

**Group Selection and Methodological Individualism:  
Compatible and Complementary**

Douglas Glen Whitman

Dept. of Economics

California State University, Northridge

18111 Nordhoff Ave.

Northridge, CA 91330

[glen.whitman@csun.edu](mailto:glen.whitman@csun.edu)

ABSTRACT: Various authors allege that the theory of group selection is inconsistent with methodological individualism, and therefore analysts must reject at least one of these principles. The present article argues for their compatibility. The meaning of methodological individualism is clarified, and the new version of group selection (articulated by Wilson and Sober (1994, 1998)) is explained. The two principles are then incorporated into a single methodological approach. Group selection affects the assumptions made about the kind of individuals who populate social scientific models, while methodological individualism requires models' conclusions to follow from the actions and interactions of those individuals.

JEL Classification Codes: B40, B52, B53, Z13

Key Words: methodological individualism, group selection, evolution

## **Introduction**

In his theory of cultural evolution, Friedrich Hayek (1991) explicitly adopts the notion of group selection as an evolutionary force that leads societies to adopt group-beneficial rules and practices. Group selection is a hypothetical process in which traits (whether genetic or cultural) evolve because members of groups that practice them experience greater reproductive success than members of groups that do not. In his earlier work (especially Hayek (1952)), Hayek emphasizes the importance of methodological individualism in the social sciences. Methodological individualism is the practice of explaining social phenomena in terms of the behaviors and interactions of individuals, rather than those of “collective” constructs such as groups, classes, or ideologies.

In an influential article, Viktor Vanberg (1986) argues that group selection and methodological individualism are incompatible, a position that has gathered the assent of other social scientists. Vanberg alleges a shift in Hayek’s thinking, “from the notion that behavioral regularities emerge and prevail because they benefit the individual practicing them, to the quite different notions [sic] that rules come to be observed because they are advantageous to the group” (1986, p. 83). Although Vanberg argues for scrapping group selection in favor of methodological individualism, others -- notably, Geoffrey Hodgson (1991) -- have lobbied for the opposite path of scrapping methodological individualism in favor of group selection. Hodgson points to “an inconsistency in Hayek’s work between, on the one hand, the ideas emanating from his individualist roots, and, on the other, his growing commitment to an evolutionary perspective. In an evolutionary context,

methodological individualism has to be either redefined or abandoned” (1991, p. 78). Hodgson agrees with Stephan Boehm that “Hayek is by no means the champion of methodological individualism that he claims to be” (Boehm 1989, quoted in Hodgson 1991, p. 78). Where Vanberg and Hodgson agree is that Hayek has erred by embracing *both* group selection *and* methodological individualism.

Some biologists, too, accept the proposition that group selection and methodological individualism cannot coexist. Wilson and Sober (1994, 1998), who have argued strenuously for the acceptance of group selection theory within evolutionary biology, have indicated an opposition to methodological individualism. After observing the prominence of group-level theorizing in the early social sciences, they note:

More recently, the group-level perspective has fallen on hard times, although to different degrees in different disciplines. Social groups are seen not as adaptive units in their own right but as by-products of individual self-interest. (Wilson and Sober 1998, p. 133)

Wilson and Sober also favorably cite D. T. Campbell’s denunciation of methodological individualism:

Methodological individualism dominates our neighboring field of economics, much of sociology, and all of psychology's excursions into organizational theory. This is the dogma that all human social group processes are to be explained by laws of individual behavior -- that groups and social organizations have no

ontological reality -- that where used, references to organizations, etc. are but convenient summaries of individual behavior.... We must reject methodological individualism as an a priori assumption, make the issue an empirical one, and take the position that groups, human social organizations, might be ontologically real, with laws not derivable from individual psychology.... (Wilson and Sober 1994, p. 599, ellipses in original)

Later, Wilson and Sober indicate a desire to “tear down the edifice of individualism” (1998, p. 330) by which they seem to mean both methodological individualism (as distinct from group functionalism) and psychological individualism (as distinct from altruism). They argue that methodological individualism owes its popularity in the social sciences to the desire to avoid the seeming mysticism of holistic approaches:

...[T]he reason for individualism’s growing attractiveness was not the confirmation of a novel empirical theory that entails that individualism must be true (Wright, Levine, and Sober 1992). The term methodological individualism is therefore apt, in that this version of individualism is not grounded in any particular scientific theory. (Ibid., p. 329)

Thus, Wilson and Sober view the evidence in favor of group selection as evidence against methodological individualism (which they think should be an empirical rather than methodological proposition). In short, proponents of group selection and proponents of

methodological individualism agree on one thing: you cannot have both. If you accept one, you must reject the other.

The ambition of this paper is to put to bed, at long last, the notion that methodological individualism (hereafter, MI) and group selection (hereafter, GS) are fundamentally incompatible. It is true that MI places some constraints on GS. Such conditions are fully consistent with the new GS theory as articulated by Wilson and Sober (1994, 1998), which resists the tendency of early GS theorists like V. C. Wynne-Edwards (1962) to assume group-beneficial evolutionary processes as the norm. But that conclusion is a far cry from Vanberg's dismissal of GS as "theoretically vague, inconsistent with the basic thrust of Hayek's [methodologically] individualist approach, and faulty judged on its own ground" (1986, p. 97), as well as Hodgson's claim that the use of both concepts constitutes "a major internal inconsistency in Hayek's work" (1991, p. 69).

In defending the compatibility of MI and GS, I am greatly assisted – and anticipated – by Zywicki (2000), which recognizes the viability of Sober and Wilson's (1998) revised version of GS and its applicability to Friedrich Hayek's (1979, 1992) theory of cultural evolution. I wish to expand on Zywicki's point by directly addressing the implications for MI of accepting GS.

The principle of MI requires that in a satisfying social scientific theory, the choices and behavior of individuals, as affected and constrained by the physical environment, must build up to the system effects at any given point in time. While the behavior of individuals is affected and constrained by the institutional environment as well, MI posits that the institutional environment itself can in principle be fully explained

by the behavior of past and present individuals (again, as affected by the physical environment). I will contend that MI, understood in this way, is indeed a *necessary* condition for a satisfying theory. But it is not *sufficient* for a good theory, because it severely underdetermines the assumptions theorists can make about the preferences and beliefs of agents. A good theorist must also ask what sort of individuals are likely to populate the system in question – a question that inevitably brings in an evolutionary standpoint. The evolutionary forces at work may include group-level as well as individual-level selection.

### **Apologia: A Brief Defense of Methodology**

It is not without trepidation that I enter a musty methodological debate, and I therefore offer the following justification for doing so. There is a popular and respectable view, in economics and many other sciences, that methodology is a distraction from the hard work of creating and testing hypotheses. Competing hypotheses should be judged by evidence, not by methodology. With respect to the GS controversy, for example, one might argue that the proof should be in the empirical pudding; the justification for GS is its predictive power, if any, not its methodological credentials. But I must disagree with this perspective, for two reasons. First, the definitions of terms such as “testing,” “evidence,” and even “hypothesis” are not self-evident, but methodology-laden. For instance, the hypothesis that “a species will exhibit characteristic X” assumes that a species is something that can be observed and studied, and likewise for characteristic X. Similarly, in the social sciences, the notion that “markets” or “societies” behave in

particular ways relies on definitions for those terms. Methodology affects what will be considered adequate definitions, and hence what will count as evidence for the hypotheses in question.

Second, methodology inevitably affects the direction of scientific research. If, for instance, GS were rejected on methodological grounds, then evolutionary researchers would be led to seek different sorts of explanations for observed phenomena. They might spend less time even looking for the kind of phenomena that might be explained by GS. If, on the other hand, GS were approved on methodological grounds, then researchers would be more inclined to search for phenomena that might be explained on those grounds and less inclined to look for alternative hypotheses. And the approach that says “methodology is not important, just let the evidence speak” is, itself, a methodology that in some measure approves of GS – and any other hypothesis, for that matter. If the pursuit of science were costless, this perspective would be unobjectionable: every potential hypothesis could be pursued without distracting from any other. But in the real world of scientific research, scarcity requires trade-offs. Methodology is the discussion about how our limited scientific resources would be most gainfully deployed. The question, then, is whether further research on GS is worthy of our effort and attention.

