

Charles Darwin Meets *Amoeba economicus*: Why Natural Selection Cannot Explain Rationality

Elias L. Khalil¹

ABSTRACT

Neo-Darwinists must recognize that organisms show behavioral flexibility in ways that cannot be explained by “hard-wired” genotype. Behavioral flexibility is nothing but rationality. How can we explain the origin of rationality, though? Neo-Darwinists assume that rationality is a trait. But this cannot be the case. Let us suppose two lineages of rationality, R1 and R2. Natural selection would favor R1 over R2 under the C1 constraints and *vice versa* under the C2 constraints. Under either constraint regime, agents would maximize fitness along the same kind of rationality, R*. But how could R*=R1 and R*=R2, when R1≠R2?

Key words: *Organismus economicus*; *Organismus automaton*; selection optimization; rationality optimization; Rationality-qua-Trait Thesis; incoherence problem; Rationality-qua-Method Thesis

JEL: B52; D0

¹ Email: elias.khalil@buseco.monash.edu.au. Department of Economics, Monash University, Clayton, VIC 3800, Australia. An earlier version appeared in the papers of the Evolutionary Economics Group (#2006-22), Max Planck Institute of Economics. The title of this paper is inspired by the title of Brian Skyrms’s [1994] paper. This paper was supported by the Konrad Lorenz Institute for Evolution and Cognition Research (Altenberg, Austria), the Max Planck Institute of Economics (Jena, Germany), and Monash University’s *Faculty Research Grant Scheme*, 2006. An older version received comments from Richard Posner, Ulrich Witt, Gerhard Müller, Werner Callebaut, Steven Orzack, Steve Abedon, Jack Vromen, Brian Charlesworth, Gordon Tullock, Timothy Crippen, Michael Ghiselin, Howard Margolis, Robert Axelrod, Richard Levins, Richard Nelson, Joseph Lopreato, R. Preston McAfee, J.S. Metcalfe, Peter Taylor, Elliott Sober, Stanley Salthe, Casey Mulligan, Franz Weissing and seminar participants at the Konrad Lorenz Institute. The current version received comments from Brian Skyrms, Jack Vromen, Michael Ghiselin, Yew-Kwang Ng, Richard Posner, Paul Griffiths, Ulrich Witt, Avi Waksberg, Martin Burd, Ellen Larsen, Deby Cassill, and seminar participants at Monash University. The paper benefited greatly from the assistance of Michael Dunstan. The usual caveat applies.

What does an amoeba do on an average day? It simply cannot afford to sit around and hope that some manna will fall from heaven. Nor can it, given the high cost of motion, afford to roam the neighborhood randomly in the hope of encountering some yeast. The amoeba must thus make efficient decisions: When it moves in search of nutrients by using its “false feet,” it moves via chemotaxis carefully towards higher food concentration gradients and away from (with some slight exceptions) toxic environments. However, when starved, these unicellular, solitary organisms have a strong incentive, or what biologists and psychologists call “signal” or “stimulus,” to undertake an appropriate response: Economists call such a response, when performed by a number of organisms together, “collective action.” When about 10^4 – 10^5 amoeba act together in such a fashion, they form something equivalent to what economists call a “club” [Eichinger *et al.*, 2005], as in the sense of a “club” or a “society” (but not an organism) [Strassmann *et al.*, 2000]. The cells that constitute the club, which looks like a slug, undergo differentiation of functions that allows the slug to act as a “bus.” The “bus” affords a more effective transportation to a new environment that would have a higher food concentration.

All organisms have to work, and they better work carefully as they search for nutrients, sexual partners, suitable habitats, and safety. All organisms must deliberate and choose among alternatives. They face an environment that is neither uniform across space nor static across time. All organisms must make decisions, and make the best decisions possible given that each decision involves a tradeoff. For neoclassical economics, the tradeoff entails that every choice, when a resource is fungible, involves a cost in the sense of an opportunity cost. To migrate to another region might afford more protection, but it involves an opportunity cost that can be unjustified by the meager marginal protection. If the making of the best decision possible allows

us to call human agents “*Homo economicus*,” it should equally allow us to call other living agents “*Organismus economicus*” or, in the above case, “*Amoeba economicus*.”

The idea that organisms are rational, called the “*Organismus economicus*” hypothesis, stands in stark contrast to the old-fashioned idea of organisms as programmed or self-operating machines, which shall be named the “*Organismus automaton*” hypothesis. If the *Organismus automaton* hypothesis is, at least for humans, unwarranted, can the theory of natural selection explain *Organismus economicus*? Natural selection can explain how a fitter particular trait in comparison to other lineages of the same trait becomes optimally spread in a population. Can it, however, explain rationality in the same manner? This is the central question of this paper.

The mention of “rationality” usually conjures the discipline of economics. The mention of “natural selection,” on the other hand, usually conjures the discipline of evolutionary biology. Thus, both terms are seen as being distinctly separate. This is a mistake. As detailed elsewhere [Khalil, 2007e], biologists use, without explicitly stating so, the concept of rationality when they discuss the fitness of a given behavior. Furthermore, economists use, without also explicitly stating so, natural selection when they discuss market equilibrium. So, we need not compare economics and biology *qua* disciplines, which usually implies that they use separate conceptual frameworks. (In any case, such comparisons have been undertaken [e.g., Hirshleifer, 1977; Hodgson, 2004, 2007].) The focus of this paper is rather on whether one kind of optimization used by natural selection theory, called “selection optimization,” can explain another kind of optimization used by rationality theory, called “rationality optimization.”²

Most biologists, at least until recently, generally shun away from the notion that

² Another paper [Khalil, 2007e] discusses in detail the difference between “selection optimization” and “rationality

organisms, including humans, are involved in “rationality optimization.” They usually embrace the *Organismus automaton* hypothesis. To clarify the conceptual issues, section 1 contrasts the *Organismus economicus* and the *Organismus automaton* hypotheses. Section 2 and 3, then, define respectively the terms rationality and natural selection. Next, section 4 discusses the notion of bounded optimization, in both rationality theory and natural selection theory, and then argues that the boundedness of rationality is not really a challenge to the optimization notion. Section 5 argues that evolutionary psychology, Herbert Simon’s process notion of rationality, or pragmatic institutionalism is also not really a challenge to rationality optimization.

These conceptual clarifications in sections 1 to 5 provide the basis for the argument, summarized in Figure 1. Namely, selection optimization is challenged by the

**Selection Optimization: *Organismus economicus* →
Rationality-qua-Trait Thesis: *Incoherence Problem* →
Rationality-qua-Method Thesis: *Uncaused Cause?***

Figure 1: The Argument

Organismus economicus (rationality) anomaly. To solve the rationality anomaly, natural selection theory, as Section 6 shows, must recognize rationality and explain it as being like any other trait. Such a “Rationality-qua-Trait Thesis” is well known to economists, imported from evolutionary biology long ago by Armen Alchian [1950] and Milton Friedman [1953]. The thesis basically argues that less rational agents are deleted by selection forces. Thus, surviving firms must be the most rational, assuming that the selection forces are sufficiently “sharp” to do so.

optimization,” and how the two concepts cut across the disciplines of economics and evolutionary biology.

However, this is questioned in Section 7, and especially Section 7.2, which is the *core of this paper*. It demonstrates why natural selection optimization cannot explain rationality optimization if we treat rationality as a trait. Following neo-Darwinism, an irrationality lineage could be, under some external constraints, the fittest trait. But to be the fittest, the organism must be maximizing some objective function, which actually makes the organism rational. This leads to a self-contradiction, similar to E.C. Escher's paintings, called here the "incoherence problem."

To avoid the incoherence problem, one must reject the Rationality-qua-Trait Thesis and propose, as Section 8 suggests, that rationality is rather a *method*, i.e., a way to organize the employment of different traits. As a method, rationality cannot be the subject of evolution. This raises the issue about the origin of rationality. Although this is an important question, it is not the focus of this essay.

The answer to this question is not the focus of this paper. However, it concludes by showing the ramifications of the "Rationality-qua-Method Thesis." If rationality cannot be caused by natural selection, this means that the natural selection paradigm has been oversold. This should open niches for alternative views of evolution. This includes, especially, evolutionary developmental biology, known as the "EvoDevo" approach [e.g., Müller & Newman, 2003]. It also includes the learning/developmental approach in economics as promoted by the Austrian school [e.g., Witt, 2003] and the Schumpeterian tradition [e.g., Schumpeter, 1989; Nelson & Winter, 1982; Dosi *et al.*, 1988; Nooteboom, 2000].

1. *Organismus economicus vs. Organismus automaton*

Until recently, most behavioral and biological scientists have fully embraced the *Organismus automaton* hypothesis. This hypothesis portrays behavior as the product of inherited programs that are oblivious to incentives. Interestingly, these scientists, except for orthodox behaviorists, would concede that, generally speaking, *Homo sapiens* is an exception. Human animals are more-or-less good shoppers, effective planners, and careful strategists in office politics--i.e., they are not oblivious to incentives. But does this mean that the upper primates are non-rational as supposed by the *Organismus automaton* hypothesis? If they are also rational, does this mean that the rest of primates are non-rational? If all primates are rational, why stop there? We do not extend the blanket of rationality to other mammals, vertebrates, animals, and eventually all organisms.

Let us suppose that a biologist takes the anti-naturalist philosophical position, viz., restricts rationality only to humans or to the species with which he or she is most familiar with [Khalil, 1996]. If so, the biologist would group all other species under a peculiar category, “non-familiar species,” and assume that they are non-rational, i.e., that the members of this category behave according to genetic programs. After all, the biologist would think: “an ant just cannot be as rational as the organisms I am familiar with” (which would typically mean humans). However, if one adopts such an anti-naturalist stand, it is a daunting task, as the paper shows, to explain the origin of rationality for the organisms in this peculiar category.

Let us suppose that a biologist takes the opposite stand, i.e., the naturalist philosophical position that either *all* organisms are rational or *none* are rational. Let us focus, obviously, on the “none are rational” position. Our biologist would postulate that all organisms act according to genetic programs: signals (what economists call “incentives”) stimulate a gene to “turn on” or

“turn off” and, correspondingly, entice a particular phenotypic behavior. Some of these supposed programmed phenotypic behaviors, our biologist would highlight, are imperfect. For instance, humans would eat non-healthy food and engage in risky behavior. Thus, humans are not rational. However, it must be noted that, contra the “none are rational” position, rationality cannot, as much as natural selection cannot, be refuted by pointing out to imperfections. Every action is costly and agents must make compromises. Some agents, such as students, might rationally choose a non-healthy diet because they want to spend their scarce resources on something else. The idea of tradeoff should not be unfamiliar to biologists: They understand that the design of organisms cannot be perfect, as every design involves a tradeoff between competing functions.

Our “none are rational” biologist would counter that organisms cannot be rational because their decisions often involve *systematic* errors. He or she would argue that organisms act according to patterns of behavior that are too routinized, i.e., cannot change in response to changes in signals (incentives). Many social scientists also use the same observation, viz., humans are the creatures of habit, to support different versions of the hypothesis that humans cannot be rational [Khalil, 2007d]. According to them, humans make also *systematic* errors—such as when they are deceived by a new advertisement gimmick on how to save money. Rationality, though, is not equivalent to error-free decision-making. As shown below, when we discuss bounded rationality and biases uncovered by behavioral economists and psychologists, error-making can be explained on the basis of rationality: Given that search is costly, humans use rules-of-thumb (routines) that have, on some occasions, negative by-products, namely, falling victim to gimmicks and failure to react to change of signals.

