



Applying evolutionary methods in economics: progress or pitfall?

Vaios Koliototis¹

© The Author(s) 2020

Abstract

The Darwinian theory of evolution has arguably become an important building block for experimental and theoretical economists. According to Burnham (J Econ Behav Org 90:S113–S127, 2013), it is possible to formulate novel hypotheses and predictions about human preferences, on the basis of what patterns of behavior would have been adaptive in the ancestral environment. After clarifying two theoretical concepts, the Adaptively Relevant Environment and fitness maximization, I argue that multiple scientifically plausible hypotheses about human preferences are compatible with evolutionary models that target behavior. Moreover, I propose a refinement of Burnham’s method based on theoretical resources provided by the indirect evolutionary approach. Economists apply or build evolutionary models of their own that target particular features of human psychology and cognition. Such models may reduce the number of plausible hypotheses to allow for rigorous scientific testing in laboratory or field experiments.

Keywords Evolution · Indirect evolutionary approach · Social preferences · Methodology

JEL classification B52 · B41 · C79 · D01 · D64

1 Introduction

Over the past decades there is a growing interest in building bridges between evolutionary biology and economics. Widely cited studies published in leading economic journals apply evolutionary arguments and theories to explain and discover features of human psychology and cognition that have a role in economic behavior (for recent reviews, see Collins et al. 2016; Burnham et al. 2016).

✉ Vaios Koliototis
55789vko@eur.nl; vkoliotot@hotmail.com

¹ Erasmus Institute for Philosophy and Economics, Erasmus University Rotterdam, Rotterdam, Netherlands

Many economists hold that the insights of evolutionary theory might be relevant for economics in that they can help improve our understanding of human preferences and behavior. Looking back into the writings of Jack Hirshleifer, a pioneer economist that examined what evolutionary biology can contribute to economics, “[t]he programmatic contention here is that such preference patterns, despite seemingly arbitrary elements, have survived because they mainly are adaptive to environmental conditions” (Hirshleifer 1977, p. 18). In the economic literature human preferences are taken as given. For Jack Hirshleifer, evolutionary theory could help economics with identifying what preferences we have and explaining why they survived selection processes (Hirshleifer 1977, p. 17).

It is still unclear, however, what method can facilitate economists in formulating evolutionary hypotheses about human preferences and behavior. Terence Burnham presents a promising evolutionary method that aspires to reconcile different sub-fields within economics (Burnham 2013). In this paper, I set aside prospects for a broad integration within economics or among the social sciences. My relatively modest aim is to clarify and refine Burnham’s proposal. It is also not my intention to present and critically assess alternative methods offered by evolutionary psychologists.¹ Instead, I examine Burnham’s method on its own terms and consider whether it can be utilized to inspire productive work at the interface between economics and biology.

In Burnham’s view, it is possible to formulate novel hypotheses and predictions about human preferences, on the basis of what patterns of behavior would have been adaptive in the Adaptively Relevant Environment (henceforth ARE). After clarifying two crucial theoretical terms that figure in Burnham’s method, the ARE and maximizing behavior, I suggest that identifying the ARE requires economists to describe a particular trait and build or pick an evolutionary model that identifies a set of selection pressures against which it would confer an evolutionary advantage. With this in mind, I further argue that evolutionary models of adaptive behavior do not routinely entail particular insights about human psychology. The problem this raises for Burnham’s method is that several scientifically plausible hypotheses about human preferences will be compatible with such evolutionary models. Finally, based on the theoretical resources provided by the indirect evolutionary approach (henceforth IEA), I argue that Burnham’s method can be more reliable if economists apply or build evolutionary models of their own that target particular features of human psychology and cognition. Such models may reduce the number of plausible hypotheses to allow for rigorous scientific testing in laboratory or field experiments.

Based on these arguments, the revised method proposed in this paper is the following:

- i. Identify preferences and overt behaviors related to them.
- ii. Apply or build an evolutionary model and use information about our ancestral past to determine the ARE of these traits;

¹ For an insightful discussion of these methods, see Griffiths (1996).

- iii. Predict the structure of preferences and how they operate in the present environment;
- iv. Test these predictions in the laboratory and the field.

The paper is structured as follows. In the next section I introduce Burnham's method, focusing on concepts that are more relevant for its application in theoretical and empirical research. Moreover, I set forth what I consider to be the main problems with his proposal. Following this, in Sect. 3, I examine what kind of evolutionary models can be applied to shed light into our psychological traits. Section 4 explores in more depth whether it is reasonable to assume that decisions in the ancestral past were based on utility maximization. Sect. 5 concludes.