## **Explaining and Defending Methodological Individualism**

*The meaning and purpose of methodological individualism*

To understand the meaning and purpose of MI, it will prove useful to consider the arguments made in its defense by Hayek (1952). Importantly, this means I am adopting a particular definition of MI that others may not share. It is possible that some versions of MI are indeed incompatible with GS. However, since it is Hayek whom Vanberg (1986) and Hodgson (1991) accuse of having strayed from his own methodological program, it seems fair to consider Hayek's conception of MI.

Hayek's defense of MI, which he also refers to as the "individualist and 'compositive' method of the social sciences" (1952, p. 61-76), depends on two notions about the role of ideas in generating human action.

The first of these is the subjective character of social scientific data. The social sciences must deal regularly with data that do not "exist" objectively as do the data of the natural sciences. The reason is that the data themselves are concepts or ideas. Hayek offers the example of the words "hammer" and "barometer":

If the reader will attempt a definition he will soon find that he cannot give one without using some term such as "suitable for" or "intended for" or some other expression referring to the use for which it is designed by somebody. And a definition which is to comprise all instances of the class will not contain any reference to its substance, or shape, or any other physical attribute. An ordinary hammer and a steamhammer, or an aneroid barometer and a mercury barometer, have nothing in common except the purpose for which men think they can be used. (Hayek 1952, p. 44-45)

Other examples abound; “money,” “property,” “government,” and so on all have definitions that require reference to the (shared) perceptions and intentions of individuals.

There do not exist physical tests for their existence.

Hayek understands that “methodological collectivism,” the alternative to MI, relies upon the apprehension of subjective wholes:

Closely connected with the “objectivism” of the scientific approach is its methodological collectivism, its tendency to treat “wholes” like “society” or the “economy”, “capitalism” (as a given historical “phase”) or a particular “industry” or “class” or “country” as definitely given objects about which we can discover laws *by observing their behaviour as wholes*. While the specific subjectivist approach of the social sciences starts, as we have seen, from our knowledge of the inside of these social complexes, the knowledge of the individual attitudes which form the elements of their structure, the objectivism of the natural sciences tries to view them from the outside; it treats social phenomena not as something of which the human mind is a part and the principles of whose organisation we can reconstruct from the familiar parts, but as if they were objects directly perceived by us as wholes. (Hayek 1952, p. 92, emphasis added)

MI does not prevent the social scientist from apprehending the existence of wholes, but rather cautions against trying to derive the laws of their behavior based on alleged observations of the whole (note the emphasized part of the quotation above). Such caution is justified by the fact that we cannot, in fact, observe wholes directly; we infer

their existence through the apprehension of particular events and phenomena. For instance, a language is never observed in its entirety; instead, it is observed in innumerable individual acts of communication that appear to draw on similar vocabulary and rules.

The second notion behind Hayek's compositive method is the distinction between *constitutive* and *speculative* ideas (Hayek 1952, p. 61-65). Constitutive ideas are those ideas that actually motivate people in their choices and behavior. They include concepts such as money and law, preferences, and beliefs about how the world works. Such beliefs may or may not be valid, but they are still relevant because they affect behavior. For instance, a superstitious belief in the good fortune brought by rabbits' feet is real, in the sense that it actually leads people to produce and buy the feet of rabbits. Such ideas (whether valid, invalid, or neither) are called "constitutive" because they constitute the social phenomena that a social scientist wishes to understand and explain.

Speculative ideas, on the other hand, are the ideas people (particularly social scientists) use to *explain* social phenomena. They are theories, models, and more rudimentary concepts used to understand the operation of the social world. Speculative ideas often take the form of "meta-ideas," in that they are ideas about the existence and interaction of other ideas. The truth status of speculative ideas, unlike that of constitutive ideas, is of the essence for the social scientist, at least if he wants his theory to have explanatory or predictive power.

The two categories of ideas are not hermetically sealed from each other. Insofar as one's speculative ideas about how the social world works affect one's behavior, they are also constitutive ideas. And to that extent, social scientists must treat them as given

just like other constitutive ideas. For instance, if consumers believe an economic theory that predicts a recession, they may alter their consumption; if voters believe a politician's claim that protectionist policies will increase their purchasing power, they may cast their votes for that politician. But the social scientist must carefully distinguish between those ideas he accepts as constitutive (he thinks real people believe them) from those he accepts as speculative (he believes them himself, at least provisionally). Otherwise, he runs the danger of accepting popular but potentially false ideas as part of his own theory. Just as it would be wrong to conclude that rabbits' feet are lucky simply because some people think so, it would be wrong to conclude that a popular economic theory is correct because some people think it is.

The danger is a serious one, as it is not uncommon for people to have anthropomorphic ideas, which attribute person-like characteristics to non-human entities. What does "society" want? What are the interests of "the state"? Both questions assume that entities composed of many individuals may be characterized as having something akin to preferences that they will act to achieve. That assumption may or may not be justified; it is the job of the social scientist to discover when and if it is.

Hayek explicitly recognized the avoidance of as-yet-unjustified speculative ideas as the root of MI:

That he [the social scientist] consistently refrains from treating these pseudo-entities as facts, and that he systematically starts from the concepts which guide individuals in their actions, is the characteristic feature of that methodological

individualism which is closely connected to the subjectivity of the social sciences.  
(Hayek 1952, p. 64)

People have ideas, often shared ones. Those ideas motivate their behavior, and their behavior results in social phenomena. But the only ideas that actually motivate behavior are constitutive. Speculative ideas only motivate behavior when they *become* constitutive, as people believe them and act on them. MI is a mental device whose primary function is to force scientists to sharply distinguish (others') constitutive ideas from (their own) speculative ideas.

To put it another way, MI is a bulwark against the accidental intrusion of unproven hypotheses into one's initial assumptions. Employing some collective entity X in one's theory means assuming, implicitly, that X is both real and subject to certain behavioral regularities. That assumption may not be justified. By forcing the analyst to carefully distinguish his own ideas from those of agents in his models, MI prevents the social scientist from relying on speculative conclusions before they have been justified by an analysis of constitutive premises.

*Methodological individualism, scientific reductionism, and emergent properties*

MI is, strictly speaking, a social scientific perspective. Its central claim regards the proper mode of explaining human social phenomena, not all phenomena in all sciences. Admittedly, there exists an affinity between MI and scientific reductionism, but there is no necessary connection. MI contends that human social phenomena can be

explained in terms of human ideas, but it takes no position on whether all human ideas can be reduced to more basic phenomena such as electrical impulses, hormones, and genes.

Indeed, social scientists often pair MI with the doctrine of methodological dualism, which contends that social scientific methods differ fundamentally from those of the natural sciences.<sup>1</sup> In the social sciences, the analysts may begin with mental phenomena such as preferences and beliefs, and use these concepts to explain social phenomena, without *necessarily* knowing the physical processes that underlie or generate the preferences and beliefs. As Ludwig von Mises argues, “we do not know how external events -- physical, chemical, and physiological -- affect human thoughts, ideas, and judgments of value. This ignorance splits the realm of knowledge into two separate fields, the realm of external events, commonly called nature, and the realm of human thought and action” (1981, p. 1). Adopting this point of view does not mean that we are uninterested in the physical processes that (may) generate mental events – only that it is possible for social science to proceed without waiting for the natural sciences to discover the minute workings of the brain first. In short, MI neither implies nor denies reductionism. MI says social phenomena can and should be understood in terms of individual psychology; it does not take a position on whether individual psychology can be understood in terms of something even more basic.<sup>2</sup>

Nonetheless, methodological individualists and scientific reductionists have something in common: they share a dissatisfaction with explanations of macro phenomena that do not make explicit reference to the micro units that comprise them. The central question here is whether higher-order entities (that is, entities comprised by

smaller-scale entities and their relationships) exhibit *emergent properties*, and if so, whether all such emergent properties derive from the properties of the constituent entities and their interaction.