This does not mean that humans do not commit irrational acts. Irrational acts include what is commonly called “recklessness” and “addiction” such as self-destructive behavior that may lead to suicide. But it is often the case that human organisms recognize such acts of irrationality and seek to rectify them. Such desire to rectify their acts of irrationality cannot be explained if one adopts the position that the irrational acts are the products of genetic programming. If they are the products of genetic programming, the irrational agents should be as happy as the carriers of the rational lineage. The fact that irrational humans seek help implies that they are, at a deeper level, rational.

There is another challenge that faces our biologist who denies the existence of rationality. Such a position faces an empirical anomaly: Namely, organisms have a variety of responses that cannot be explained by “turned on” and “turned off” genetic programs. There is simply too little time available for genetic instructions when an organism has to flee from a predator or toxic substance or has to pursue prey or seek nutrients. We simply have enormous phenotypic plasticity that cannot be explained by genetic blueprint, as the EvoDevo research program, reviewed below, attests [West-Eberhard, 1989; Raff, 1996; *passim* Müller & Newman, 2003].

However, even when we can explain a phenotypic response to an incentive by tracing it to the “turning on” of a gene or a collection of genes, it is not straightforward how to model this genetic “turning on.” Namely, it does not necessarily mean that the product (the phenotypic behavior), is the outcome of an *exogenous* genotypic change, as is argued in the *Organismus automaton* hypothesis. Rather, the genotypic “turning on” can be modeled as an endogenous response. That is, the decision of the organism would be based mainly on a cost-benefit calculation and the genetic “turning on” simply assists the organism in carrying out the already

decided, optimum action. If the organism decides a different action in response to a change of incentives, it would “turn on” an alternative gene for assistance. The genetic assistance, in this endogenous genetic “turning on” view, merely saves energy and time: It helps the organism to use ready-made repertoires, corresponding to the changing environment, in order to avoid the reinvention of the wheel each time.

Such an endogenous view of genetic “turning on” certainly needs to be explored, but cannot be done so here. It is only mentioned to show that genetic “turning on” does not necessarily entail the *Organismus automaton* hypothesis. It suffices to add that the plausible endogenous view of genetic “turning on” is perfectly *analogous* with a particular, rational view of action vis-à-vis cultural repertoires. Namely, human agents firstly decide on what is the rational action, i.e., undertake a forward-looking calculation of cost and benefit. They then search deep into their past looking for a suitable cultural trait (equivalent to genetic “turning on”) that helps them to avoid the reinvention of the wheel in dealing with the given problem. So, the claim that genetic “turning on” is endogenous is analogous to the claim that culture is not exogenously deposited as claimed by some institutional economists [Khalil, 2007c,f]. Cultural repertoires or motifs are rather revived from the past only insofar as they advance forward-looking interest.

On the other hand, the thesis that cultural repertoires, routines, or genetic “turning on” are exogenous—the *Organismus automaton* hypothesis—begs the question: Why did a genetic “turning on” take place in response to one incentive, but not in response to other incentives? For instance, if a predator spots some prey, the gene for “pursuit” would not “turn on” if the prey is too far. Why is there no response in this case? The neo-Darwinian answer is that the signal (i.e.,

what economists call an incentive) would be too weak to stimulate a response. On the other hand, if the prey is too close, the predator may again not pursue the prey if it senses that the prey is bigger than normal and, hence, cannot be subdued easily. The neo-Darwinian answer, again, is that the “incentive” is not strong enough.

So, ultimately, the predator is making a calculation of costs and benefits. This is precisely what economists call rationality when they study humans. Why should we then call it “as-if” rational decision making when we study nonhuman organisms? Why not use the same framework for all organisms—and hence avoid the arbitrary human/nonhuman divide [Khalil, 1996]?

As shown below [see Sober, 1998], both hypotheses, the *Organismus economicus* and *Organismus automaton*, predicts the *same* optimum allocation of resources. Why would we want, then, to adopt the round-about *Organismus automaton* hypothesis? Why should we violate Ockham’s razor with a redundant theoretical route, i.e., claim that minute rational calculations are only “as-if” rational and attribute them to a genetic program selected by natural selection? In addition, we would have to undertake a somewhat unwarranted theoretical route, to avoid further violations of Ockham’s razor, and postulate that humans, given they are not special, also do not behave rationally. In short, the *Organismus automaton* hypothesis makes the theory unnecessarily round-about. The only advantage of such round-aboutness is rather metaphysical: The theory allows one to not need to account for rationality in nature, which was the main reason behind the triumph of Darwinism [Dennett, 1995; Hodgson, 2002].

On the other hand, some biologists have already moved somewhat towards the *Organismus economicus* hypothesis [Ghiselin, 1974, 1992; Schoener, 1971; McFarland, 1977;

Smith & Winterhalder, 1992; Detrain *et al.*, 1999; Cassill, 2003; Franks *et al.*, 2003; Vermeij, 2004; Raby, 2007; Hurley & Nudds, 2006]. Many of them have moved on their own, without being prompted by economists who have been advancing the bioeconomics agenda [Tullock, 1971, 1994; Landa, 1986; Landa & Wallis, 1988; see Khalil & Marciano, 2007]. These biologists were not responding to the Ockham's razor argument explicitly. Rather, they were merely responding to the burgeoning evidence of phenotypic plasticity, as seen even in plants and brainless organisms, that undermines the old-fashioned *Organismus automaton* hypothesis. The *Organismus automaton* hypothesis basically advances the "all-in-the-genes" view. For this view, there is a sufficiently high one-to-one correspondence between the genotype and the phenotype. Although natural selection theory works as long as there is some non-zero correspondence between genotype and phenotype, the assumption of high one-to-one correspondence is necessary *by definition* for our biologist who insists on tracing every minute difference in phenotype to genotypic instructions, assuming that such instructions are exogenous.

The *Organismus automaton* hypothesis, which supposes a one-to-one correspondence of genotype and phenotype, has been debated extensively under the rubric of reductionism and holism throughout the history of biology [Mayr, 1982; Khalil, 1993]. The consensus is that it is simply impossible to trace every minute variation in the phenotype to heritable genetic material. After all, phenotypic plasticity is too ubiquitous to dismiss [Gould, 2002]. Furthermore, such phenotypic ubiquity is not limited to humans or their favored "smart" animals. The claim that humans, or their favored organisms, are the only one to be free from the shackles of the genes is therefore simply the product of species-centric social science [Khalil, 1996, 1997a].

Phenotypic plasticity amounts to the responsiveness of the organism (in terms of organ

development or behavior) to incentives, which is the focus of development biology. This position can be pushed to underemphasize the role of genes [Levins & Lewontin, 1985]. But one does not have to go that extreme. The basic finding is that the organism is not an automaton in the face of the environment [Lewontin, 1984]. Rather, the organism varies its decisions in light of the change of incentives. So, this paper assumes the *Organismus economicus* hypothesis—viz., it is best to model actions as the result of calculation of costs and benefits. This means that inherited patterns of behavior and bodily capital (organs) can be taken as given constraints, and are thus somewhat similar to environmental constraints, in rational decision making. So, within the inherited and environmental constraints, actions can vary. The same genotype could therefore give rise to a variety of actions. Organisms make decisions in light of incentives, which is the core meaning of rationality as explained below.

2. Again, then, What is Rationality?

If rationality exists, and it exists across all organisms, how then could brainless organisms, from *Amoeba economicus* to plants, ever be rational? But this question is based on the assumption that rationality depends on neural structures such as the brain. This paper challenges this assumption. Rationality has nothing to do with the brain or with mental/neural activities.

The last statement would strike most, if not all, social scientists as surprising, if not audacious. Let us start with the origin of brains. Following Rodolfo Llinás [2001], plants never developed brains because they do not need them. They do not need them because they do not move. Brains, as he argues, are needed by animals that move in order for them to make predictions about the environment as they move quickly. Of course, plants do move when they

climb rocks, seek light, and so on—a subject that fascinated Charles Darwin [1989]. However, such motion is greatly limited and, hence, according to Llinás, would be inefficient for them to develop a brain. So, neural structures are something that organisms, such as *Amoeba economicus* and plants, can live without while still can make decisions.

Still, almost all social and biological scientists would find the above statement surprising, if not shocking. Many scientists explicitly conflate rationality with mental abilities such as intelligence [e.g., Plotkin, 1994, pp. 125-127; Ng, 1996; Herman; 2006; Pepperberg, 2006]. The conflation is not new; it is asserted without argument in key philosophical treatises [e.g., Dewey & Tufts, 1908, p. 306]. One purpose of this paper is to challenge the conflation of rationality with intelligence. Darwin [1998, chs. 3-4] has called intelligence “mental power.” The term “mental power” signifies that intelligence is a trait and, as such, does not differ from “body power” such as muscles. If rationality has zero correlation with body power such as muscles, rationality then has zero correlation with mental power such as intelligence.

As defined by the American Psychological Association, intelligence is an “ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought.”³ Intelligence is a mental faculty not different from memory, color recognition, and musical talent. It is not even different from quasi-mental faculties such as personality temperament. These faculties are traits and, hence, ultimately are similar to sharpness of beaks, length of legs, hearing, eyesight, and so on. As for intelligence, it is an ability concerning the processing of information/time. Such ability is costly. Insofar as rationality is not confused with sharpness of beaks, which is also

³ “Intelligence: Knowns and Unknowns.” Report of a Task Force of the Board of Scientific Affairs, American

costly, it should not be confused with intelligence.

If rationality is not intelligence, what then is it? Rationality is simply the idea that the organism changes its behavior in response to the change of constraints, i.e., incentives. When the organism does so, it tries to maximize the product given the constraints. If the product is calories per unit of effort, the organism chooses an action, such as hunting in area A as opposed to area B, based on expectation or causality pattern. The causality pattern is usually called cognitive “belief” when held by human animals. The causality-pattern specifies what to expect if one action is taken rather than another. So, it is about cause-and-effect relationships in the environment. The causality pattern, in turn, is also the product of rational decision, viz., it must be the best causality pattern given the information. Further, the information is also the product of rational decision, viz., it must be the best information given the expected cost and benefit of search.

So, we have three layers of rationality: First, the agent decides on the best information given the cost and benefit of search. Second, the agent decides on the best causality pattern given the information. Third, the agent decides on the best action given the causality pattern and other resources. The best action amounts to the most effective way to coordinate/allocate activity or resources. Rationality, at all layers, amount to the same thing: the organization of the diverse traits, including intelligence, so that the organism can make the best decision possible in light of the internal and external constraints.⁴ Economists call such behavior optimization. It

Psychological Association, released August 7, 1995: http://www.lrainc.com/swtaboo/taboos/apa_01.html

⁴ The distinction among the three layers of rationality should not entail that there are different kinds of rationality, as Alex Kacelnik [2006] advocates. For Kacelnik, philosophers and psychologists are concerned with one kind of rationality because they focus on the rationality of beliefs, while economists are concerned with another kind of rationality because they focus on the rationality of action. As advocated here, we only have one kind of rationality, viz., getting the best product given the constraints.

was called earlier “rationality optimization” to distinguish it from “selection optimization,” which refers to the diffusion or spread of the most effective trait throughout the population.

Rationality is underpinned by a few assumptions: First, the preferences must be coherent (transitivity axiom) and, second, the agent must be decisive on how to rank them (completeness axiom). These two assumptions, along with minor technical ones [see Kreps, 1990], are merely about consistency of the agent’s goal and hence considered as *necessary* conditions of rationality. The first assumption, coherence, assures us that agents do not make logical errors: If agents prefer X to Y, and prefer Y to Z, then they would prefer X to Z. The second assumption, decisiveness, assures us that agents do not stand by helplessly, i.e., in doubt of what they want.