2 Burnham's method and the adaptively relevant environment

This section begins by presenting Burnham's proposal and focuses on two key concepts that are in need of disambiguation: The Adaptively Relevant Environment and maximizing behavior. Next I examine whether it is possible to search for and subsequently frame hypotheses and predictions about human preferences, using only evolutionary models that explain patterns of behavior.

Burnham's research strategy involves the following four steps:

- i. Provide a model of the Adaptively Relevant Environment (ARE);
- ii. Solve for maximizing behavior in the ARE and predict the structure of preferences;
- iii. Predict how these evolved preferences in the ARE operate in the present environment;
- iv. Test these predictions and refine.

According to Burnham, economists begin their research by examining the human ancestral environment, trying to find therein insight into what sorts of behaviors could have evolved by selection. In particular, economists identify human behavior that was maximized in the ARE. For Burnham, this is sufficient to infer hypotheses about our psychology and cognition and predict the structure of preferences that produce the relevant pattern of behavior. After examining how these preferences operate in the current environment, a test of these predictions is performed in the laboratory or the field.

Burnham's method is related to the following evolutionary hypothesis. Human behavior is the product of psychological traits that evolved in the ancestral environment. We, however, live in an environment that is very different from the one in which these traits evolved. Back in the Pleistocene savannah, for example, there were no fast food restaurants, contraceptives, antibiotics or automobiles. In a novel environment, our psychological traits often result in dysfunctional behaviors. If significant changes in the environment actually occurred, one might expect that many human behaviors would be detrimental to our survival or reproduction. Kanazawa

(2003) calls this the Savannah Principle, Hagen and Hammerstein (2006) the mismatch hypothesis, Richerson and Boyd (2006) the adaptive-lag hypothesis and Burnham and Johnson (2005) the evolutionary legacy hypothesis (henceforth ELH).

The standard example is our preference to eat food containing sugar, fat and salt. This preference evolved in an environment where such substances were important to health, though they were difficult to obtain. In current environments however it is easy to buy and store food with sugar and fat. Our preference to eat such food evolved in the ancestral environment and is still present today, with obvious negative effects to our health.

At first sight, Burnham's proposal is neat with clear-cut steps that are easy to follow. But on closer look, it turns out that key concepts are ambiguous and in need of clarification.

An important concept in Burnham's method is the ARE (Irons 1998). In evolutionary research, many traits or characters (e.g. psychological, cognitive, morphological, behavioral) evolved in the context of the selection pressures that operated in the ancestral past. Selection pressures typically refer to features of the environment that exert an influence on reproductive success, driving the evolution of traits. They can take many forms, such as climatic conditions, habitat, mates, migration, geographical area, availability and quality of food, energy resources, predators, diseases, population size as well as other conspecifics.

The ARE describes the main selection pressures thought to have led to the evolution of a particular trait throughout its evolutionary history². However, evolutionary researchers should not equate the ARE with the Pleistocene, a period that our ancestors spent as hunter-gatherers in the African savannah, prior to the emergence of agriculture about 10,000 years ago. Some features of the African savannah may have had little to do with the evolution of particular traits. The ARE is also not necessarily tied to a particular epoch like the Pleistocene or the Holocene. Some traits have a much longer evolutionary history or have evolved much more recently in the last 10,000 years. Instead, the ARE describes key features of different environments that influenced the evolution of a particular trait.

How can economists reconstruct the ARE? Burnham is silent on this issue. As I see it, evolutionary theory can help us track the selection pressures that might have led to the evolution of particular traits. An economist interested in evolutionary research has to describe in some detail a particular trait and pick or build an evolutionary model that identifies a set of selection pressures against which that particular trait could have evolved. If the selection pressures identified by the evolutionary model are supported by empirical evidence, evolutionary researchers can infer that they are part of the ARE.

² Evolutionary researchers often apply the distinct concept of environment of evolutionary adaptedness (EEA for short), the environment "that man inhabited for two million years until changes of the past few thousand years led to the extraordinary variety of habitats he occupies today" (Bowlby 1969, p. 58). In contrast to the EAA that describes a uniform background to which our traits adapted in the ancestral past, ARE is trait-specific.

This implies that one has to add a step in Burnham's method. Knowledge of a trait comes first and the application of an evolutionary theory to determine the ARE follows. Evolutionary theory, as any other theory, typically aims to explain traits that are known to be present in contemporary or past populations. Without some initial information about the target trait, it is not possible to describe it in some detail and build an evolutionary model that identifies its ARE.