MI admits the existence of emergent properties of social systems (such as the price system, the state, the institution of money).<sup>3</sup> But MI insists that, at least in principle, those properties can be explained by the actions and interactions of individuals who take part in those systems. The ideas and behavior of individuals may be affected by institutional arrangements, such as firms and legislatures; but those institutions themselves can be explained by the ideas and behavior of present and past individuals.<sup>4</sup> Moreover, the MI theorist is suspicious of alleged emergent properties that have not yet been so explained. For instance, if someone asserts that a legislature's actions manifest a "general will" of the public<sup>5</sup> or the maximization of some well-defined social welfare function<sup>6</sup>, the MI theorist will be skeptical until a process is found by which a "general will" or social welfare function could emerge from individual choices and behavior. Arrow's Theorem (Arrow 1951) is the celebrated conclusion that no such social welfare function exists, if we require such a function to satisfy a number of intuitive conditions. Black's Median Voter Theorem (Black 1958), on the other hand, identifies some narrow conditions under which the preferences of a particular voter may be regarded as the preference set that guides legislative choices.

The scientific reductionist has similar, though not identical, concerns. He, like the methodological individualist, wishes to place foundations under emergent properties. He wishes to explain the behavior of molecules in terms of atoms and their interaction; he wishes to explain atoms in terms of subatomic particles and their interaction; and so on.

There is, however, an important difference: in the natural sciences, the existence of many emergent properties can be verified independently of their explanation.<sup>7</sup> A molecule can be isolated and observed under lab conditions, whether we know the relations of its component atoms (or even that such component atoms exist) or not. A social welfare function, on the other hand, cannot be observed. It does not exist in the same sense as a molecule; its reality, if any, is more akin to that of a logical deduction or a mathematical calculation. Without logical proof, we have no strong independent reason to believe it exists at all.

Yet emergent characteristics of social systems *do* exist. Prices do have a tendency to reflect scarcity. Black's Median Voter Theorem does provide a reasonable explanation for some kinds of legislative outcomes. In certain types of banking systems, such as those with fiat money, bad money does tend to drive out good (Gresham's Law). Far from denying the usefulness of such conclusions, MI embraces them *if* they have been justified in terms of the choices and behavior of individuals. If they have not yet been justified, the MI theorist will remain skeptical; they *might* be true, but they have the status of unproven hypotheses.

#### *Methodological individualism as a minimal requirement*

MI does not purport to provide sufficient conditions for good social science. Rather, it guards against one type of error. Other errors must be avoided by means of other principles, or just careful thought and research. Here, I wish to make four points about what MI does *not* require.

First, MI does not specify the content of agents' ideas. MI says the theorist should begin with a notion about the constitutive ideas held by the agents being modeled, including their preferences and beliefs, but it places no constraint on what type of preferences and beliefs they might have. A theory in which agents are assumed to love work and despise leisure might be criticized as unrealistic, but it would not violate MI.

It follows that MI does not specify whether agent preferences will be selfish, altruistic, or something else entirely.<sup>8</sup> This point deserves emphasis, because the association of MI with selfishness seems to have motivated Wilson and Sober's opposition to it.

This image of 'selfish individuals' may also [like the 'selfish gene'] be metaphorical, but it is more insidious because it can be taken literally. In other words, it is possible to believe that individuals really are intentional systems *motivated entirely by self-interest* and this is, in fact, the individualistic perspective that pervades the social sciences. (Wilson and Sober 1994, p. 601; emphasis added)

Wilson and Sober's conflation of MI with self-interest has been noted by Eric Alden Smith:

... MI per se does not entail an assumption of selfishness or self-interest. Put as simply as possible, MI asserts that the properties of any group result from the actions and interactions of its individual members (Elster 1983; Smith &

Winterhalder 1992); this can remain true even if these individuals are altruistic, dedicated to the collective good, and so on. (Smith 1994, p. 637)

Second, MI does not specify exactly how individuals make choices. They might be perfectly rational actors, boundedly rational actors, rule followers, or even automatons. Naturally, the assumptions made in this regard will have substantial consequences for the predictions made, and some assumptions might be criticized as unrealistic. Economists steeped in the rational-choice tradition are apt to criticize models that place too much emphasis on cultural programming while ignoring people's responsiveness to incentives (thus treating people as either rule followers or automatons). And that criticism might be valid – as might a criticism of economic approaches that assume rational agents with unreasonably high calculative abilities. On these matters, MI is silent.

Third, MI does not deny the existence of collective wholes. It simply requires their behavior to be explained in terms of their components' behavior, as holistic explanations run the risk of ascribing agency where it does not exist. “Even where individual actions and intentions are coordinated to achieve a collective outcome, there is no justification for attributing agency to the collectivity, as if it had eyes that see and hands that move” (Smith 1994, p. 637).

Fourth, MI does not require the exclusive use of spontaneous order or “invisible hand” explanations. Vanberg (1986) sees MI and spontaneous-order explanations as closely connected: “[Methodological individualism] is behind Adam Smith's notion of the ‘invisible hand’ as it is behind Adam Ferguson's conception of social institutions as

‘the result of human action but not of human design’ (Hayek, 1967, pp. 96 ff.)” (Vanberg 1986, p. 80). But while spontaneous-order explanations of social outcomes are generally consistent with MI, they are not required by it. An explanation that relies on a collective decision-making process, such as voting or consensus-building, is consistent with MI if the choices of individuals within the process can be explained.

For instance, consider the question of why people drive on one side of the road. An invisible-hand explanation would focus on the instability of any other proposed equilibrium. Rational individuals who know that 51 percent or more of the population tend to drive on the right will choose to drive on the right themselves – and as a result, everyone will eventually drive on the right (or the left, if the initial majority happens to lean that way). But this is not the only account consistent with MI. An explicit collective decision could also have done the job: the religious or political leaders of the society decreed that everyone should drive on the right. This is, indeed, what happened in Sweden in 1963 (Lucas 2002). There is nothing un-MI about this story, as long as we recognize that individuals make the laws and individuals follow them. The explanation of a collective choice on something more controversial than selecting the side of the road would need to include some reason why dissenters are willing (or forced) to go along with the outcome, as well as the reasons why the designated decision-makers made the choice they did.

Because MI has so often been used in conjunction with particular assumptions and approaches, including rationality, self-interest, and spontaneous order, we must be especially careful not to confuse them. MI’s association with other methodological

notions and social scientific perspectives is mostly attributable to the dominance of MI in economics, but the connection is historical, not definitional.

## **The Rejuvenated Theory of Group Selection**

### *The old group selection theory*

In its earliest manifestation, GS theory was little more than an assertion, backed by some empirical observations, that organisms tended to evolve group-beneficial traits – that is, traits yielding benefits in terms of long-run reproduction for all or most other members of a group, even while (possibly) creating costs for the individual organism expressing the trait. For instance, Wynne-Edwards (1962) observed a tendency of animal populations to limit their population density to the optimal level allowed by the physical environment, even when individual organisms within the population could benefit themselves by reproducing more at the expense of the rest.

If intraspecific [within-species] selection was all in favour of the individual, there would be an overwhelming premium on higher and ever higher individual fecundity, provided it resulted in a greater posterity than one's fellows.

Manifestly this does not happen in practice; in fact the reproductive rate in many species, and recruitment of adults in others, is varied according to the current needs of the population. (Wynne-Edwards 1962, p. 19).

Economists will recognize the situation described as a kind of tragedy of commons<sup>9</sup>; in essence, Wynne-Edwards observed that some animal populations appeared to have overcome the commons problem.

But how could the problem have been solved? Subsequent works in evolutionary biology, most notably Williams (1966), argued that it would *not* be solved because the group-level solution is vulnerable to subversion from within. Any trait that allows an individual organism to contribute more offspring to the next generation *relative to other organisms without the trait* will tend to grow in the population, and any trait that causes it to leave fewer offspring relative to others will thus tend to be pushed out. As a result, group-beneficial traits cannot survive in the long run, despite Wynne-Edwards's statement that when group-level and individual-level selection conflict, "group-selection is bound to win" (Wynne-Edwards 1962, p. 20).

This powerful critique led many evolutionary biologists to reject GS theory for much of the latter half of the century. Wilson and Sober (1994, 1998) have since resurrected GS theory. The new theory, while repudiating the notion that group-selection is "bound to win," elucidates some conditions under which it *can* win. I discuss the new GS below.