These assumptions are not sufficient to define rationality as understood here. Rationality is action undertaken to maximize an objective function given the constraints. Agents are rational insofar as they modify action in response to a change in constraints (i.e., incentives). Such behavioral flexibility, exhibited by brainless organisms as well as organisms with brains, is the core of rationality as defined here—irrespective of whether one conceives rationality as bounded or unbounded as discussed below.

Of course, organisms, mostly human organisms, choose in many cases suboptimal behavior [Khalil, 1997b]. To uphold the rationality approach does not entail that decisions are *always* rational as supposed by the revealed preference approach [Samuelson, 1938].⁵

The actions of plants are rational insofar as they react to incentives, i.e., they exhibit behavioral flexibility in response to a change in constraints. Plants react to incentives when they

⁵ One can define “irrationality” as the failure to act according to one’s best interest, which is recognized in the literature on weakness of will. There is an attempt to argue that weakness of will is actually rational. But there is a growing literature, aided with the rise of behavioral economics and behavioral decision theory, that recognizes

change their orientation if re-planted, react to changes of soil moisture, or adjust behavior when exposed to new predators. The repertoires of behavior flexibility can be minute as in the case of the behavior of amoeba and other simple organisms. Nonetheless, amoeba is rational insofar as they do not “walk away” from what they usually consider to be food, or “walk towards” what they usually consider as toxic matter—notwithstanding that some ingestion of toxic matter can be beneficial to make themselves unappealing to predators. If so, plants and amoeba do not need brains to make rational decisions. In simple organisms and in organisms that do not need to move extensively such as plants, there is no single specialized organ to take care of coordination. Instead, in brainless organisms, coordination of functions must be undertaken by less specialized tissues. If this is the case, we should thus not conflate rationality with the brain, as much as we should not conflate mobility with limbs or digestion with stomachs. Rationality, mobility, and digestion can be the prerogative of general-purpose parts of the body such as simple cells communicating chemically with their environment or among a group of cells.

3. Again, then, What is Natural Selection?

As a result of refinements in light of genetics, the modern synthesis of Darwin’s [2006] theory and genetics, known as “neo-Darwinism,” can be stripped to three *basic* principles: 1) blind mutation; 2) blind random assorting of genes during meiosis and chromosomal crossovers; and 3) differential replication. We can lump the first two principles into a single one, viz., the principle of random origin of variety [Khalil, 1993]. According to this principle, traits are the result mainly of the genotype, although environmental factors can also play a role. The traits,

weakness of will as an aberration from rationality. But this debate is better left for another arena [Khalil, 2007b,c].

the phenotype, can undergo change and development as a result of learning and environmental factors. But such phenotypic change, even when it involves somatic genotype change, cannot feedback and manipulate the germline/germ genotype (e.g., the DNA of male semen or a female egg). This one-way causality is known as the Weissman barrier, which is the central dogma of molecular biology. Contrary to Lamarck's thesis on the inheritance of acquired characteristics, the Weissman barrier insists that any acquired characteristic cannot influence the germline.⁶ Thus, the germline is the start of the causal process, and protein cannot manipulate it. Of course, the germline undergoes change—but only as a result of blind mutation. The mutations are neither the product of learning nor the *ex post* benefit of the trait.

To be clear, while germline mutations are blind, they are not totally random, such as in the case where a lioness can give birth to a cub with bird-like wings. After all, the previous scheme or plan of the organism, or what Stephen Gould and Richard Lewontin call [2000] “structure,” imposes limits on the range of viable mutational changes. Within the range, mutation is “blind” in the sense that it does not take place in light of *ex post* payoff. However, if a new lineage does happen to be fitter than others, it would replicate itself more successfully. That is, it would copy itself in a higher number of viable offspring than other less fit traits. Through a number of generational rounds, the fittest trait becomes more frequent than other trait lineages in the population. For neo-Darwinism, evolution amounts to the inadvertent change of the population—neither as the result of any preordained plan, nor as the result of the actors' intention.

⁶ New findings suggest some exceptions to the Weissman barrier. For instance, Edward Steele *et al.*, [1998] have found that the immune system response to new viral invasions, which involves somatic genetic mutations, can be passed on to the germline, i.e., to progeny. If true, this would be one instance of the Lamarckian thesis of the inheritance of acquired characteristics.

There is empirical evidence that may or may not support the seemingly convincing natural selection theory. While empirical and experimental studies are, of course, highly relevant [Orzack & Sober, 1994], this paper is not about whether the data support selection theory. It is rather about the logical coherence of the natural selection explanation when it comes to explain rationality. Also, this paper ignores many issues in evolutionary theory. For instance, it ignores the debate around the unit of evolution: is it the gene, the organism, or a colony of organisms? (The unit of selection, along with the status of the “species” concept, has bedeviled the recent philosophy of biology [Wilson & Sober, 1989; Brandon & Burian, 1984; Sober, 1984].) Whatever is designated as the highest unit of selection, or what is the relation between population and species, the question of optimization will still persist. That is, the question of optimization is orthogonal to the definitions of unit of selection, species, and population.

4. Optimization: Unbounded or Bounded?

For rationality theory, optimization is the maximization of the objective function. Evolutionary biologists call such optimization the maximization of fitness. As such, fitness means which organism, among its competitors for the same niche, is best adapted to that niche, i.e., the fittest. For selection theory, on the other hand, optimization is the maximization of the diffusion of the lineage of the fittest organism in the population. Economists call such optimization equilibrium, i.e., the equalization of rate of profit as capital moves from low- to high-performing units. That is, low-performing units cannot persist, in the long run, along with the rational/fit performers.

So, we have two concepts of optimization [Khalil, 2007e]: As mentioned earlier, the

optimization that the agent undertakes is called “rationality optimization,” while the optimization that nature (market) undertakes is called “selection optimization.” As shown in Figure 2, both of these concepts of optimization, amazingly, can be expressed in two different flavors: as bounded

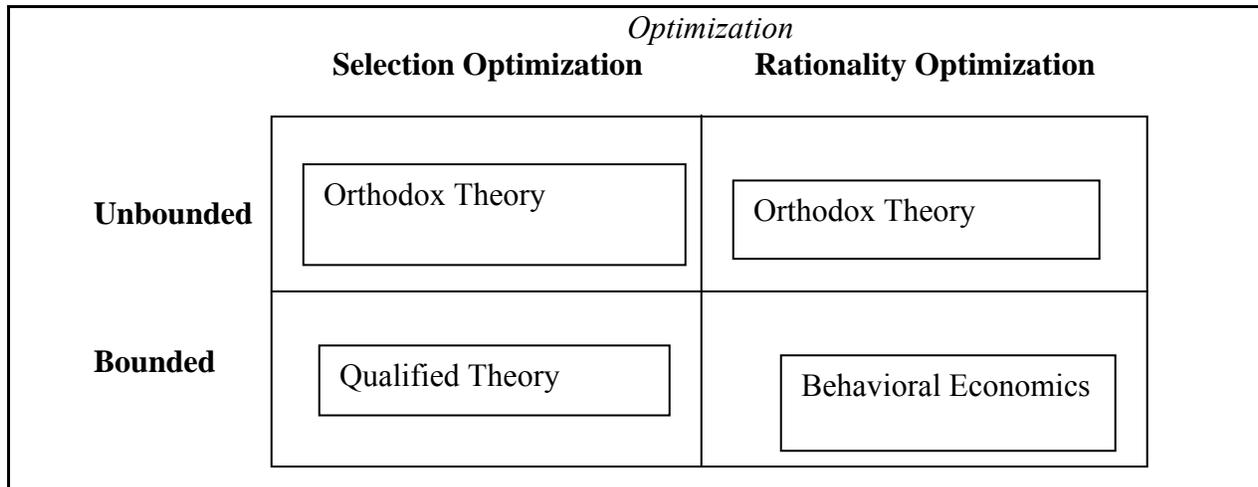


Figure 2: Optimization: Unbounded and Bounded

optimization and unbounded optimization. There are thus four possible combinations. In **unbounded selection optimization** [e.g., Dawkins, 1976, 1982], a population becomes fully adapted to its environment when the fittest lineage, in comparison to other lineages in the population, becomes the dominant one. The optimization is “unbounded” in the sense that there are no enduring obstructions, transaction costs, or information imperfections that prevent the population from reaching full adaptation, i.e., becoming fully dominated by the fittest lineage. Likewise, with the absence of obstacles, market competition ensures the dominance of the best technology so that there is uniform rate of profit across the economy. In **unbounded rationality optimization**, on the other hand, agents make the best decisions given the information available and the cost of further search. The optimization is “unbounded” in the sense that there are no obstacles that prevent the agent from calculating the optimum allocation.

Both positions are disputed by the bounded view. Concerning **bounded selection optimization**, “soft” Darwinians such as Stephen Jay Gould [1977] argue that the outcome of natural selection is often haphazard and engenders a population with ill-fitted organisms. This is

the case because of stickiness or friction in the operation of natural selection, which saps the selection forces. So, we cannot assume that the fittest allele will win the day and become frequent enough in the population. The fittest allele can become more frequent, but short of being optimally frequent. That is, it might have to share the total pool with less fit alleles.

In fact, a population can never reach a unanimity in terms of the dominance of one allele. Some variety is actually recommended, as advocates of the unbounded view quickly point out. Such advocates argue that diversity acts as a hedge “fund,” which, although maybe being an “ill-investment” given the current environment, is in fact optimal given the vagaries of environmental conditions. It is always good for a population to preserve such a hedge fund (diversity) so that, in the case of drop of temperature or invasion of a disease, some individuals will not be wiped out. Such a statement is framed in a structural-functionalist (holistic) thesis, but it can easily be operationalized into a non-holist or individualist thesis: The environment is actually mired with uncertainty, where each “state of the world” has a different probability of coming into being. So, it is selection-optimal in the unbounded sense to sustain a diversity of alleles. In this manner, each lineage is “fitted” to its temporally-spread ecosystem state of the world or, in other words, to its geographically-spread eco-niche. So, the diversity of alleles is actually an equalizing equilibrium, i.e., expressing efficient or unbounded selection optimization [Roughgarden, 2004].⁷ The notion of “equalizing equilibrium” is identical to the notion of “equalizing differences” in economics [Rosen, 1981]. Examples of equalizing differences include the diverse prices of homes indicating their distance from the center of town, the diverse

⁷ Roughgarden celebrates the diversity of population, which she calls the “rainbow.” But she seems to conflate the diversity of a trait, such as the co-existence of beaks with differing sharpness, with the number of a gender in a population. The latter is not about the diversity of same trait, but rather the complexity of specialization of the process of reproduction.

wages of employees indicating their differences in skill, and the diverse rents of land indicating location or fertility.

However, it is also possible under bounded selection to interpret the persistence of diversity as something suboptimal, i.e., indicative of “flabby” (bounded) selection optimization forces.⁸ Likewise, sagging market selection optimization (competition) can allow firms within the same industry to coexist for a long time with various kinds of technologies (and hence productivities), as Daniel Luria [1996] documents. (On the other hand, as argued above, such a diversity of technologies can be conceived, instead, as optimal in the face of fluctuating market conditions.) But this should not concern us here. The point is that bounded market selection optimization is possible, at least theoretically. Such boundedness could feasibly arise as a result of increasing returns or high transaction costs (friction). It could also arise as a result of path dependency, in which accidental events, such as which technology is first introduced, matter because first-comers can carve a path of change that makes them withstand competition from newcomers that turn out to have fitter technologies [Dosi *et al.* 1988, Arthur, 1989].

