ethical project as a conscious, critical, stance that can reform our evolved morality according to our normative ideals. That is clearly the fundamental task for evolutionary ethics — indeed, for any plausible moral theory. In this paper I have tried to take some initial steps to think a bit more carefully about how evolutionary and critical morality might be related. I believe that is this context NK models are helpful.⁹⁵ When we depict the problem in these terms, we can more

94. *Ibid.*, p. 500.

95. Although of course they are not the only way to analyze this difficult problem. When we employ a good model we see things differently, and so we see things that would have escaped us. But the model

clearly see the conditions under which adaptive and moral landscapes are plausibly correlated. And this, in turn, places us in a position to more rigorously deal with other recurring questions, about what an evolutionary analysis implies about the respect for our current morality and its reform based on our normative thinking. By focusing on Hayek I have tried, firstly, to bring out the surprising richness and subtlety of his analysis of evolved morality (things that are still not sufficiently appreciated) but, more importantly, to demonstrate that the sort of correlation I analyze is not a mere abstract possibility, but one which, if not precisely implicit in his thought, can be easily reached from it. Like all living things the Hayekian project is subject to evolutionary pressures; here, at least, I hope evolution and progress march together.

*Philosophy
University of Arizona*

can also blind us to important features of the phenomenon. Like any good tool, a good model helps greatly in the right context, but a one-tool toolkit quickly runs into problems. Clichés can be insightful: to one who only has a hammer, everything really can look like a nail.