Here, however, I wish to observe that even the old GS can coexist, albeit uneasily, with MI. No one ever doubted that *if* some or all members of a group displayed certain types of traits (such as a tendency to limit offspring or to warn of predators), the entire group could benefit. The claim of those opposed to GS is simply that such "altruistic" members will cease to exist in the long run, because they will be weeded out. In the context of cultural evolution (the only context in which MI is applicable), this would be

equivalent to the claim that individuals could not possess certain kinds of preferences and beliefs, because such individuals would tend to be driven out by individuals with different preferences and beliefs. Now, that argument may indeed be true under some conditions, but it is not an argument required by MI. As discussed earlier, MI does not specify the content of agents' ideas.

An objection to the old GS based on MI can be constructed by observing a key difference between biological and cultural evolution. In biological evolution, traits are “built into” organisms, and therefore a trait will be expressed as long as there are individual organisms born with it in a given generation; whereas in cultural evolution, the individual organism may *choose* to adopt a different trait. Thus, an MI-consistent account must consider whether individuals will, in fact, choose to retain group-beneficial traits. When addressing this question, we should recall that MI does not require a specific decision-making mechanism (like rationality) for individuals; they could be rule-followers, for instance. Notably, Hayek's (1988, p. 21) theory of cultural evolution downplays the role of reason while emphasizing the human capacity for imitation, which lends support to a rule-following conception of human nature. This perspective may partially explain why Hayek willingly embraced the old GS.

The point here is not to defend the old GS. The point is that the error of the old GS was only tangentially related to MI. The real fault of the old GS was its failure to explain how a process of group-level selection could ever overcome the process of individual-level selection. That task was left to the new GS theorists.

*The new group selection theory*

In explaining and defending the new GS, it will be helpful to consider explicitly the mathematical reasoning offered by Sober and Wilson (1998). Much of Sober and Wilson's discussion focuses on the evolution of altruistic traits, i.e., traits which are both group-beneficial *and* create a cost to the individual organism. Although altruistic traits are not a necessary element of GS theory, they do pose the greatest challenge to that theory. (Later, I will consider treating altruism as a feature of social equilibria, rather than as a feature of traits.)

The simplest explanation for how GS could enable the evolution of altruistic traits relies on an application of Simpson's Paradox, which says that the proportion of observations with some property in an aggregate group may be smaller than the proportion of observations with that property in every subgroup comprising the aggregate. For instance (to use one popular example<sup>10</sup>), hospital A may save a larger fraction of all its patients than hospital B, even though hospital B saves a larger fraction of its elderly patients than hospital A *and* a larger fraction of its young patients than hospital A. This could occur because hospital B treats a disproportionately large number of elderly patients, who have a lower survival rate than young patients.

As applied to group selection, Simpson's Paradox indicates that an aggregate population might display an increase in the proportion of its members with an altruistic trait, even though each subgroup within the population experiences a decrease in the proportion of members with that trait. This occurs because the subgroups with more altruists contribute disproportionately to the aggregate population.

Suppose there exist two subsets of a larger population of some organism. Each subset has 100 members. Both subsets have some members who are disposed toward altruism: they will make personal sacrifices in terms of reproduction for the benefit of the group. But subset A has a larger fraction of altruists than subset B – say, 80 percent versus 20 percent. The overall population thus has 50 percent altruists.

Each subset will reproduce for one generation before being combined into a single larger population. In the meantime, the proportion of altruists in the each subset will shrink, because they contribute fewer offspring than do the non-altruists. To be specific, let us say each non-altruist contributes (on average) 10 percent more offspring than does each altruist. As a result, subset A's second generation will have 78.43 percent altruists (versus 80 percent before), and subset B's second generation will have 18.52 percent altruists (versus 20 percent before). Thus, the proportion of altruists in each subgroup declines.

Then the two subsets combine into a single population, but not in equal numbers. The group with a larger fraction of altruists will contribute disproportionately because they benefited more from the actions of the altruists. Suppose that the size of subset A increased by a ratio of 2 (to 200 members), while B increased by a ratio of only 1.2 (to 120 members). Then the combined population has 320 members, with 179 altruists (157 from A and 22 from B). The fraction of altruists in the aggregate population has increased, from an initial 50 percent before to just over 55 percent now. In short, the altruistic fraction of the whole group increased, even while the altruistic fraction of each subgroup decreased. This is Simpson's Paradox at work.

Two aspects of this process bear emphasis. First, the result depends crucially on the numbers chosen. If I had posited a wider disparity in the number of offspring contributed by altruists and non-altruists, then the fraction of altruists in the whole population could have decreased. If I had posited a smaller disparity in growth rates of the subsets, then again, the fraction of altruists in the whole population could have decreased. The point – one that the new GS theorists have emphasized – is that the importance of GS depends on the *relative strength* of within-group and between-group selection.

Second, without additional conditions being imposed, the process described only works for one generation. Suppose that, after recombining into a single population, the population never split up again. Then there would only be one group, and within-group selection would result in a diminishing fraction of altruists in each generation. Or suppose that the population did split up again, but the altruists were distributed equally among subsets, so that each group had the same fraction of altruists. Without variation in the representation of altruists, the groups would not contribute disproportionately to the whole population when they combined again.

In order to make the process work iteratively, we need additional conditions on the evolutionary process. Specifically, the altruists must periodically be reconcentrated so that some groups have greater representation of altruists than others. This can happen through chance if the groups are small enough. As an extreme example, suppose the whole population fissions into a large number of two-member groups – that is, mating pairs. Random mating will result in some fraction of two-altruist pairs, some fraction of a two-non-altruist pairs, and some fraction of half-altruist pairs. Variation thus results

automatically from a random process. This is the case of kin selection, encompassed as a special case of GS.<sup>11</sup>

Even if groups are somewhat larger than two members, there will still be random variation among groups in the representation of altruists. But as the “sample size” gets larger, groups come to resemble the overall population, with insufficient variation to cause differential contributions to the whole population due to altruism. If groups are large enough, then GS requires the existence of a sorting process whereby altruists have a better-than-random tendency to group together.

#### *Addressing potential objections to the new group selection theory*

What objections might an MI theorist have to the version of GS theory just presented? The difficulty with answering this question is that, as presented, the GS theory is one of natural science, whereas MI is a principle of social science. Nonetheless, it is not difficult to imagine an application of the argument to cultural evolution; indeed, Wilson and Sober (1998) do attempt to extend their approach to the cultural context (p. 149-158). In this section, I will grapple with some possible objections from the MI standpoint, keeping in mind the earlier observations about the actual content of the MI principle.

The first objection to the new GS theory, as with the old GS theory, is that it is implausible that any individual will be altruistic. This objection is easily dismissed: MI does not specify what sort of preferences and beliefs individuals must have. As noted earlier, MI has no necessary connection with the assumption of self-interest, despite often

being used alongside it. Altruism is one sort of preference an individual might have, perhaps as a result of biological evolution, perhaps as a result of cultural transmission of values. The critic of GS might argue that such preferences (however acquired) will tend to be weeded out when they appear. Whether the weeding process will always be effective, however, is the question at hand. To assert that it will is to beg the question. And even if the assertion turns out to be correct, it is not an assertion dictated by MI.

A second objection might be that individual incentives are being ignored. This turns out, however, to be a variant of the first objection. Incentives are inherently a matter of preferences, as what constitutes a sufficient incentive for one individual may not be sufficient for another. An individual that cares about maximizing its own offspring, and nothing else, will not, if it has a choice, be altruistic toward unrelated individuals. But MI does not specify that individuals must value producing their own offspring. To assert that they do – or more precisely, that they have traits leading them to act as if they do – is again to beg the question. The same rebuttal would apply to the assertion that individuals must value some specific goal other than maximization of offspring, such as maximization of a selfish utility function.

A third objection might be that the theory *asserts* the existence of group-level emergent properties. First, this is simply untrue, as the new GS theory recognizes that the conditions for group-level emergent properties will not always exist. A greater differential in reproduction between altruists and non-altruists, for example, could reverse the results and lead to the evolution of selfishness. The absence of a fission-and-regrouping process could also reverse the results. Second, as discussed earlier, MI does not deny the existence of emergent properties; it merely asks for their justification in

terms of the relations among lower-level entities. That is precisely what the Simpson's Paradox version of GS does: it explains how a specific type of natural selection under special conditions can lead to group functionalism.