Considering **bounded rationality optimization**, neoclassical economists hijacked Herbert Simon’s [1957] notion of “bounded rationality” or “satisficing” to argue the following: The mental capacity of the agent is bounded by its intelligence, memory, and other cognitive abilities. Such a bounded mental capacity entails particular transaction costs to search for information, to digest the optimum information to come up with optimum beliefs, and to use the optimum beliefs effectively to come up with the optimum action. It is very costly for the agent to come up with the optimum decision. So, agents adopt institutions, heuristics, or rules that

⁸ Such a possible “inefficient” selection optimization (i.e., sub-optimal equilibrium) differs from the notion of

work most of the time, but occasionally cause some errors. Such heuristics, on average, allow agents to do better, given their limited capacity, than if they were to optimize unboundedly at each instant.

If agents are going to optimize in each instant as supposed by unbounded rationality, they might avoid some pitfalls, but it would be enormously expensive. So, they adopt heuristics that, as by-product, make behavior deviate from the predictions of unbounded rationality. These deviations have been uncovered by behavioral economics and behavioral decision research [Camerer *et al.*, 2004]. Even research in nonhuman animal behavior has shown that such agents also exhibit bounded rationality [e.g., Schuck-Paim *et al.*, 2004]. The point is rather how to interpret such deviations. As mentioned already, and following the argument of Gerd Gigerenzer [2005, 2006], such shortcomings or errors in logical reasoning are occasional failures of rules of thumb or “heuristics” that are, on average, effective. It pays on average to have heuristics, and the price of such shortcuts is that one occasionally makes mistakes. If so, we should not get rid of the “quirks” of cognition if it also means sacrificing the heuristics that are effective most of the time. Even if we “improve” computations and minimize errors in logical reasoning, we are not improving rationality. We are rather employing a different technology that allows us to make different decisions. Given that such technology is costly, one adopts it only when the payoff is greater than the cost.

Put differently, if we take a broader account of all relevant information, given the expected costs and benefits of such information, to achieve the best solution, sometimes it is better to settle for the second-best. For instance, if a person calculates each time whether to take

“inefficient” rationality optimization (i.e., irrationality) along the same line that differentiates these two kinds of

the stairs or the elevator, he or she would, on average, lose more resources than if he or she would settle on the second-best and adopt the habit of taking one means.⁹ So, even when the agent acts according to second best, i.e., acts with bounded rationality, he or she is *using* unbounded rationality.

So, debates, often voiced, of whether a behavior expresses bounded or unbounded rationality is ultimately an empirical question. The disagreement often misses the implied agreement, viz., bounded optimization does not actually undermine the notion of optimization *per se*. When one advances the thesis that rationality optimization is bounded, one has to assume unbounded rationality optimization in order to account for bounded rationality optimization.

These issues are replicated with debates of whether evolution can generate false beliefs. For instance, Christopher Stephens [2001] calls such inaccurate beliefs “better-safe-than-sorry arguments.” Such beliefs prompt agents to adopt risk-averse behavior when it appears that accurate beliefs would lead agents, at least in some occasions, to amass greater wealth or progeny. However, accurate beliefs that prompt agents to be risk neutral may actually, on average, decrease wealth or progeny if the greater accuracy entails costs unjustified by marginal benefits. Here, again, when theorists want to show the limits of optimization, they have to use optimization. Likewise, Stephen Stich [1985, 1990] argues that suboptimal beliefs can be adaptive: In evolution, what matters is not finding the truth, but finding beliefs that afford greater chances of survival and, hence, leaving greater offspring on average than other beliefs. If agents

optimization, as discussed below.

⁹ In economics, some authors have also pointed out situations of market stickiness that resemble the stickiness that hinders evolutionary selection optimization as pointed out by Gould. As a result of nonergodic feedbacks concerning innovation, market niches and other variables, the potentially more productive firms do not even appear [Khalil, 2000].

act very cautiously and develop inaccurate beliefs that engender risk-aversion that maximizes chances of survival, they can be favored by natural selection over agents who attend to the facts and develop more accurate beliefs that put them at greater risk. Again, such inaccurate beliefs do not indicate sub-optimality. They rather indicate the bounded nature of optimality.

5. Challenging Rationality Optimization: Simon's Procedural Decision Making

Herbert Simon spent the latter part of his career insisting that his notion of bounded rationality had been misinterpreted (hijacked) by the neoclassical approach [Callebaut, 1998; Mousavi & Garrison, 2003]. To salvage his notion, Simon [1977] devised an alternative term, "procedural rationality," to stress that he is actually challenging both *unbounded* and *bounded* rationality optimization. To wit, the procedural rationality notion supersedes the bounded/unbounded distinction. For Simon, agents are simply so enmeshed within their *particular* environment or *particular* experience that their decision making is, rather, experiential or procedural. The procedural rationality notion ultimately means that agents would continue a particular routine, procedure, or habit of production as long as there is no sudden shock that upsets the *status quo* [Khalil, 2007d,g]. Only with a shock, the agent would search outside the box for an alternative routine or procedure. Thus, agents do not stand outside their environment, and be involved in the kind of rational calculation about rules and heuristics, as assumed in neoclassical economics.

This procedural view of decision making coheres with the modularity approach generally [e.g., Callebaut & Rasskin-Gutman, 2005], and with a particular strand of evolutionary psychology [Cosmides & Tooby, 1994; Sperber, 1994; Samuels 1998, 2000]. For instance, Leda

Cosmides and John Tooby argue that rationality is *not* a “content-independent” faculty. That is, rationality cannot be defined independently of the particular experience and, hence, it consists of whatever habit of behavior that the agent finds feasible, which is Simon’s notion of satisficing. Paul Glimcher *et al.* [2005], likewise, argue that we should not start with axiomatic rationality, but rather with evolutionary history that takes particular brain circuits, physiological structure, and the relevant environment seriously. The problem with Cosmides, Glimcher and other advocates of the evolutionary approach is that they confuse rationality with what rationality uses, viz., eyesight, brain circuits, and other abilities. It is similar to the earlier discussed confusion of rationality with the intelligence ability.

This recent upsurge of evolutionary explanation, which stresses historical experience, is not new. It was advanced in 1922 by John Dewey [1957] and even earlier by the founders of American Pragmatism such as William James [1893; see Khalil, 2003, 2007d; Mousavi & Garrison, 2003]. The pragmatic approach greatly influenced Thorstein Veblen’s [1998] critique of the rationality approach of economics. The pragmatic/evolutionary approach emphasizes that what appears as rational action is actually the result of experience that has engendered useful routines on how to interact with the environment. There is no such thing as global rationality; there is only habituation where the organism and the environment are united in, what John Dewey and Arthur Bentley [1999] call, “transaction” or “union” of the known (environment) and knowing (the actively involved organism). So, agents behave in particular ways because such ways are how they behaved in the past and, hence, cannot express “content-independent” rationality.

The idea that there is no content-independent rationality undermines the notion of

rationality. As such, it states that natural selection can explain rationality because rationality does not exist in the first place. So, the modular, evolutionary psychology, or pragmatic view solves the problem of the origin of phenotypic plasticity by denying that agents are rational. However, even if we go along with the procedural view, agents still would have to make a decision in light of an environmental shock; agents still have to adopt the best routine possible in such situations. Thus, we cannot entertain the Simon/Dewey/modularity line of argument.¹⁰

6. The Rationality-qua-Trait Thesis

If we accept the *Organismus economicus* hypothesis, can natural selection account for rationality? The usual answer is that if we consider rationality as a trait, the “Rationality-qua-Trait Thesis,” natural selection theory can explain it as it does explain other traits.

The Rationality-qua-Trait Thesis is common to economists. For instance, as Alchian [1950] and Friedman [1953] have postulated, if firms are irrational, they would fold very quickly. So, what is left to survive, if we assume *unbounded* selection optimization, are the rational agents. And given that genetic mutations cannot improve on 100 per cent rationality, i.e., simply being rational, what is left to survive must be the lineage of the *most* rational agents. However, even if we assume *bounded* selection optimization, the most rational agents will still be the dominant lineage.

Even philosophers, such as W. Quine [1969] and Jerry Fodor [1981, p. 121], point out that if humans are irrational, they would have long become extinct. And the same applies,

¹⁰ Actually, the Simon/Dewey argument leads to the following problem: If action is simply a habit, it cannot draw a distinction between technological learning and scientific learning [Khalil, 2007d]. In technological learning, the agent learns how to manipulate the environment in order to attain a goal. In scientific learning, there is no goal or benefit at hand: The agent only wants to know how the environment works. Dewey argues that the two types of learning are “habit.” If so, agents adopt a technology just because of an accident, not because of figuring out a better way to attain the goal. As such, the Simon/Dewey line cannot afford an endogenous theory of evolution of technology.

apparently, to nonhuman animals. Some thinkers question whether *all* nonhuman animals are rational [Witt, 2003; see Khalil, 1996]. However, even if we restrict rationality to some organisms, the neo-Darwinian explanation is still that it is a trait favored by selection optimization.

To wit, as Elliott Sober [1996, 1998; in Orzack & Sober, 2001] argues, rationality optimization should not be feared. Rationality optimization generates, at first approximation, the same allocation of resources as selection optimization. Sober, to be sure, recognizes that the structure of rationality optimization differs from the structure of selection optimization. Nonetheless, for Sober, both kinds of optimization engender the same outcome.

Sober proposes a thought experiment that he calls the “heuristic of personification.”¹¹ If the organism is given the chance to choose a trait present in the population, it would choose the best trait, i.e., the most efficient given the constraints. For instance, if a zebra were given a choice, holding everything else constant, it would choose the trait of speed over slowness under some conditions, and *vice versa* under other conditions. For Sober, given that both kinds of optimization engender the same outcome, there is no substantial difference between rationality optimization and selection optimization.

To illustrate this further, let us start with a hypothetical population that has a large frequency of agents with an idiosyncratic taste for, e.g., prolonged dance ritual that reduces their fitness. Such agents would be misallocating resources in comparison to the agents who are free from such an idiosyncratic taste. Assuming an unbounded optimization mechanism, the agents who survive, and leave the greatest offspring, are the ones with the non-idiosyncratic taste lineage (i.e., the ones who have the fitness-enhancing taste). Such agents would rationally choose an allocation of resources that does not differ from what natural selection “chooses.”

But still, even if rationality optimization must coincide with selection optimization, can

¹¹ The following paragraph is similar to one in Khalil [2007e].

selection optimization *explain* rationality optimization?

7. The Incoherence Problem

As laid out in Figure 1, the task is to show why the Rationality-qua-Trait Thesis faces a problem of incoherence. The problem of incoherence appears at three layers. The first layer discussed below concerns the assumption that traits undergo progress, i.e., one can rank traits *without reference* to environmental conditions or incentives (“constraints”).

The idea of progress is antithetical to both selection optimization and rationality optimization. If one subscribes to selection optimization, the second layer discussed below, one cannot identify progress or what is the best trait in the abstract, without reference to the constraints. Selection optimization simultaneously determines, in light of the constraints, what is the best type and, consequently, what must proliferate throughout the population if there were no exogenous obstacles. This is in contrast to the ignoring of the constraints in the idea of progress.

Likewise, if one subscribes to rationality optimization, the third layer discussed below, one cannot identify progress or what is the best technology in the abstract. That is, one cannot rank different production possibility frontiers, based on technological progress, without reference to the constraints. Rationality optimization simultaneously determines, in light of the constraints, what is the best type and, consequently, what is the efficient use of the resources if there were no imperfect information and other exogenous obstacles. This is again in contrast to the idea of progress.