The second key concept in Burnham's method is "maximizing behavior". The term is however ambiguous since Burnham does not clarify the maximand: Is it fitness, utility, income, well being or happiness? In what follows I will make an attempt to shed some light on this term by taking into consideration how "maximizing behavior" is typically understood in evolutionary biology.

Evolutionary scientists often suggest that a type of behavior is adaptive if it maximizes fitness, i.e. reproductive success or a currency conducive to survival and reproduction. A generalization of Darwinian fitness in case of social behaviors is inclusive fitness, the sum of the direct fitness gained through the influence of an individual's behavior on her production of offspring and the indirect fitness derived from contributing to the production of offspring of genetically related individuals.

A recent theoretical controversy revolves around the generality of inclusive fitness maximization and questions whether it can be used as a tool for understanding the evolution of human and animal traits. According to Alan Grafen's work, evolutionary researchers can expect fitness maximizing behavior to evolve only under particular conditions (Grafen 2006, 2014). An important theoretical conclusion is that inclusive fitness maximization does not apply to cases of frequency-dependent selection and non-additive fitness effects.³ Whether Grafen's results can be generalized and extended to the more realistic nonadditive case is an open theoretical question and has not yet been settled.

It is beyond the scope of this paper to advance particular theoretical arguments for or against the principle of inclusive fitness maximization (e.g. Lehmann et al. 2015). Given that this principle is controversial and its generalization has not yet been achieved, there is no reason to incorporate it in an evolutionary methodology that can be applied by economists and other social scientists. Moreover, such a principle is not necessary to the study of the evolution of human behavior. Instead of assuming that an organism's behavior maximizes inclusive fitness, evolutionary researchers can apply standard evolutionary models that examine whether particular traits were favored or disfavored by natural selection.

Let us stop for a moment and see where we have come. Burnham presents a methodology whereby economists begin their research with the ARE and identify therein maximizing (adaptive) behavior. A detailed examination of two key concepts, the ARE and fitness maximization, suggests the following refinement of Burnham's method. To determine the ARE economist have to start their investigation with a particular trait. And instead of taking for granted the controversial principle of

³ Alan Grafen, one of the proponents of the optimization programme in evolutionary biology, writes that "the assumption of additivity is made throughout this paper, but is not in general a realistic assumption. In many applications, non-additivity is an important part of the problem" (Grafen 2006, p. 543).

(inclusive) fitness maximization, they can apply an evolutionary model that explains why this trait evolved in the ancestral past. After identifying the ARE of that trait, one can infer hypotheses about the presence of particular psychological traits. With these points in mind, one can identify two problems with Burnham's method that call for another refinement.

First, according to the ARE concept, economists cannot simply assert that we live in a novel environment, distinct from the Pleistocene environment, as Burnham and other evolutionary researchers have done in a variety of cases. Economists might anticipate that modern industrial societies are different from Pleistocene hunter-gatherer societies. However, it is entirely possible that some key features of ancestral environments that led to the evolution of a specific trait have remained roughly the same. The ARE is always identified for a particular trait and different traits often evolved against a different ARE. Without a detailed analysis of the trait of interest and the relevant features of past and present environments, economists cannot know to what extent the ARE differs from relevant features of the current environment.

Second, there is a problem of inference in the second step of Burnham's method, the move from models that target behavior to hypotheses about human preferences. What is important to emphasize is that when evolutionary theorists build models that assume fitness-maximization, their target of analysis is behavior, not human psychological or cognitive traits. Evolutionary game theory and evolutionary optimality models for example do not suggest that evolutionary processes select psychological traits that generate fitness-maximizing behavior. Instead, they remain agnostic about the specific neural, cognitive or psychological traits that guide animal and human behaviors. What these models explain is why patterns of behaviors are selected in the ancestral past. To the extent that multiple underlying psychological traits result in the same fitness adaptive behavior, Burnham's method will not be a powerful guide in telling which hypotheses about human psychology are more plausible. To the economist applying this method, many hypotheses about preferences appear to be just as good as any other.