A fourth objection might be that the conditions required for the process to work – a large enough reproductive advantage for subgroups with more altruists, a small enough reproductive advantage for non-altruists over altruists within subgroups, a process of fission and regrouping within a larger population – are unrealistic or too rare to be important. This, however, is an empirical objection, not an objection based on the principles of MI. One might make a similar objection against Black's Median Voter Theorem, on grounds that the assumption of a one-dimensional policy spectrum is unrealistic. Whether that's true or not, it does not mean the Theorem violates MI. MI is a necessary, not a sufficient, condition for a satisfying social scientific theory.

#### *A simple game theoretic extension*

While GS theory is often deployed to explain the (possible) existence of altruistic behavior, GS is not *defined* by altruism. GS may be manifested in cases where altruism is not required; in such cases, within-group selection may be weak or non-existent, and thus group-beneficial behavior evolves easily. Two simple games can be used to make this point.

Public good problems, and other situations in which selfish behaviors can lead to inferior results for all (and hence for the group), are often represented in the form of a prisoners' dilemma, as shown in figure 1. In this figure, S is a dominant strategy for both

participants in the game, and therefore S/S is the predicted equilibrium, even though both players would be better off if the outcome were A/A.

FIGURE 1 ABOUT HERE

FIGURE 2 ABOUT HERE

In figure 2, we have what is commonly called a coordination game. Just as in figure 1, both players are better off with A/A than with S/S. But in this game, there are two equilibria: A/A and S/S. Which equilibrium will occur depends crucially on the history of the game. In a biological context, a population could end up in the inferior S/S equilibrium simply because a large enough majority of the founding population's members had the S trait. In a choice context, no one would have an incentive to depart (unilaterally) from the S strategy. Absent group selection, there is no basis for assuming that either equilibrium would prevail. With group selection, however, we can imagine a process that leads to the emergence of the superior outcome: subgroups of the overall population could reach different equilibria, and those groups in the superior A/A equilibrium would outproduce or outperform those groups in the S/S equilibrium.

Figures 1 and 2 differ in just one respect: the payoff from playing S when the other player plays A. Anything that reduces the size of this payoff in figure 1 moves it closer to figure 2, until eventually figure 1 is transformed from a prisoners' dilemma into a coordination game.

How does such a transformation occur in the real world? The mechanisms are well known, even if the processes of transformation are not. In some prisoners' dilemma

situations, such as the overuse of a common resource (an example to be discussed more fully in the next section), privatization is a potential solution. But other solutions exist as well, one of the most common being the establishment of social norms to punish uncooperative behavior. The use of “trigger” strategies and similar devices can provide a means of enforcing cooperative behavior in repeated prisoners’ dilemmas, in which case the cooperative equilibrium and the uncooperative equilibrium of the repeated game are equivalent to the multiple equilibria of the coordination game.

If prisoners’ dilemmas can be transformed into coordination games, GS might seem superfluous in the explanation of how group-beneficial outcomes can evolve. That conclusion is unwarranted, for two reasons. First, as argued above, attaining the group-beneficial equilibrium in a coordination game is not to be taken for granted. Some groups might reach the superior equilibrium while others fail. (Some cartels succeed in enforcing high prices, while others do not.) Second, obeying norms that transform one game into another may itself sometimes involve a degree of unselfish behavior. If the members of a society are expected to impose a punishment on a norm-violator, the act of imposing the punishment can be costly to the person imposing it. In such a case, the enforcement of norms to solve a public good problem is itself a kind of public good; this is known as the second-order public good problem (Wilson and Sober 1998, p. 144). How can the second-order public problem be overcome?

The answer suggested by the new GS theory is that agents’ attitudes may be *internal* to the model. If some societies have members with inherent or ingrained preferences that reinforce group-beneficial norms – including punishment of non-cooperators, toleration and social support of property claims, and so on – those societies

will contribute a disproportionately large number of members to future generations relative to other societies. In cultural evolution, this effect can be exaggerated by the ability of individuals to change their allegiances. Individuals with the ability to choose among societies will naturally choose those that have managed to solve (or mitigate) their public-good problems.

Furthermore, some second-order public good problems involve less costly activities than the first-order public good problems they are intended to solve. The lower is the costliness of performing group-beneficial activities, the weaker will be the pressure of within-group selection relative to between-group selection. Wilson and Sober use this argument to explain the sharing of meat in hunter-gatherer societies, even though the act of hunting imposes disproportionate costs on the hunter, including the possibility of death. Such societies typically have a variety of rewards for hunters who do share their meat, such as privileged mate selection, and punishments for those who don't, such as ostracism.

Let's call hunting and sharing the primary behavior and the rewards and punishments that others confer on hunters the secondary behaviors (Ellickson 1991). ... The evolution of the secondary behavior (the rewards and punishments) can be analyzed in exactly the same way as the evolution of the primary behavior. By causing another individual to perform an altruistic primary behavior such as hunting and sharing, the secondary behavior increases the fitness of the entire group. At the same time, the secondary behavior is likely to require at least some time, energy, or risk for the individual who performs it. The

secondary behavior therefore requires group selection to evolve... (Wilson and Sober 1998, p. 143-144)

The broader point is that within-group selection interacts with between-group selection. Some genetic or cultural traits may evolve because they weaken or even eliminate within-group selection, thereby increasing the likelihood of group-beneficial traits evolving.

### **Integrating Evolution and Methodological Individualism**

#### *A combined approach*

As the foregoing discussion makes apparent, MI leaves a lot of things unspecified. It does not say what people believe, what people want, or what means they will use to achieve their goals. That may seem unsatisfying, but one principle cannot be expected to do all the work. MI is one methodological principle; it is not the only one.

In this section, I will outline a methodological approach that incorporates both MI and evolutionary forces – including, at least potentially, group selection. This approach is not original to this article; earlier statements (or at least hints) of this approach can be found in the work of Langlois (1983, 1985, 1986), Koppl and Whitman (2003), and most likely other works unknown to this author. The central notion of the approach is that any (satisfying) social scientific model will have two components: a story about what sort of individuals populate the model, and a story about how individual actions and interactions lead to the overall system effects predicted or explained by the model.

In the former story, evolutionary considerations take center stage, because they shape what kind of assumptions the modeler is justified in making about agents' preferences, initial beliefs, and mental capacities. Are individuals selfish or altruistic, or some combination of both? Are they rational, boundedly rational, or rule-following?<sup>12</sup> The answers to these questions depend on the selective mechanisms – or to use Langlois's (1986) language, "system constraints" – that agents face or have faced historically.<sup>13</sup> Some mechanisms have operated to shape all human beings at a genetic level in the distant past, such as the need to avoid predators, the need to procreate, the need to acquire sustenance, etc. Other mechanisms are more recent and culturally specific. One example of a culturally specific selective mechanism is the discipline of profit and loss in capitalist economies: firms that persistently make losses tend to be liquidated. As a result, existing firms are more likely to exhibit the characteristic of profit maximization, or at least loss avoidance, than they would be in the absence of the system constraint. (Notice that this selective constraint requires the existence of other social phenomena, such as money and property, whose existence may also demand social scientific explanation.)

The latter story, in which individual actions and interactions lead to the overall system effects, is where MI enters the picture. By deciding what sort of individuals populate his model, the modeler implicitly creates one or more ideal types (Whitman and Koppl 2003). The choices of agents in the model must follow from their assumed characteristics, combined with whatever constraints are imposed by the physical and institutional environment. Note again that the origin of specific institutions may also

require explanation, though not necessarily by the model in question – one scientist cannot explain everything!

MI is a minimal condition, satisfied with relative ease. A theorist could assume a variety of things about agents in his model, creating ideal types whose behavior follows from whatever characteristics he has assumed. But there is no guarantee that the predictions or explanations will be accurate, because the initial assumptions may be wrong. Suppose, for instance, that a theorist begins by assuming the existence of altruistic agents. If there are in fact not many altruists in the relevant population, then the model's predictions are likely to be inaccurate. MI is thus necessary but not sufficient for the construction of a fully satisfying model. In order to make a judgment about the likely accuracy of the model, we would need to know something about the system constraint. Only this kind of knowledge can indicate what sort of agents actually populate the social world to be explained.