7.1 The Progress View

Let us adopt a “folk biology” definition of evolution. Namely, let us assume that a trait such as 80 per cent rationality is superior to 60 per cent rationality in the abstract sense, i.e., without reference to constraints. Given that we are ignoring here optimization subject to constraints, the term “evolution” therefore means progress. The idea of progress does not fit with the neo-

Darwinian, selection optimization view, as will be shown next. However, putting aside the neo-Darwinian objection, the idea of progress entails that we can measure the improvement of the trait as if it is costless, i.e., we can ignore environmental conditions or incentives or, in short, constraints. So, we can compare the productivity of one trait vs. another as if these traits are technology fallen from heaven and, hence, assume that productivity differs from efficiency.¹²

The criterion of progress may vary, depending on the trait under question. For instance, one can postulate that eyesight progresses according to the improvement of sharpness of image. For a vehicle, the criterion of progress is the speed of the vehicle. In either example, the trait undergoes betterment or improvement without reference to the cost of such betterment or improvement. Such improvement or progress takes place through mutation or invention. The invention/mutation allows the agent to increase its productivity in amassing greater surplus or offspring. The source of the invention/mutation can be a random shock or intention [Khalil, 2007a]. Whatever the source, it is irrelevant to the main issue. Namely, according to the progress view, we can rank the increase of productivity without regard to constraints.

From the progress view, if the Rationality-qua-Trait Thesis is granted, rationality must be the subject of improvement as the case of all other traits. But can this be the case? There is usually an upper limit to the improvement of the speed of, e.g., a particular species of ants given usually by its basic structure. Such structure includes the basic proportion of limbs to body shape. One can also think of rationality improvement, as the rationality of agents inches closer towards the maximum 100 per cent rationality.

However, there is a difference. If the basic structure of the ants under focus changes, such as acquiring longer limbs, the upper limit of speed changes likewise. One cannot state the same with regard to rationality. Rationality is not structure-dependent. That is, rationality cannot attain a new upper limit as a result of change of the structure of the species. Rationality is

¹² The assumption that we can separate productivity from efficiency underpins my earlier paper [Khalil, 2000]. The paper discusses the selection of productivity as separate from rationality.

rather an abstract, structure-free mode of behavior defined, as stated earlier, in terms of reaction to incentives. Such a mode of behavior is not defined in reference to any structure. Once behavior is depicted as 100 per cent rational, it is not possible to improve it any further by changing the structure of the organism as in the case of speed, intelligence, or any other trait. That is, there is an upper limit to the improvement of rationality—defined by optimization rationality—where such a limit is structure-free. In contrast, the upper limit of the improvement of traits such as speed, intelligence, or eyesight is structure-dependent. So, rationality cannot be a trait in the same sense as speed, eyesight, and sharpness of beaks. Simply put, the progress framework cannot support the Rationality-qua-Trait Thesis.

7.2 The Selection Optimization View

However, the Rationality-qua-Trait Thesis is incoherent on another level, namely, the neo-Darwinian view that rejects the idea of progress.

For neo-Darwinism, the idea of progress is meaningless because it assumes that traits and improvements are costless. Even a trait such as intelligence comes at a price. For an organism, it might be better to be less intelligent but faster in running, given that intelligence consumes a great amount of resources. And even if we face a puzzle, such as the continuous improvement of intelligence in vertebrate evolution, it can be explained in terms of the parallel increase of diversity of the ecosystem, as Yew-Kwang Ng does [1996]. With the increase of diversity, the marginal value of intelligence (which is, as discussed previously, not the same as rationality) increases.¹³

So, for neo-Darwinism, whether a trait is better than another depends on the environment. If a trait, as simple as speed, undergoes improvement, it must be at the expense of another trait

¹³ But then one has to explain how the ecosystem tends to move to greater diversity over time. It is tautological to state that it moves to greater diversity because of an arms race: The rise of intelligence in the ecosystem increases the payoff of an organism to have intelligence. This *assumes* that the initial rise of intelligence was the fittest trait, which is exactly the question that one wants to answer.

such as eyesight. So, one cannot measure improvement in the abstract sense, i.e., without regard to the tradeoff.

Let us suppose a genetic mutation affords a better mix or a better lineage of speed and eyesight. Such betterment can only be judged with reference to the particular environmental constraints. Under different constraints, the mutation can actually be a regression. Given that the mutated lineage is betterment, the greater fitness of such a lineage can become dominant in the population only if there are no obstructions to selection optimization. If there are no obstructions, fitness amounts to the adaptation of a population to its particular environment in light of the fitness of the particular lineage. When fitness of the population increases, it means that the fittest lineage, as measured in relation to environmental conditions, has proliferated throughout the population. The judgment of what is the appropriate trait mix and the judgment of what is the dominant trait mix in the population are thus both the same judgment: Both are determined by comparing the varied traits in the population according to their relative effectiveness in a particular set of constraints [Dawkins, 1976; Mayr, 1982]. For instance, if a long beak at the expense of another trait is the fittest given other available alleles and environmental conditions, a short beak can become the fittest under different conditions. It is impossible to rank the long beak and the short beak according to some criterion of progress without reference to the constraints.

Thus, within the temple of neo-Darwinism, we cannot state that 100 per cent rationality is better than 60 per cent rationality (i.e., irrationality) without reference to the constraints. Both are simply different alleles of rationality and, hence, cannot be ranked without reference to particular constraints. So, it is possible for irrationality, under some conditions, to be better than 100 per cent rationality. We can thus, at best, call the different shapes of rationality by different names—similar to how we call the different shapes of beaks. All shapes of rationality are equivalent in the sense that they cannot be ranked innately. They can only be ranked in reference to particular constraints. As such, there is no rationality. There are only agents

programmed by different shapes of rationality or modules.

As discussed earlier, such a position cannot be sustained in light of the fact that agents do respond to incentives, i.e., behavior is elastic with regard to incentives: Action is not simply an expression of a module oblivious to changing environment. So, the reality that the same agent responds to different incentives differently leads to the idea that rationality exists. As such, the neo-Darwinists view it as a trait subject to the usual selection pressures. Is such a view feasible?

Let us say that there is one zebra that carries the gene for inconsistency and indeciveness to a great extent that makes it a 60 per cent rationality carrier. This Zebra is obviously less fit than 80 per cent rationality carriers. But let us say that new constraints arise that allow the 60 per cent rationality carrier to be fitter than the 80 per cent rationality carriers. Given that the 60 per cent rationality carrier is the fittest, it must be maximizing its objective function better than the 80 per cent variant. This leads to a contradiction: How could the 60 per cent variant more rational than the 80 per cent variant?

A neo-Darwinist theorist would “solve” the contradiction in the following way: It is wrong in the first place to differentiate rationality according to degrees, as if they can be lined up according to some ideal criterion of progress. This is because neo-Darwinists must reject the notion of an “ideal” trait, given that they disagree with the idea that evolution progresses towards such an ideal trait without reference to constraints.

So, we should see whether the point is robust, i.e., without any allusion to which lineage is more rational. Let us construct two thought experiments, one with regard to a real trait, such as the length of beaks, and the other with regard to rationality-qua-trait:

Thought Experiment I: Let us assume that a short beak (SB) is fitter than a long beak (LB) under constraints C_s , i.e., constraints favorable for the short beak lineage. The opposite is true under constraints C_l , i.e., constraints favorable for the long beak lineage. We assume here that SB and LB are non-identical lineages. We also assume that they have non-trivial fitness values

(i.e., have a different impact on outcome) which differ from each other. Given C_s , SB agents would have more offspring or surplus than LB agents. Given C_l , LB agents would have more offspring or surplus than SB agents.

In Thought Experiment I, one lineage can be fitter than the other because, under particular constraints, it affords greater production of surplus than the other lineage. Such fitness calculation can be undertaken by applying the *same* rationality and, hence, the difference in performance reflects only the difference in the traits under given constraints. A lineage is thus deemed the fittest because it allows the agent, under particular constraints, to produce more surplus than competing lineages. This superior fitness is spread, via selection optimization, throughout the population. This is an additional optimization mechanism that does not concern us here.

In Thought Experiment II we introduce rationality as supposedly the subject of selection. Let us assume two lineages of rationality and call them “short rationality” and “long rationality.” The terms are chosen arbitrary to rule out *absolute* rationality ranking à la progress view of evolution:

Thought Experiment II: Let us assume that short rationality (SR) is fitter than long rationality (LR) under constraints C_s , i.e., constraints favorable for the short rationality lineage. The opposite is true under constraints C_l , i.e., constraints favorable for the long rationality lineage. We assume here that SR and LR are non-identical lineages. We also assume that they have non-trivial fitness values (i.e., have a different impact on outcome) which differ from each other. Given C_s , SR agents would have more offspring or surplus than LR agents. Given C_l , LR agents would have more offspring or surplus than SR agents.

Thought Experiment II explicitly acknowledges rationality, which results in behavioral flexibility. So, agents here must be using rationality in the determination of their choice or

performance. This is not *necessarily* the case in Thought Experiment I where the “choice” could have been determined, amazingly, through some automaton mechanism à la *Organismus automaton* hypothesis.

In their determination of choice in Thought Experiment II, let us assume that LR agents are using “LR*” rationality, while SR agents are using “SR*” rationality. Under C_1 constraints, LR agents would choose an action, let us call it LA_1 , which differs from the action they would choose under C_s constraints, let us call it LA_s . Likewise, under C_1 constraints, SR agents would choose an action, let us call it SA_1 , which differs from the action they would choose under C_s constraints, let us call it SA_s . As mentioned earlier, such behavioral plasticity is the core of rationality. So,

$$V^{LR*}(LA_1) > V^{SR*}(SA_1) \dots \text{(i.e., under } C_1 \text{ constraints)} \quad (1)$$

$$V^{SR*}(SA_s) > V^{LR*}(LA_s) \dots \text{(i.e., under } C_s \text{ constraints)} \quad (2)$$

whereas V is the fitness or value function. Given a particular regime of constraints, the value depends *apparently* on two variables: the particular rationality method used (the superscript) and the particular trait employed in the action (LA or SA).

The *key* question is whether we actually have two variables. In particular, do we have different rationality methods or is $SR^* \neq LR^*$ possible? The two methods cannot be other than identical. The optimum behaviour might be different for different functions—but the *measure* of such of differences must be the same. For instance, two different mountains could have different or identical summits. But the way we *measure* a summit—viz., in relation to sea level—is identical for both mountains.

Let us assume the contrary, i.e., $SR^* \neq LR^*$. Then, we cannot at all compare the two value functions. We cannot determine which is greater. It would be similar to comparing two distances in different units, when it is impossible to translate one unit of measure to the other. The fact that we can compare two value functions, $SR^* \equiv LR^*$. Then, we will just call it “R*.”

If so, there is a problem with Thought Experiment II. It uses rationality in two

incompatible functions. First, it uses rationality as a faculty not different from how it uses intelligence, muscle power, or beak length. In this function, we could have different lineages of rationality. Secondly, it also uses rationality as R^* , i.e., as a uniform measure or way to calculate the optimum. So, a lineage of rationality-qua-trait is deemed fitter than all other lineages through the usage of R^* . This double-function of rationality does not arise with regard to the trait under focus in Thought Experiment I. The beak-qua-trait was only used as a trait; it was *not* used to determine the optimum, i.e., fitness value.