The more general problem is that economists cannot simply assume there is only one psychological trait associated with adaptive behavior. Such a one-to-one mapping is very unlikely. It is plausible, I think, to accept that psychological traits are typically involved in behavior and it is through behavior that these traits confer to an organism a fitness advantage. Any given adaptive behavior however can be the result of a variety of psychological traits and each of them might have a different evolutionary history. The implication is that economists cannot lose sight on the numerous relations between adaptive behavior and psychological traits. And unless there are reasons independent of evolutionary models to expect that there are not many different psychological traits that can potentially support the same adaptive behavior, Burnham's method would not be a reliable guide in formulating hypotheses about human psychology.⁴

⁴ Although Burnham does not present additional resources that can help researchers to constrain hypotheses about psychological traits, using knowledge from cognitive psychology is a reasonable suggestion. This, however, raises the following problem: Cognitive science is not unified at the theoretical level. There are multiple theories (e.g. modularity hypothesis, heuristics and biases, folk-psychology, dual-pro-

Before I tackle this problem in the next section, it is useful to give empirical content to Burnham's proposal. He argues that his methodology provides the following novel hypothesis: High levels of altruistic behavior observed in economic experiments (and contemporary societies) arise from human psychological and cognitive traits that evolved in an ancestral environment of repeated interactions, and reputation was constantly at stake.

To arrive to this hypothesis, evolutionary researchers start with a general description of the environment faced by our Pleistocene hunter-gatherers: They lived in small groups and repeated interactions with other group members may have been an important feature of their lifestyle. In this sort of environment, it would have been adaptive for our ancestors to cooperate and protect their reputation. Burnham and other evolutionary researchers suggest that altruistic behavior in one-shot and anonymous interaction in contemporary environments is the result of human psychology fine-tuned to operate in repeated and public interactions. Evolutionary researchers conduct a series of experiments designed to test predictions derived from this hypothesis.⁵

What I emphasized so far is that evolutionary researchers have to apply an evolutionary model to determine the ARE for a particular trait. Instead, evolutionary researchers provide a very general description of the hunter-gatherer lifestyle, without a detailed analysis of the actual selection pressures that led to the evolution of social behavior (e.g., Burnham and Johnson 2005; Haley and Fessler 2005). In particular, it is not clear which evolutionary theory and therefore which adaptive behaviors and selection pressures these researchers consider to derive their hypothesis from.

For some researchers, direct reciprocity theory (Trivers 1971) is assumed to be largely correct and cooperation among non-kin evolved in repeated interactions (Burnham and Johnson 2005, p. 130). In evolutionary biology, however, reputational effects are analyzed by indirect reciprocity theory (e.g. Nowak and Sigmund 2005). Both indirect and direct reciprocity models can explain the evolution of human cooperation. They demonstrate that cooperation evolves because it ultimately results in a net fitness benefit. The behaviors however examined by these theories are very different. The evolution of cooperation by direct reciprocity is often related to the success of "Tit-for-Tat" (TFT) in computer tournaments (Axelrod 1984). This type of behavior cooperates on the first round and thereafter it repeats the previous behavior of its opponent. In contrast, indirect reciprocity theory implicitly assumes that the same individuals will not meet again (Nowak and Sigmund 2005). Interactions are not repeated, but one shot. In particular, reputation is based on an image

Footnote 4 (continued)

cess theory etc.) and no consensus view has emerged as to which of these theories is more plausible. I discuss this issue in more detail in Sect. 4.

⁵ Empirical support for this hypothesis is mixed. Many experiments demonstrate that reputation-related cues such as displaying images of eyes promote cooperative behavior (e.g. Haley and Fessler 2005; Burnham and Hare, 2007). Other experimental studies suggest that being observed by such cues do not affect experimental behavior (e.g. Fehr and Schneider 2010). There is currently an ongoing debate about interpretation of experimental results (a short review can be found in Nettle et al. 2013).

score that increases if a player cooperated in past interactions with other players and decreases if they did not do so. While in indirect reciprocity models cooperative behavior is based on how an individual has behaved against others, in direct reciprocity models cooperation rests on how an individual has behaved with the same partner in the past.

Suppose that advocates of the ELH manage to solve this issue and direct reciprocity provides the explanation for the evolution of human cooperation. Now according to Burnham's method, it is possible to predict our preferences from a conditional cooperative behavior like Tit-for-Tat.⁶ However, this behavior does not single out a particular type of preference. For example, an evolutionary explanation based on Tit-for-Tat does not allow economists to suggest that cooperation is motivated by selfish preferences. While it can be the case that the preferences evolved were selfish, conditional cooperation in repeated interactions could also be served, among others, by a genuine altruistic preference (Trivers 1971, p. 51), an impure altruistic preference (Andreoni 1990) or by a preference for fair outcomes. Hence, knowing that cooperation in our species most likely evolved by direct reciprocity does not provide a particular insight into what preferences have evolved in the ancestral environment.