GS fits into this framework as one kind of system constraint that might be, or historically might have been, effective in shaping the set of relevant agents. The crucial question is whether a GS process constitutes a strong or weak system constraint. The answer will depend on a variety of factors, including (a) the relative strength of the advantage conferred on individuals living in groups with members who perform group-beneficial (altruistic) tasks, (b) the relative disadvantage suffered by the individuals who perform those tasks, (c) the tendency, or not, of the population to separate into smaller groups that periodically regroup, and (d) the tendency, or not, of those people exhibiting group-beneficial characteristics to form groups with each other.

Inasmuch as GS may operate counter to individual-level selection, these two may be considered distinct system constraints that can wholly or partially offset each other. The stronger is one relative to the other, the more effective will be the weeding process, which allows the theorist to make more specific assumptions about the agents in his model. For instance, very strong individual-level selection combined with weak group-level selection would tend to weed out altruism, so in this context the theorist would be justified in assuming selfish agents. If the strength of the two forms of selection were reversed, the theorist would be justified in assuming altruistic agents. However, it is conceivable that the two forces could balance each other, leading to a polymorphic equilibrium wherein both types (selfish and altruistic) are substantially represented in the population. Indeed, Wilson and Sober (1994) suggest this is the most likely outcome: “We should expect a diversity of motives in the human repertoire that is distributed both within and among individuals” (p. 601). In this context, the theorist is compelled to make weaker assumptions: rather than assuming selfish or altruistic agents, he will have to allow for both.

### *Incorporating institutions*

The role of institutions in this structure deserves special attention. As noted above, the institutional environment constrains the actions of individuals. Typically, institutions involve patterns of actions or strategies adopted by current and previous members of society, which impose themselves on new members of a society by punishing (explicitly or implicitly) the adoption of alternative strategies. For example, imagine a

new member of a society in which people drive on the right side of the road. Although it is within the new driver's choice set to drive on the left, he nonetheless will choose to drive on the right. Any individual who chose to drive on the left would tend be eliminated from the society. And the fact that today's new drivers choose to drive right perpetuates the institution, so that tomorrow's new drivers will do so as well. Thus, the existing equilibrium creates a system constraint that affects both who is in the society and how they will choose to behave.

The same may be said, albeit with less confidence, about norms that have evolved to solve free-rider problems, such as respect for private property. Such norms may induce individuals to behave in a manner that appears "altruistic" even though the behavior is in fact motivated by self-interest. In this sense, we may say that altruism is not just a property of individual preferences (although it may be that as well), but is rather a property of a social equilibrium. The tendency to act in a group-beneficial manner may be transmitted from generation to generation not merely by genes or socialization, but by the incentive for new generations to conform their behavior to that of overlapping previous generations.

For MI to be satisfied within a single model, it is enough to show that, given the existing institutions, individuals will act in a way that sustains them (or, possibly, undermines them). Someone analyzing the behavior of American drivers need not explain why America is a society that drives on the right, only why new drivers' behavior will sustain that norm. But the MI approach insists that the emergence of an institution, such as right-side driving or enforcement of property rights in land, can *in principle* be explained in terms of the behavior of individuals in the past, as constrained by the

physical and institutional environment that faced them. Obviously, evolutionary forces can affect the evolution of institutions; GS, in particular, would be felt in the differential contribution of groups with different institutions to the overall human population.

In what follows, I will offer two examples of phenomena whose explanation might be assisted by the use of GS thinking: profit-maximizing firms, and norms for avoiding overuse of common resources. For both phenomena, I will offer speculative arguments about how GS might have been involved in their emergence. But I do not claim that these are phenomena are *in fact* attributable to GS; first, because other explanations may prove superior, and second, because further research would be required to show that the conditions for GS actually existed in these cases. The point here is to offer simple GS accounts of economic phenomena that are consistent with the strictures of MI. These two examples will also help show how the interplay of within-group and between-group selection can influence social outcomes.

*Example: the theory of the firm*

To what extent are we justified in assuming that firms act to maximize profits? Firms are, after all, collectives. A firm that acts to maximize profit will clearly prosper relative to other firms. But according to the critics of GS theory, that is not sufficient, as any collective may be vulnerable to subversion from within. As Zywicki (2000) argues,

Firms are mechanisms for carrying out team production. (Alchian and Demsetz 1972). The firm with the best survival prospects, therefore, is the one that

maximize[s] the team's joint product while minimizing the costs of this team production. But the creation of this team production structure opens the door for shirkers who seek to free ride on the efforts of the other members of the team. The hierarchical structure of a firm may be understood as a mechanism for minimizing these shirking costs so as to prevent free riding. (p. 85)

The standard critique of GS is that within-group selection will always outstrip between-group selection. If this critique were correct, then we might *never* be justified in assuming a profit-maximizing firm. But there is no reason to assume, a priori, that one form of selection is stronger than the other. If opportunities for shirking and free-riding within the firm are relatively small, and the discipline of profit and loss is relatively strong, then the outcome could indeed be profit-maximizing firms. Firms that fail to control their internal incentive problems will tend to be eliminated by competition with other firms; as a result, most employees will find themselves working in firms that address incentive problems with some measure of success.

It might be objected that firms typically implement specific structures designed to deal with shirking and free-riding, including stock options for managers, incentive pay, frequent monitoring, and so on. These structures make the "altruistic" behavior not truly altruistic; e.g., the frequently monitored worker has an incentive not to slack. Nonetheless, there is still a free-rider problem, because the enforcement of norms to solve a public good problem is itself a kind of public good; this is the second-order public good problem. If there are monitors whose job is to prevent shirking, why don't the monitors themselves shirk?

The obvious reply is that the monitors themselves are given incentives by their superiors, who are given incentives by their own superiors, all the way up the chain to the top of the organization. At the top we will find residual claimants – owners – who have a financial incentive to implement whatever cost-effective structures are needed to overcome shirking problems. Two responses are relevant here. First, the emergence of a residual claimant structure can itself be seen as an evolved response to the public good problem that arises from team production without a residual claimant. By reducing (but not eliminating!) the free-riding problem, residual claimants reduced the strength of within-firm selection, thus giving greater relative strength to between-firm selection. Firms with residual claimants outperformed those without them. Without the selective pressure of profit-and-loss, which operates at the firm level, residual claimants might not have been necessary. Second, while sole proprietorships have unitary residual claimants, what are we to make of corporations and partnerships? When responsibility is distributed, the possibility of free riding arises again. Since there exist multiple means of trying to control opportunistic behavior within the firm, and since there is an external constraint (the discipline of profit and loss) that rewards firms that control destructive internal behavior, we predict evolution of the most effective devices for internal control of incentive problems. GS favors practices that minimize the strength of within-group selection.

The firm is an instructive example, because it shows how emergent properties of higher-order entities can be used as a kind of shorthand. While the firm is not actually a single individual, for some purposes it may be usefully treated as an individual acting with the particular goal of maximizing profit. Profit maximization is an emergent

property of a system in which groups organized for team production compete in the context of strong profit-loss discipline. But it may not always be acceptable to make the assumption of profit maximization (consider firms in communist economies, where the discipline of profit and loss is absent). This is the constraint imposed by MI: the behavior of individuals within a firm must build up to the behavior of the firm itself. The willingness of economic theorists to look within the firm to examine internal incentive problems is the hallmark of the MI approach. But that is not *sufficient* for a compelling theory of the firm, as we should also enquire into the evolution of the firm's internal culture, which guides the attitudes and behaviors of the individuals who comprise it. GS theory provides a reason why firm-beneficial attitudes persist: those firms that failed to inculcate such attitudes were less likely to survive.