The double-function of rationality in Thought Experiment II leads to incoherence. The calculation of fitness of rationality-qua-trait under C_s regime or C_l regime involves the use of the *same* kind of rationality, R^* . So, we have a contradiction at hand: *How could $R^*=SR$ and $R^*=LR$, while $SR \neq LR$?*

There are apparently six, as far as I am aware, solutions to the contradiction. However, all six solutions turn out to be unsuccessful:

- i. Agents do not follow a uniform rationality mechanism (R^*) when they are carriers of the diverse rationality-qua-trait. So, the SR agents must be using the SR mechanism while the LR agents must be using the LR mechanism. Let us suppose that the SR and LR agents, i.e., agents with different abilities, need to calculate the distance between points A and B. They would come up with different estimates in correspondence with their different lineages (abilities). This should be obvious if the trait under question is eyesight. The long-sighted agent, for example, might be better than the short-sighted agent in estimating distances, but *vice versa* in discerning color. But when the two agents use the trait, viz., either eyesight or rationality-qua-trait, they must use it in *the best possible way*. It does not make sense to state that the short-sighted agent uses the short-sighted mechanism by examining the object for a day, while the long-sighted agent uses the long-sighted mechanism by examining the object for a week. If the agents take different periods of time to process the

information, then these periods of time are lineages of another trait, let us say intelligence. If so, we should have included the differences in intelligence at the start of analysis. Now, given all the differences in eyesight and intelligence, the agents must have used their given traits in the *best possible way* in estimating the distance between A and B. That is, they *would not undertake irrational acts* that would hinder their estimation. The “best possible way” or “would not undertake irrational acts” means that both agents used the same rationality mechanism, R^* .

- ii. Agents do not use rationality at all. While this *Organismus automaton* solution solves the contradiction, it takes us back to Thought Experiment I. We would be back facing the empirical anomaly *Organismus economicus*.
- iii. Agents use “satisficing” rationality that, for Herbert Simon, differs from optimizing rationality. So, unlike the first solution, both agents might use R^* , and, unlike the second solution, they do use some kind of rationality. But such a kind of rationality is not about finding the best product. This solution is proposed by evolutionary psychology, modularity programming, and pragmatism [Khalil, 2007d]. For these approaches, organisms start with useful routines, satisficing habits, or modules adapted to particular set of stimuli [Khalil, 2007d]. Ultimately, though, this routine solution simply reverts back to Thought Experiment I: Natural selection can explain rationality because there is no rationality to start with. This is actually the logical conclusion of Ng’s [1996] model: To explain rationality via natural selection, Ng has to conflate it with other traits such as mental ability, and hence there is no rationality, aside from intelligence, that needs to be explained.
- iv. Agents, following the idea bounded rationality in neoclassical economics, follow rules-of thumb or heuristics. Namely, agents adopt SR and LR institutions to *economize* on transaction costs (from search to use of limited mental traits such as memory and intelligence). But notice the word “economize.” It entails that, at a

- deeper level, there is a global rationality, R^* . Thus, we are back to the contradiction.
- v. The incoherence problem is actually the result of taking the rationality trait in the abstract sense. For instance, if we take “vision” in the generic sense, do we have the same problem as what we are faced with regarding rationality? If so, the issue is not really rationality but rather any trait taken in the generic sense. This means that natural selection cannot explain vision *per se*. It can only explain a particular sharpness of vision or a particular angle of vision. Likewise, we can say natural selection can explain a particular sharpness of rationality or a particular angle of rationality. So, rationality has no general meaning. That is, it does not exist as such. But if this is the case, we have solved the incoherence problem by denying that rationality exists, and hence we are back facing the empirical anomaly.
 - vi. Rationality is costless and, hence, it is not subject to substitution. Natural selection would always favour higher grade of rationality. While it might be the case that rationality is costless, it cannot be the case that natural selection always favors rationality over irrationality. For natural selection to favour something, it must be a trait and, hence, must be costly. So, it would be inconsistent for a neo-Darwinian theory to claim that the most rational lineage would be favored over all other lineages by natural selection simply because it is innately better. This ignores the issue of constraints, as mentioned above.

The last apparent solution actually characterizes the arguments of Alchian [1950], Friedman [1953], and (neo-Darwinian) evolutionary economists in general: They ignore that rationality, once seen as a trait subject to selection forces, cannot be assumed to be costless or, what is the same thing, innately superior to irrationality within the neo-Darwinian framework. What matters in market competition (selection optimization) is that the market will allow the survival of firms that have the highest rate of profit (highest fitness). If rationality is a trait, and hence costly, it cannot be assumed that profitability is a positive function of rationality. That is,

it cannot be assumed that the highest profitable firm is also the highest firm in the scale of rationality. It is possible that, given that rationality is a trait and hence costly, a less rational firm would be fitter than a more rational firm. Market competition (selection optimization) would favor the best rationality lineage *only under particular constraints*. Under other constraints, the less rational agent—analogue to low-level technology—can have greater net advantages over the more rational agent—analogue to high-level technology.¹⁴ A consistent neo-Darwinist cannot thus assure us that natural selection favors rationality.

In conclusion, the logical incoherence of the Rationality-qua-Trait Thesis cannot be avoided as long as one accepts that rationality exists. Therefore, neo-Darwinists are caught between a rock and a hard place. When they deny the *Organismus economicus* hypothesis, they simply explain fitness as the outcome of genetic programs where, given the constraints, some programs are more successful than others. But then they face the rock: the empirical anomaly of rationality in the sense that agents change their behavior in response to signals that cannot be explained by “turned on” and “turned off” genes. To avoid the rock, they accept the *Organismus economicus*. But then they face the hard place: the problem of incoherence in the sense that rational agents cannot also be carriers of different lineages of rationality.

7.3 The Rationality Optimization View

Let us approach the incoherence problem from the view point of rationality optimization. Let us assume that agents can make intertemporal allocation that involves investment which means the improvement of one’s traits. For instance, humans can invest in their education and acquire skills to become carpenters or neurosurgeons. If rationality is a trait, such agents should also be capable of choosing the degree of rationality suitable to them.

¹⁴ My earlier essay [Khalil, 2000], with its distinction between productivity and efficiency, suffers from the same mistake: It assumes that a highly productive technology is inherently superior to low productive technology, in total abstraction from the constraints.

To state that agents can choose a trait entails two assumptions. First, the chosen trait must be substitutable or fungible with other traits. If agents can only be carpenters, they have no choice. The fact that agents can make choices, the resources used to attain one product (such as tables) can be used to attain another product (such as logged timber). Second, the trait must be scarce. If it is free, such as air, there is no need to make a choice in the first place.

The two assumptions also underpin the Rationality-qua-Trait Thesis. So, rationality must be costly and fungible. Concerning fungibility, the agent decides between investment in resources to nurture rationality or investment in the same resources to nurture, e.g., the skill of carpentry. Let us assume that the agent chooses greater skill in carpentry at the expense of rationality. Let us assume that after investing in carpentry, there was a totally unexpected shock. The expected return from rationality suddenly rises, while the expected return from carpentry remains the same. If the difference of returns is high enough, and the transaction cost of switching is low enough, the agent should forget the sunk cost of the earlier decision and start anew investment in more rationality at the expense of carpentry. But how could the agent make a rational decision of switching if he or she is already irrational?

More broadly, can an agent, starting with any degree of rationality (X), trade it with another trait (Y)? For such substitution to take place, he or she must use the very skill (X). This leads to self-contradiction: If rationality is a trait, one cannot use rationality-qua-tool to make decisions on whether to invest more on rationality-qua-trait at the expense of an alternative trait. This finding stands irrespective of whether one proposes that rationality is bounded or is unbounded. That is, the finding is orthogonal to the issue of adoption of short-cuts or heuristics that may lead, in particular instances, to sub-optimal (“irrational”) choices. As discussed earlier, agents adopt heuristics because of bounded rationality, i.e., to avoid expensive search or deliberation costs. The adoption of heuristics is not a decision about trading rationality *per se*, but rather is about whether to use calculation in each instance or to adopt a rule/habit, given that calculation requires costly search and time-consuming calculation. Some thinkers [e.g., Stich,

1985; Stephens, 2001] confuse rationality *per se* issue and the issue of adoption of beliefs (heuristics). For agents to decide on the adoption of heuristics, they use rationality *per se*. And when they adopt general rules, which in some cases or in hindsight turn out to be fallacious, they are not giving up rationality *per se*. In fact, they retain rationality *per se* in order to adjust, give up, or adopt more heuristics.

8. Rationality-qua-Method Thesis

Given the incoherence problem, we face two options. First, we deny that rationality exists: Agents simply act according to programs as supposed by the *Organismus automaton* hypothesis. That is, organisms do not change decisions in response to changing constraints. Given that rationality exists, however, the only option available is to reject the Rationality-qua-Trait Thesis.

Rationality is not an element that occupies the set of constraints or the set of objectives. It is rather a *method* of how to relate these two sets in order to produce the highest benefit. The proposed Rationality-qua-Method Thesis entails that rationality is decision making that advances the interest of the organism given the internal constraints (traits) and external constraints (environmental resources). Otherwise, if rationality is an element (trait) that occupies one set only, it leads to a contradiction: How could then rationality allocate rationality efficiently? In this light, all organisms, irrespective of their lineages, are rational. Insofar as they have different internal constraints, such as intelligence or memory capacity, they adopt different actions while facing the same environmental constraints. The common environmental constraints define the diverse lineages, or their carriers, as in competition with each other. Under some environmental constraints, sharper intelligence or a sharper beak is favored by natural selection, i.e., the carrier of the trait has an absolute advantage over competitors. Under other environmental constraints, the carrier of greater intelligence might be unfit relative to other carriers.

So, natural selection does not select rationality, it rather selects the fittest lineage given the environmental conditions. If the focus is intelligence, the fittest lineage could be a low-

intelligence. But the carrier of low-intelligence is as rational as the carrier of high-intelligence. The set of constraints is the determinant which one is the fitter or what premium is worth paying for higher intelligence. So while every organism can be assumed to be involved in rationality optimization, their behavior is as diverse as the lineages. The set of identical environment constraints would determine which lineage should gain the upper hand..

As discussed above, to assume that organisms behave rationally does not mean that their action is rational in all circumstances. This would amount to a tautology, which is entailed in the “revealed preference” hypothesis [Samuelson, 1938]. According to the hypothesis, to state that a consumer chooses bundle A, while he or she can afford bundle B, must mean that he or she prefers A to B. However, this would violate some common observations about weakness of will, whose obvious manifestation is addiction and recklessness in humans.

But our focus here is not on the limits of rational choice theory [see Khalil, 2007b,c]. Rather, it is on the status of rationality *per se*. It has been shown that for rationality to be classified as a trait, it must be non-fungible. It must also be a non-scarce “resource.” To be scarce, rationality must be costly. It is the case that intelligence and other mental activity are costly. But this does not establish that rationality is costly. In fact, if rationality is costly, it must be a trait, which leads to the incoherence problem mentioned above. So, by deductive reasoning, rationality cannot be costly. This does not mean that the search for information is costless. The agent needs information to form belief. And the agent cannot keep on searching forever because search is, by definition, costly while the expected benefit from extra information decreases with further search. So, the agent must economize on the search effort. While the search effort is costly, the decision when to stop searching is costless. So, to state that rationality is costless does not mean that search itself is costless.

A good example of how rationality is costless is the phenomenon known as “moral hazard.” When the government started to impose seat belt laws, for instance, economists predicted that agents would drive less cautiously, resulting in more accidents but less (thanks to

the seat belt) fatalities per accident [Peltzman, 1975]. The tradeoff between safety and time gained by greater speed was never conscious. Also, agents who have insurance tend to adjust to the insurance, and adopt less-careful habits with regards to locking doors, making sure that the stove is off, and so on. Likewise, Brian Goff *et al.* [1997] show that baseball players in the American League adjusted their behavior in light of the 1973 change of the designated hitter rule.

Going to the classic example of Robinson Crusoe, he has to make a decision on whether to continue fishing or switch to hunting. So, he adopts a heuristic that states that, if he catches zero fish in the first two hours, then it is better to switch to hunting for the rest of the day. The establishment of this rule, like the search for resources, is costly. But the method used to determine the rule itself, i.e., rationality, is not costly. While the agent can economize on calculation, searching, and the use of other traits, he cannot economize on being economical. To make a choice implies that one is economical. Therefore, it is incoherent to assert that one can make a choice, or nature can make it on one's behalf, on how much rationality to use.