This example points to the following problem. Although we might have reasons independent from direct reciprocity models to think that the underlying preferences are, for example, selfish, knowing that the correct explanation for the evolution of cooperative behavior is provided by direct reciprocity theory and that a conditional behavior like Tit-for-Tat evolved in the ancestral past provides no particular insight into human preferences. This might allow someone to claim that even if the modified version of Burnham's method generates a novel hypothesis that ends up being true after experimental research, it is only accidentally true. That is, it is true independently of the evolutionary model applied and this undermines the idea that Burnham's method can systematically provide new knowledge about human psychology.

The conclusion of this section is that hypotheses about human preferences are often inferred from an evolutionary model that focuses on fitness outcomes of behaviors. However, it is not possible to ascertain the presence of particular motives, preferences, desires or intentions based on such models. Put differently, it will routinely be possible to identify for some behavior and its ARE, multiple scientifically plausible hypotheses about human psychology. Based on such evolutionary models alone, it is hard to see how scientific progress could be made in determining which of these hypotheses are worth of experimental research. In the section that follows, I consider a way to refine Burnham's method and reduce the search space to a smaller number of hypotheses.

⁶ To be clear, Tit-for-Tat is not always evolutionary successful in an indefinitely iterated Prisoner's Dilemma (e.g. Binmore 1998).

3 Evolutionary models in economics: evolutionary game theory and the indirect evolutionary approach

At the core of Burnham's method is evolutionary theory and this would be the focus in this section. Recall from the previous discussion that evolutionary researchers apply evolutionary models to determine the selection pressures for particular traits. Before evolution models are applied, one has to determine which traits are likely candidates for natural selection explanations. An open question remains: What kind of models should economists apply to infer hypotheses about preferences? We saw that knowing the right evolutionary model for some behavior does not tell economists what kind of preferences are more plausible from an evolutionary perspective. In what follows my aim is to present an alternative way of evolutionary theorizing that can successfully rule out all but a few hypotheses about human preferences.

It is possible to distinguish between two relevantly different types of evolutionary models. In one type the primary target of analysis is a psychological or cognitive trait internal to the organism. These models investigate evolutionary processes that produced the human psychologies that, in turn, contribute to human behaviors. Other types of models remain agnostic about human psychology and cognition and attempt to explain the evolution of particular behaviors.

To illuminate this distinction, in what follows I examine evolutionary game theory (henceforth EGT) and models of the IEA. What distinguishes the IEA from EGT is the shift of focus from behaviors to underlying preference patterns. EGT provides evolutionary explanations of behaviour (or behavioural strategies). In contrast, models of the IEA explain the evolution of preferences directly and behavior indirectly. And as we will see, such models can narrow down the range of plausible hypotheses about human psychology that can be examined in the laboratory or the field.

3.1 Evolutionary game theory

Evolutionary research applies several approaches to modeling. Though one particular cluster of models can, however, be singled out: That of EGT. As we have seen in the previous section, Burnham's proposal is just an effort to hypothesize and predict preferences from evolutionary games. Models of direct or indirect reciprocity theory are assumed to be correct, at least in their general outlines. Based on such models, economists attempt to predict which human preferences have evolved in past environments.

EGT and evolutionary dynamics were originally developed to explain why certain types of (supposedly genetically determined) behavioral strategies⁷ are promoted by natural selection and stabilize in a population (e.g. Maynard Smith 1982). An evolutionary game refers to evolutionary processes such as natural selection and

⁷ A strategy refers to "a behavioral phenotype; i.e., it is a specification of what an individual will do in any situation in which it may find itself" (Maynard Smith 1982, p. 10). In this paper a strategy is "a way of behaving" (Kitcher 1985, p. 89), manifest behavior traceable to genes, often conditional to the "situation" an organism is in.

environment conditions such as repeated interactions or population structure that made a significant positive difference to the fitness of some organism that has a particular behavior.⁸

From an economist's perspective, what are the specific characteristics of evolutionary games? In a short but illuminating presentation of evolutionary models in sociobiology and game theory, D'Arms et al. (1998) observe that evolutionary game theorists seek to explain "behavior by pointing to adaptive advantages for those who engage in it, without attempting to explain how exactly tendencies to behave in the relevant way are embodied in a psychology" (D'Arms et al. 1998, p. 87). The question they attempt to answer is why a given type of behavior has evolved due to the winning process of selection, compared to many other behaviors that could have evolved.