The application of GS to the theory of the firm is far from unproblematic. One condition for the new GS to work in biology is the differential reproduction of groups; what is the cultural analog of reproduction in the case of firms? Since firms are potentially infinitely lived, "reproduction" might have to be interpreted as persistence. Or reproduction might take the form of firms that imitate the structures and practices of existing firms. I will not attempt to answer here whether these conditions are sufficiently similar to the conditions necessary for GS in biological evolution. Profit-maximizing firms *might* be seen as the outcome of selection among groups, but other explanations might perform better. For example, residual claimants could be the accidental by-product of private property, an institution that evolved earlier and possibly for different reasons.

*Example: the tragedy of the commons*

The tragedy of the commons (Hardin 1968) is a classic case in which individually rational behavior leads to an undesirable outcome for the group (or more precisely, for each member of the group). The over-exploitation of common assets is not just a human phenomenon, as animal populations may overburden their local environments by excessive reproduction. Indeed, this was precisely the problem addressed by V. C. Wynne-Edwards (1962) in his early version of GS theory. Wynne-Edwards did not, like Wilson and Sober, ground his theory in an understanding of the interplay between within- and between-group selection. Nonetheless, the commons tragedy provides fertile ground for the application of the new GS theory. To what extent can populations of animals or humans be expected to display group-beneficial strategies in their use of assets? The answer to this question bears on the validity of, for instance, Harold Demsetz's (1967) theory of property rights, which asserts that "property rights arise when it becomes economic for those affected by externalities to internalize benefits and costs" (p. 97).

In the human context, the most commonly recognized solution to the commons problem is to privatize the asset in question (say, land), so that individual members of the population have exclusive stewardship over smaller plots. This solution is not as simple as it might appear, as it typically requires some kind of enforcement apparatus. Pure self-enforcement might be sufficient, and if it were, privatization would occur on its own. But if some kind of social enforcement is necessary, then a second-order public good problem arises: why should individual members contribute to the costly enforcement process by refraining from good opportunities to steal, refusing to trade with known property rights violators, and so on?

In this context, an opponent of GS theory might conclude that privatization will not occur. But that conclusion would only follow from a situation in which between-group selection is too weak relative to within-group selection. It might turn out, on the contrary, that the individual incentive to deviate from property-enforcing norms is relatively small compared to the gains from cooperation. If so, then societies composed mostly of individuals with an inherited or ingrained inclination to follow those norms will prosper relative to other societies. They would contribute a disproportionate number of members to future societies, thereby spreading the property-enforcing norms.

Of course, privatization is not the only potential solution to the commons problem. Another possibility is the enforcement of norms that directly limit individuals' consumption of the common resource. Such norms are common in primitive societies. Their enforcement, just like the enforcement of property-norms, may be subject to a second-order public good problem; and yet they persist. One possible explanation is that the enforcement of group-beneficial norms is eased by the inherent tendency of individuals to detect and punish people perceived as "cheaters" – a tendency reinforced by a GS process.

For an analysis of the tragedy of the commons to satisfy MI, the social outcome – whether overuse or efficient use of assets – must result from the attitudes of individuals and the constraints they face. If individuals are entirely selfish, then we do not predict a solution to the commons problem except in the presence of small numbers and very low monitoring costs. If individuals are entirely altruistic, then we predict commons problems to be solved with little effort. If individuals have both selfish and altruistic attitudes, then we predict commons problems to be solved (sometimes) through a

combination of monitoring and individual forbearance. But which of these assumptions about individual attitudes is most appropriate? In other words, what sort of agents should populate our economic models? Here, evolutionary theory, and GS theory specifically, can be of assistance. The stronger was between-group selection in the shaping of human psychology, the more justified we are in assuming some degree of altruism in the current population of humans.

As in the case of firms, the application of GS to the tragedy of the commons is not unproblematic. Among other things, there is the historical question of whether private property did, in fact, emerge at a time when human populations consisted of small competing groups, and whether those groups exchanged members in a fashion approximating the fission-and-reconcentration pattern required for group selection. In addition, even if norms for the avoidance of the tragedy of the commons (such as private property) could emerge from a GS process, other processes could possibly explain the same transformation. Such processes could include deliberate or non-spontaneous decision-making, such as the choice by a legislature to establish private property rights in assets previously owned in common. Even if GS explains the emergence of private property in some cases, it does not explain it in all cases.

## **Conclusions**

The notion that MI and GS conflict derives from an overextension of both principles. MI is taken to imply self-interest, rationality, spontaneous order, and denial of the existence of groups. GS is taken to imply that groups will *always* exhibit group-

beneficial characteristics. Both characterizations are mistaken. MI is a relatively minimal requirement that says group properties must be built up from individual actions and interactions, and GS (at least in its modern form) says that groups will exhibit group-beneficial characteristics only under specific circumstances that do not always obtain.

In this article, I have offered *speculative* discussions about incentive problems in the context of firms and common-pool resources. I have suggested, in each case, that GS could be invoked as a (partial) explanation of how such problems are solved. But the main point of the article – that there is no contradiction between GS and MI – does not depend on the correctness of these examples.

Once MI and GS are properly defined and contained, the appearance of conflict disappears. Both can be incorporated into a single methodological approach based on individual choice in the presence of system constraints. MI dictates that individual choices lead to social outcomes; GS is one force (among many) that determines what sort of individuals are present in the system. To accept this position, one need not agree that GS is a *significant* force in biological or cultural evolution. That is ultimately an empirical question, one that should not be foreclosed by the rejection of GS theory on erroneous methodological grounds.

---

---

## Endnotes

<sup>1</sup> Later, Hayek (1967) adopts a more nuanced approach, wherein the dividing line is not between natural and social sciences, but between simple and complex phenomena. As a result, even some natural sciences might have to adopt a form of methodological dualism, because their basic elements cannot easily be reduced to more constituent parts. “Where we have to deal with a complex situation in which observation discloses only very limited regularities, be it in the ‘applied’ branches of physics or in biology or in the social sciences, we usually ask to what extent our existing knowledge of the forces at work, or of the properties of some of the elements of the complex, may account for what we observe. We endeavor to find out whether this may be derived by deduction from what we know about the behaviour under simpler conditions of some of the factors involved” (Hayek 1967, p. 10). However, there are other features of the social sciences that distinguish them from all natural sciences: “Another and perhaps more important peculiarity of the social sciences is due to the fact that here the recognition of the different kinds of facts rests largely on a similarity between the observer and the observed persons” (Hayek 1967, p. 18, note 15).

<sup>2</sup> To be more precise, MI says social phenomena should be reduced to individual psychology (and environmental factors), while MD says no *further* reduction is required. Elsewhere (Hayek 1952b), Hayek argues that it may be *impossible* to reduce human action to something even more basic. But this position is distinct from both MI and methodological dualism, neither of which takes any position on this matter.

---

<sup>3</sup> Langlois (1983) refers to the position that emergent properties of social systems do not exist as “naïve methodological individualism,” to distinguish it from the version of MI which seeks the explanation of emergent properties. “Far from denying the importance of emergent phenomena, the economist and philosopher Friedrich A. Hayek ... reminds us that the entire *objective* of the social sciences is to explain how the behavior of individuals leads to orderly patterns and institutions that none had consciously planned – to explain, in other words, the emergent results of individual action” (Langlois 1983, p. 583). However, as I argue below, MI does not *require* the absence of conscious planning.

<sup>4</sup> Hayek (1952) unfortunately does not make this point explicit. It is clear, however, that many of the “wholes” that Hayek says require compositive explanations are institutions that individuals may take as given, such as money and language. See, e.g., p. 71: “It makes no difference for our present purpose whether the process [of emergence of an order] extends over a long period of time, as it does in such cases as the evolution of money or the formation of language, or whether it is a process which is constantly repeated anew, as in the case of the formation of prices or the direction of production under competition.”

<sup>5</sup> See, generally, Rousseau (1762).

<sup>6</sup> See Arrow (1951) and subsequent literature.

<sup>7</sup> But this is not always true even in the natural sciences. To the extent that they deal with complex emergent phenomena that cannot be observed directly, some natural sciences may face a problem similar to that of the social sciences. See footnotes 1 and 2.

---

<sup>8</sup> Not to imply that altruism and selfishness are *only* properties of preferences. They might also be properties of institutions or social equilibria. See the final section on “Integrating Evolution and Methodological Individualism.”