As a method, rationality is an endemic aspect of decisions. This "Rationality-qua-Method Thesis" entails that rationality is the way agents reconcile the objective function in light of the constraint function. And for the organism to do so, the method or the bridge between the two functions cannot be an element in either function. The method is rather about the use of non-method elements, such as eyesight, muscles, beak, and other tools needed for survival.

9. Conclusion

Why can't natural selection explain rationality? We have two answers. First, if rationality *is* "as-if" rationality (i.e., modular programs selected by natural selection), there is no rationality and, hence, there is nothing to explain. Second, if rationality exists, it would have a double-function, viz., as a trait and as a method of decision making. As a trait, what is the fittest lineage of rationality varies depending on the constraints. But as a method, there is no variation; it must

be uniform across constraints. This leads to a contradiction.

So, neo-Darwinism finds itself between a rock and a hard-place. To attain logical coherence, it quickly fall back to the position that the way organisms behave is selected by natural selection. Thus, there is no rationality; there are instead only instincts or programs of behavior. This means roughly one-to-one correspondence between phenotype and genotype. But this position (the rock) faces the empirical anomaly that agents are rational: They change action without being prompted by an exogenous genetic program. If there is a “turning on” of a genetic program, it takes place endogenously, i.e., in response to signals. To solve the empirical anomaly, neo-Darwinism must accept that organisms are rational. But this leads to the contradiction (hard place): How could rationality be of different lineages while, at he same time, be uniform?

This paper illustrated this incoherence in three ways. First, if we examine the matter from a non-Darwinian perspective, rationality is a trait that is susceptible to improvements. Such improvements can approach greater perfection along an abstract scale of progress. This means that, similar to other traits, it is always possible to improve rationality. But rationality by definition cannot be improved beyond 100 per cent rationality, as opposed to other traits. This thus entails that rationality cannot be a trait.

Second, within the neo-Darwinian framework, rationality cannot be seen as part of progress. To start with, there is no room for the notion of progress in neo-Darwinism. Neo-Darwinism assumes that traits are fit only with respect to given constraints. We cannot rank traits in the abstract. If rationality is a trait, natural selection might, under some constraints, favor agents with, say, 60 per cent rationality. Such irrational agents, given that they are the fittest, must therefore be making optimal decisions. This is a contradiction in terms.

Third, if rationality is a trait and if the agent can allocate resources intertemporally to improve such a trait, one can choose less rationality to obtain another trait. If so, one must be using rationality as a tool to make such optimal allocation of resources. How could one,

however, use rationality to make optimal decisions when it is already compromised and given up in favor of other traits?

Once one accepts that rationality is not the child of natural selection, one is forced into new vistas of understanding evolution beyond the confines of natural selection theory. This does not need to entail the invitation of medieval scholastics and the idea of “uncaused cause” as Darwinists [e.g., Dennett, 1995; Hodgson, 2002] have warned. It is true that Darwin undermined the metaphysics of divine intervention, but by accepting that rationality is uncaused by natural selection does not mean that it is beyond scientific study. It only means that such a study has not begun, and in fact was prevented from beginning given the dominance of the natural selection paradigm.

To wit, diverse evolutionary biologists have questioned the dominance of natural selection [Khalil, 1993]. A great portion of these heterodox evolutionary biologists accords a greater role to the individual and its development than is allowed by orthodox natural selection theory. The attention paid to the evolutionary role of the development of the individual, known as the EvoDevo approach, corresponds with the thesis of this paper about rationality. Rationality, after all, is about cognition and decision making that has consequences with regard to the development of the individual, which is the focus of the EvoDevo approach [Amundson, 2006].

While the term “EvoDevo” is relatively new, early pioneers of EvoDevo include Waddington [1953, 1957], Løvtrup [1974], Gould [1977] and Stanley [1979; see Khalil, 1993]. More recent literature includes the work of Matsuda [1987], West-Eberhard [1989], Raff [1996; *passim* Müller & Newman, 2003] and Hall *et al.*, [2004; see Callebaut & Rasskin-Gutman, 2005]. This literature is invariably critical of the *Organismus automaton* hypothesis. (In economics, the learning/development approach can be traced back to Joseph Schumpeter [e.g., Nelson & Winter, 1982; Dosi *et al.*, 1988; Nooteboom, 2000] and includes an Austrian strand of thought [e.g., Witt, 2003].)

The EvoDevo literature amply shows the plasticity of the form of organism. The form is the result of ontogenesis (development) which is partially a function of the environment—a theme championed by Lamarckism. The Lamarckian theme should be music to ears of the advocates of rationality theory who also stress the relevance of the environment in influencing behavior, which economists call “incentives.” In fact, mainstream, neoclassical economics has been called the science of incentives in that it is concerned with how agents change their behavior in light of changes in the environment. Of course, Darwinian evolution also stresses the role of the environment, but only as a selection force, and not as an inducement for the agent to shape its own development. That is, for Darwinian evolution, the environment enters only *ex post* to validate or invalidate an invention. In contrast, for developmental evolution, the environment enters also *ex ante* to inform and influence the invention itself.

References

- Alchian, Armen A. "Uncertainty, Evolution, and Economic Theory." *Journal of Political Economy*, June 1950, 58:3, pp. 211-221.
- Amundson, Ronald A. "EvoDevo as Cognitive Psychology." *Biological Theory*, Winter 2006, 1:1, pp. 10-11.
- Arthur, W. Brian. "Competing Technologies, Increasing Returns, and Lock-in by Historical Events." *Economic Journal*, March 1989, 99, pp. 116-131.
- Bonner, John Tyler. *The Evolution of Complexity by Means of Natural Selection*. Princeton, NJ: Princeton University Press, 1988.
- Brandon, Robert N. and Richard M. Burian. *Genes, Organisms, Populations: Controversies Over the Units of Selection*. Cambridge, MA: The MIT Press, 1984.
- Cairns, John, Julie Overbaugh, and Stephan Miller. "The Origin of Mutants." *Nature*, September 1988, 335:6186, pp. 142-145.
- Callebaut, Werner. "Self-Organization and Optimization: Conflicting or Complementary Approaches?" In Gertrudis van de Vijver, Stanley N. Salthe, and Manuela Delpos (eds.) *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization*. Dordrecht: Kluwer Academic, 1998, pp. 79-100.
- _____ and Diego Rasskin-Gutman (eds.). *Modularity: Understanding the Development and Evolution of Natural Complex Systems*. Cambridge, MA: MIT Press, 2005.
- Camerer, Colin, George Loewenstein, and Matthew Rabin (eds.) *Advances in Behavioral Economics*. New York: Russell Sage Foundation; Princeton, NJ: Princeton University Press, 2004.
- Cassill, Deby. "Skew Selection: Nature Favors a Trickle-Down Distribution of Resources in Ants," *Journal of Bioeconomics*, 2003, 5: 2&3, pp. 83-96.
- Cosmides, Leda and John Tooby. "Beyond Intuition and Instinct Blindness: Toward an Evolutionary Rigorous Cognitive Science." *Cognition*, 1994, 50, pp. 41-77.
- Darwin, Charles. *The Power of Movement in Plants*. New York: New York University Press, (1880) 1989.
- _____. *The Descent of Man*, intro H. James Birx. Amherst, NY: Prometheus Books, (1871) 1998.

_____. *The Origin of Species*, a various text, 1st pbk ed, edited by Morse Peckham. Philadelphia: University of Pennsylvania Press, (1859) 2006.

Dawkins, Richard. *The Selfish Gene*. New York: Oxford University Press, 1976.

_____. *The Extended Phenotype: The Long Reach of the Gene*. San Francisco: W.H. Freeman, 1982.

Dennett, Daniel C. *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.

_____. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon & Schuster, 1995.

Dewey, John. *Human Nature and Conduct: An Introduction to Social Psychology*. New York: Modern Library, (1922) 1957.

_____ and James H. Tufts. *Ethics*. New York: H. Holt and company, 1908.

_____ and Arthur F. Bentley. *Knowing and the Known*. In Rollo Handy and E.C. Harwood (eds.), *Useful Procedures of Inquiry*. Great Barrington, MA: Behavioral Research Council, (1949) 1999.

Detrain, C., J.L. Deneubourg, and J.M. Pasteels (eds.). *Information Processing in Social Insects*. Basel: Birkhäuser Verlag, 1999.

Dosi, Giovanni, Christopher Freeman, Richard Nelson, Gerald Silverberg, and Luc Soete (eds.). *Technical Change and Economic Theory*. London & New York: Pinter, 1988.

Eichinger, L. *et al.*, "The Genome of the Social Amoeba *Dictyostelium discoideum*." *Nature*, 5 May 2005, 435, issue no. 7038, pp. 43-57.

Fodor, Jerry A. *Representations: Philosophical Essays on the Foundations of Cognitive Science*. 1st ed. Cambridge, MA: MIT Press, 1981.

Franks, Nigel R., Eamonn B. Mallon; Helen E. Bray; Mathew J. Hamilton, Thomas C. Mischler. "Strategies for Choosing between Alternatives with Different Attributes: Exemplified by House-Hunting Ants." *Animal Behaviour*, 2003, 65, pp. 215-223.

Friedman, Milton. *Essays in Positive Economics*. Chicago: University of Chicago Press, 1953.

Ghiselin, Michael T. *The Economy of Nature and the Evolution of Sex*. Berkeley: University of California Press, 1974.

_____. "Biology, Economics, and Bioeconomics." In Gerard Radnitzky (ed.) *Universal Economics: Assessing the Achievements of the Economic Approach*. New York: ICUS Publications, 1992, pp. 55-104.

Gigerenzer, Gerd. "I Think, Therefore I Err." *Social Research*, Spring 2005, 72:1, pp. 195-218.

_____. "Bounded and Rational." In R.J. Stainton (ed.) *Contemporary Debates in Cognitive Science*. Oxford, UK: Blackwell, 2006, ch. 7, pp. 115-133.

Glimcher, Paul W., Michael C. Dorris, and Hannah M. Bayer. "Physiological Utility Theory and the Neuroeconomics of Choice." *Games and Economic Behavior*, 2005, 52, pp. 213-256.

Goff, Brian L., William F. Shughart II, and Robert D. Tollison. "Batter Up! Moral Hazard and the Effects of the Designated Hitter Rule on Hit Batsmen." *Economic Inquiry*, July 1997, 35, pp. 555-561.

Gould, Stephen Jay. *Ontogeny and Phylogeny*. Cambridge, MA: Harvard University Press, 1977.

_____. *The Structure of Evolutionary Theory*. Cambridge, MA: Harvard University Press, 2002.

_____ and Richard C. Lewontin. "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Program." In David R. Keller and Frank B. Golley (eds.) *The Philosophy of Ecology: From Science to Synthesis*. Athens, GA: University of Georgia Press, 2000, ch. 19, pp. 263-276.

Hall, Brian K., Roy D. Pearson, and Gerd B. Müller (eds.). *Environment, Development, and Evolution: Towards a Synthesis*. Cambridge, MA: MIT Press, 2004.

Herman, Louis M. "Intelligence and Rational Behaviour in the Bottlenosed Dolphin." In Susan Hurley and Matthew Nudds (eds.). *Rational Animals?* Oxford: Oxford University Press, 2006, (ch. 20) pp. 439-467.

Hirshleifer, Jack. "Economics from a Biological Viewpoint." *Journal of Law and Economics*, April 1977, 20:1, pp. 1-52.

Hodgson, Geoffrey M. "Darwinism in Economics: From Analogy to Ontology." *Journal of Evolutionary Economics*, 2002, 12:3, pp 259-281.