<sup>9</sup> In fact, Wynne-Edwards illustrated the basic problem with the case of overfishing, an example also used by Hardin (1968) to illustrate the tragedy of the commons.

<sup>10</sup> See, for instance, Zhang (n.d.).

<sup>11</sup> “As we have seen, however, kin selection is a special case of a more general theory – a point that Hamilton (1975, 1987) was among the first to appreciate. In his own words, ‘it obviously makes no difference if altruists settle with altruists because they are related ... or because they recognize fellow altruists as such, or settle together because of some pleiotropic effect of the gene on habitat preference.’” (Wilson and Sober 1998, p. 134).

<sup>12</sup> This sentence is not meant to indicate that rule-following is necessarily irrational, only that it might be.

<sup>13</sup> Langlois (1985) also refers to the notion of a “compositional principle,” such as selection or an invisible hand principle, that leads from individual actions to social results. Langlois’s approach is similar, though not identical, to that taken here.

## **Acknowledgements**

I wish to thank Roger Koppl, an anonymous referee, and all of the commentators for their helpful analysis and suggestions.

## References

- Agassi, J., 1975. Institutional Individualism. *British Journal of Sociology* 26, 144-155.
- Alchian, A. A., Demsetz, H., 1972. Production, Information Costs, and Economic Organization. *American Economic Review* 62, 777-795.
- Arrow, K. J., 1951. *Social Choice and Individual Values*. John Wiley and Sons, New York.
- Arrow, K. J., 1994. Methodological Individualism and Social Knowledge. *American Economic Review Papers and Proceedings* 84, 1-9.
- Black, D., 1958. *The Theory of Committees and Elections*. Cambridge University Press, Cambridge.
- Boehm, S., 1989. Hayek on Knowledge, Equilibrium and Prices: Context and Impact. *Wirtschaftspolitische Blätter* 36, 201-13.
- Caldwell, B., 2001. Hodgson on Hayek: A Critique. *Cambridge Journal of Economics* 25, 539-53.

Demsetz, H., 1967. Toward a Theory of Property Rights. *American Economic Review* 57, 13-27. Reprinted in Katz, A. W. (Ed.), 1998, *Foundations of the Economic Approach to Law*. Foundation Press, New York, 93-102.

Ellickson, R. C., 1991. *Order Without Law: How Neighbors Settle Disputes*. Harvard University Press, Cambridge, Massachusetts.

Elster, J., 1983. *Explaining Technical Change: A Case Study in the Philosophy of Science*. Cambridge University Press, Cambridge.

Hardin, G., 1968. The Tragedy of the Commons. *Science* 162, 1243-1248.

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

Hayek, F. A., 1952a. *The Counter-Revolution of Science: Studies on the Abuse of Reason*. LibertyPress, Indianapolis.

Hayek, F. A., 1952b. *The Sensory Order*. University of Chicago Press, Chicago.

Hayek, F. A., 1967. Degrees of Explanation, in: Hayek, F. A. (Ed.), *Studies in Philosophy, Politics and Economics*. University of Chicago Press, Chicago.

Hayek, F. A., 1979. *Law, Legislation and Liberty, Volume 3: The Political Order of a Free People*. University of Chicago Press, Chicago.

Hayek, F. A., 1991. *The Fatal Conceit: The Errors of Socialism*. University of Chicago Press, Chicago.

Hodgson, G. M., 1991. Hayek's Theory of Cultural Evolution: An Evaluation in Light of Vanberg's Critique. *Economics and Philosophy* 7, 67-82.

Koppl, R., Whitman, D. G., 2004. Rational-Choice Hermeneutics. *Journal of Economic Behavior and Organization*, forthcoming.

Langlois, R., 1983. Systems Theory, Knowledge, and the Social Sciences, in: Machlup, F., Mansfield, U. (Eds.), *The Study of Information: Interdisciplinary Messages*. John Wiley, New York.

Langlois, R., 1985. Knowledge and Rationality in the Austrian School: An Analytical Survey. *Eastern Economic Journal* 9, 309-330.

Langlois, R., 1986. The New Institutional Economics: An Introductory Essay, in: Langlois, R. (Ed.), *Economics as a Process: Essays in the New Institutional Economics*. Cambridge University Press, Cambridge.

Lucas, B., 2002. Which side of the road do they drive on? Web page:

[http://www.travel-library.com/general/driving/drive\\_which\\_side.html](http://www.travel-library.com/general/driving/drive_which_side.html), accessed 8

April 2003.

Mises, L. von, 1981 [1957]. Theory and History: An Interpretation of Social and Economic Evolution. The Ludwig von Mises Institute, Auburn University, Alabama.

Mises, L. von, 1963. Human Action: A Treatise on Economics, 3<sup>rd</sup> Edition. Contemporary Books, Inc., Chicago.

Schutz, A., 1962 [1954]. Concept and Theory Formation in the Social Sciences, in: Natanson, M. (Ed.), Collected Papers I: The Problem of Social Reality. Martinus Nijhoff, the Hague.

Schutz, A., 1964a [1943]. The Problem of Rationality in the Social World, in: Broderson, A. (Ed.), Collected Papers II: Studies in Social Theory. Martinus Nijhoff, the Hague.

Schutz, A., 1964b [1946]. The Well-Informed Citizen, in: Broderson, A. (Ed.), Collected Papers II: Studies in Social Theory. Martinus Nijhoff, the Hague.

- Schutz, A., 1964c [1960]. The Social World and the Theory of Social Action, in:  
Collected Papers II: Studies in Social Theory, Broderson, A. (Ed.), Collected  
Papers II: Studies in Social Theory. Martinus Nijhoff, the Hague.
- Schutz, A., 1967 [1932]. The Phenomenology of the Social World. Translation by  
Walsh, G., Lehnert, F. Northwestern University Press, Evanston, Illinois.
- Smith, E. A., Winterhalder, B., 1992. Natural Selection and Decision-Making: Some  
Fundamental Principles, in: Smith, E. A., Winterhalder, B. (Eds.), Evolutionary  
Ecology and Human Behavior, Aldine de Gruyter.
- Smith, E. A., 1994. Semantics, Theory, and Methodological Individualism in the Group-  
Selection Controversy. Behavioral and Brain Sciences 17, 636-37.
- Vanberg, V., 1986. Spontaneous Market Order and Social Rules: A Critical  
Examination of F. A. Hayek's Theory of Cultural Evolution. Economics and  
Philosophy 2, 75-100.
- Whitman, D. G., 1998. Hayek contra Pangloss on Evolutionary Systems. Constitutional  
Political Economy 9, 45-66.
- Williams, G. C., 1966. Adaptation and Natural Selection: A Critique of Some Current  
Evolutionary Thought. Princeton University Press, Princeton, New Jersey.

Wilson, D. S., Sober, E., 1994. Reintroducing Group Selection to the Human Behavioral Sciences. *Behavioral and Brain Sciences* 17, 585-608.

Wilson, D. S., Sober, E., 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Cambridge, Massachusetts.

Wright, E., Levine, A., Sober, E., 1992. Marxism and Methodological Individualism, in: *Reconstructing Marxism – Essays on Explanation and the Theory of History*. Verso Books, London, 107-128.

Wynne-Edwards, V. C., 1962. *Animal Dispersion in Relation to Social Behavior*. Oliver & Boyd, Edinburgh.

Zhang, J., n.d. What is Simpson's Paradox? Web page, <http://www.ma.iup.edu/~zhang/simpson.html>, accessed 8 April 2003.

Zywicki, T., 2000. 'Was Hayek Right About Group Selection After All?' Review Essay of *Unto Others: The Evolution and Psychology of Unselfish Behavior*, by Elliot Sober and David Sloan Wilson. *Review of Austrian Economics* 13, 81-95.

**Figure 1: Prisoners' Dilemma**

		Player 2	
		<b>A</b>	<b>S</b>
Player 1	<b>A</b>	5, 5	1, 6
	<b>S</b>	6, 1	2, 2

**Figure 2: Coordination Game**

		Player 2	
		<b>A</b>	<b>S</b>
Player 1	<b>A</b>	5, 5	1, 1
	<b>S</b>	1, 1	2, 2