_____. *The Evolution of Institutional Economics: Agency, Structure, and Darwinism in American Institutionalism*. London: Routledge, 2004.

_____. "Biology and Economics: A Very Long Engagement." *Journal of Bioeconomics*,

2007, forthcoming.

Hurley, Susan and Matthew Nudds (eds.). *Rational Animals?* Oxford: Oxford University Press, 2006.

James, William. *The Principles of Psychology*. New York: H. Holt, 1893.

Kacelnik, Alex. "Meanings of Rationality." In Susan Hurley and Matthew Nudds (eds.). *Rational Animals?* Oxford: Oxford University Press, 2006, (ch. 2) pp. 87-106.

Khalil, Elias L. "Neo-classical Economics and Neo-Darwinism: Clearing the Way for Historical Thinking." In Ron Blackwell, Jaspal Chatha, and Edward J. Nell (eds.) *Economics as Worldly Philosophy*. London: Macmillan, 1993, pp. 22-72.

_____. "Economic Action, Naturalism, and Purposefulness." *Research in the History of Economic Thought and Methodology*, 1996, 14, pp. 119-140.

_____. "Economics, Biology, and Naturalism: Three Problems Concerning the Question of Individuality." *Biology & Philosophy*, April 1997a, 12:2, pp. 185-206.

_____. "Buridan's Ass, Uncertainty, Risk, and Self-Competition: A Theory of Entrepreneurship." *Kyklos*, 1997b, 50:2, pp. 147-163.

_____. "Survival of the Most Foolish of Fools: The Limits of Selection Theory." *Journal of Bioeconomics*, 2000, 2:3, pp. 203-220.

_____. "The Context Problematic, Behavioral Economics and the Transactional View: An Introduction to 'John Dewey and Economic Theory'." *Journal of Economic Methodology*, June 2003, 10:2, pp. 107-130.

_____. "Action, Entrepreneurship and Evolution." In Michel Weber (ed.) *Handbook of Whiteheadian Process Thought*. Frankfurt: Verlag, 2007a, forthcoming.

_____. "The Origin of Sin: A Two-Moment Theory of Integrity, Fairness and Other Internal Constraints." A mimeo, 2007b.

_____. "Moral Outrage: A Two-Moment Theory of Nationalism, Entrepreneurship and Other Internal Motivations." A mimeo, 2007c.

_____. "The Origins of Routines: Distinguishing Technological Routines from Cognitive Routines." *Revue de Philosophie Économique*, 2007d, forthcoming.

_____. "Natural Selection and Rational Decision: Two Concepts of Optimization." A mimeo, 2007e.

_____. "The Roadblock of Culturalist Economics: Douglass North and the Retreat of Economics." A mimeo, 2007f.

_____. "The Cognitivist Fallacy: The Heckman Puzzle." A mimeo, 2007g.

_____ and Alain Marciano. "Organomics *contra* Neo-Darwinism." A mimeo, 2007.

Kreps, David M. *A Course in Microeconomic Theory*. Princeton, NJ: Princeton University Press, 1990.

Landa, Janet T. "The Political Economy of Swarming in Honeybees: Voting-with-the-Wings, Decision-Making Costs, and the Unanimity Rule." *Public Choice*, 1986, 51, pp. 25-38.

_____ and Anthony Wallis. "Socio-Economic Organization of Honeybee Colonies: A Transaction-Cost Approach." *Journal of Social and Biological Structures*, July 1988, 11:3, pp. 353-363.

Levins, Richard and Richard Lewontin. *The Dialectical Biologist*. Cambridge, MA: Harvard University Press, 1985.

Lewontin, Richard C. "Adaptation." In Elliott Sober (ed.) *Conceptual Issues in Evolutionary Biology: An Anthology*. Cambridge, MA: MIT Press, 1984, pp. 235-251.

Llinás, Rodolfo R. *I of the Vortex: From Neurons to Self*. Cambridge, MA: MIT Press, 2001.

Løvtrup, S. *Epigenetics*. New York: Wiley, 1974.

Luria, Daniel. "Why Markets Tolerate Mediocre Manufacturing." *Challenge*, 1996, 39:4, pp. 11-16.

Matsuda, Ryuichi. *Animal Evolution in Changing Environments with Special Reference to Abnormal Metamorphosis*. New York: Wiley & Sons, 1987.

Mayr, Ernst. *The Growth of Biological Thought: Diversity, Evolution, and Inheritance*. Cambridge, MA: Harvard University Press, 1982.

McFarland, D.J. "Decision Making in Animals." *Nature*, 1 September 1977, 269, pp. 15-21.

Mousavi, Shabnam and Jim Garrison. "Toward a Transactional Theory of Decision Making: Creative Rationality as Functional Coordination in Context." *Journal of Economic Methodology*, June 2003, 10:2, pp. 131-156.

Müller, Gerd B. and Stuart A. Newman (eds.). *Origination of Organizational Form: Beyond the*

- Gene in Developmental and Evolutionary Biology*. Cambridge, MA: MIT Press, 2003.
- Nelson, Richard R. and Sidney G. Winter. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press, 1982.
- Ng, Yew-Kwang. "Complex Niches Favour Rational Species." *Journal of Theoretical Biology*, 1996, 179, pp. 303-311.
- North, Douglass C. *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press, 1990.
- Nooteboom, Bart. *Learning and Innovation in Organizations and Economies*. Oxford: Oxford University Press, 2000.
- Orzack, Steven Hecht and Elliott Sober. "Optimality Models and the Test of Adaptationism." *American Naturalist*, March 1994, 143:3, pp. 361-380.
- _____. (eds.) *Adaptationism and Optimality*. Cambridge: Cambridge University Press, 2001.
- Peltzman, Paul. "The Effects of Automobile Safety Regulation." *Journal of Political Economy*, August 1975, pp. 677-725.
- Pepperberg, Irene M. "Intelligence and Rationality in Parrots." In Susan Hurley and Matthew Nudds (eds.). *Rational Animals?* Oxford: Oxford University Press, 2006, (ch. 21) pp. 469-488.
- Plotkin, Henry. *Darwin Machines and the Nature of Knowledge*. Cambridge, MA: Harvard University Press, 1994.
- Quine, W. *Ontological Relativity and Other Essays*. New York: Columbia Univ. Press, 1969.
- Raby, C.R., D.M. Alexis, A. Dickinson, and N.S. Clayton. "Planning for the Future by Western Scrub-Jays." *Nature*, 22 February 2007, 445, pp. 919-921.
- Raff, Rudolf A. *The Shape of Life: Genes, Development, and the Evolution of Animal Form*. Chicago: University of Chicago Press, 1996.
- Rosen, Sherwin. "The Economics of Superstars." *The American Economic Review*, December 1981, 71:5, pp. 845-858.
- Roughgarden, Joan. *Evolution's Rainbow: Diversity, Gender, and Sexuality in Nature and People*. Berkeley: University of California Press, 2004.

Samuels, Richard. "Evolutionary Psychology and the Massive Modularity Hypothesis." *British Journal for the Philosophy of Science*, 1998, 49, pp. 575-602.

_____. "Massively Modular Minds: Evolutionary Psychology and Cognitive Architecture." In Peter Carruthers and Andrew Chamberlain (eds.) *Evolution and the Human Mind: Modularity, Language, and Meta-cognition*. Cambridge: Cambridge University Press, 2000, ch. 2, pp. 13-46.

Samuelson, Paul. "A Note on the Pure Theory of Consumers' Behaviour." *Economica*, 1938, 5, pp. 61-71.

Schoener, Thomas W. "Theory of Feeding Strategies." *Annual Review of Ecology and Systematics*, 1971, 2, pp. 369-404.

Schuck-Paim, Cynthia, Lorena Pompilio, Alex Kacelnik. "State-Dependent Decisions Cause Apparent Violations of Rationality in Animal Choice." *PLoS Biology*, December 2004, 2:12, pp. 2305-2315.

Schumpeter, Joseph A. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*, trans. by Redvers Opie, with a new intro. by John E. Elliott. New Brunswick, NJ: Transaction Publishers, (1912) 1989.

Simon, Herbert A. *Models of Man*. New York: John Wiley, 1957.

_____. *Models of Discovery*. Dordrecht: Reidel, 1977.

Skyrms, Brian. "Darwin Meets the Logic of Decision: Correlation in Evolutionary Game Theory." *Philosophy of Science*, 1994, 61, pp. 503-528.

Smith, Eric Alden and Bruce Winterhalder. "Natural Selection and Decision-Making: Some Fundamental Principles." In E.A. Smith and B. Winterhalder (eds.) *Evolutionary Ecology and Human Behavior*. New York: Aldine de Gruyter, 1992, pp. 25-60.

Sober, Elliott. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Chicago: University of Chicago Press, 1984.

_____. "Evolution and Optimality: Feathers, Bowling Balls, and the Thesis of Adaptationism." *Philosophic Exchange*, 1996, Number 26, pp. 41-57.

_____. "Three Differences Between Deliberation and Evolution." In Peter A. Danielson (ed.) *Modeling Rationality, Morality, and Evolution*. New York: Oxford University Press, 1998, pp. 408-422.

Sperber, Dan. "The Modularity of Thought and the Epidemiology of Representations." In Lawrence A. Hirschfeld and Susan A. Gelman (eds.) *Mapping the Mind: Domain Specificity in*

Cognition and Culture. Cambridge: Cambridge University Press, 1994, pp. 39-67.

Stanley, S.M. *Macroevolution: Pattern and Process*, 1979. San Francisco: Freeman.

Steele, Edward J., Robyn A. Lindley, Robert V. Blanden. *Lamarck's Signature. How Retrogenes Are Changing Darwin's Natural Selection Paradigm*. Reading, MA: Perseus Books, 1998.

Stephens, Christopher L. "When is it Selectively Advantageous to Have True Beliefs? Sandwiching the Better-Safe-than-Sorry Argument." *Philosophical Studies*, 2001, 105:2, pp. 161-189.

Stich, Stephen. "Could Man be an Irrational Animal?" *Synthese*, 1985, 64, pp. 115-135.

_____. *The Fragmentation of Reason*. Cambridge, MA: MIT Press, 1990.

Strassmann, Joan E., Yong Zhu, and David C. Queller. "Altruism and Social Cheating in the Social Amoeba *Dictyostelium discoideum*." *Nature*, 21 December 2000, 408, pp. 965-967.

Tauber, Alfred I. (ed.) *Organism and the Origins of Self*. Dordrecht: Kluwer, 1991.

Tullock, Gordon. "The Coal Tit as a Careful Shopper." *American Naturalist*, 1971, 105, 77-80.

_____. *The Economics of Non-Human Societies*. Tucson, AZ: Pallas Press, 1994.

Veblen, Thorstein. "Why is Economics not an Evolutionary Science?" *Quarterly Journal of Economics*, July 1898, 12:3, pp. 373-397.

Vermeij, Geerat J. *Nature: An Economic History*. Princeton, NJ: Princeton University Press, 2004.

Waddington, C.H. "Genetic Assimilation of an Acquired Character." *Evolution*, 1953, 7, 118-6.

_____. *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*, with an appendix by H. Kacser. London: Allen & Unwin, 1957.

West-Eberhard, Mary Jane. "Phenotypic Plasticity and the Origins of Diversity." *Annual Review of Ecological Systems*, 1989, 20, pp. 240-278.

Wilson, David Sloan and Elliott Sober. "Reviving the Superorganism." *Journal of Theoretical Biology*, 1989, 136, pp. 337-356.

Witt, Ulrich (ed.). *The Evolving Economy: Essays on the Evolutionary Approach to Economics*. Edward Elgar, Cheltenham, 2003.