Evolutionary game theorists, however, seek to explain types of behavior by examining the fitness benefits accruing to individuals, without providing an explicit account of preferences and decision-making that influence behaviors.⁹ To be clear, it is recognized that it is important to understand the fine detail of human psychology (Trivers 1971; Axelrod 1984). However, cognitive or psychological traits are not an integral part of EGT. They are studied after building an evolutionary model and they are typically conceived as mere add-ons in the evolutionary explanation these models provide.

That said, this particular feature of EGT is also characteristic of much of evolutionary theorizing in economics. In a recent in-depth study on the evolution of human cooperation, Samuel Bowles and Herbert Gintis build evolutionary models based on parochialism, gene-culture coevolution theory and strong reciprocity (Bowles and Gintis 2013). These evolutionary models focus on fitness and behavior, without examining questions of motivation and decision-making (ibid.: p. 168). In particular, they do not assume that individuals in a population face a choice between alternative options. Instead, evolution simply programs or "hard-wires" individuals with behavior conditional to norms. Moreover, these evolutionary models do not speak out on the issue of what passions, sentiments or neural mechanisms underlie social behaviors. In the ending chapters of their book, Bowles and Gintis present a distinct set of models that examines processes of norm internalization and social emotions like guilt and shame that bring about a particular type of behavior like altruistic punishment (Bowles and Gintis 2013, Chapters 10 and 11).

Let us take stock. In Burnham's method, models based on EGT are applied with the aim to provide novel hypotheses to explore in economic experiments. These evolutionary models focus on behavior, typically assumed to be genetically transmitted

⁸ Of particular importance to EGT are models that explore dynamic processes of Darwinian selection such as the replicator dynamic. However, dynamic processes often examined in EGT "do not build in any assumptions on behavior or knowledge, other than the basic principle of differential selection – apparently successful behavior increases its representation in the population, while unsuccessful behavior does not" (Mailath 1998, p. 1355).

⁹ For example, according to Axelrod "there is no need to assume that the players are rational. They need not be trying to maximize their rewards. Their strategies may simply reflect standard operating procedures, rules of thumb, instincts, habits, or imitation" (Axelrod 1984, p. 18).

across generations. Moreover, the description of behavior in terms of human decision-making and human psychology is either very coarse or completely absent. To be clear, this does not make EGT trivial or empty of explanatory content. However, EGT is disconnected from any scientific programme that studies human decision-making and psychology. In the evolutionary explanations of EGT, nothing needs to be said (and in most cases, nothing is actually said) of how our mind functions. To determine what preferences follow from such evolutionary models, additional assumptions have to be made that are not part of the explanations these models provide.

It becomes also clear what kind of evolutionary theory can be more reliable in formulating hypotheses about human psychology. Evolutionary models that start from a more detailed description of human behavior in terms of decision-making and human psychology can also deliver a more refined description of the ARE. In turn, this can significantly reduce the range of plausible hypotheses about human psychology that can be systematically examined in the laboratory or the field.

3.2 The indirect evolutionary approach

One way to overcome the limitations of EGT and refine Burnham's method is the following: economists can actively engage in evolutionary modeling and give evolutionary explanations of their own in which human psychology and decision-making receive their due. In fact, economists often analyze preferences and decision-making from an evolutionary perspective, asking why we have particular psychological traits and seeking answers in terms of the contributions these traits have made to reproductive success. A case in point is the so-called Indirect Evolutionary Approach (IEA) initiated by Güth and Yaari (1992) and subsequently applied to investigate the evolution of other-regarding preferences (e.g. Akçay et al. 2009; Alger and Weibull 2013, 2016; Heller and Mohlin 2019).

The IEA aims to provide an evolutionary explanation of preferences patterns by asking what sort of selection pressures lead to their evolution and maintenance in a population. A robust mathematical result of the IEA is that self-regarding preferences, a standard assumption in much economic research, is not satisfactory from an evolutionary perspective. While social preferences often result in cooperative behavior and can be easily exploited, in the selection pressures formalized by these models, a combination of self-regarding and social preferences is more evolutionary stable than self-regarding preferences.

Let me give an example. In a series of papers Alger and Weibull build evolutionary models inspired by the work of Güth and Yaari (e.g. Alger and Weibull 2013, 2016). They also assume that individual behavior is based on utility maximization, while evolutionary success is driven by payoffs related to reproductive success. They demonstrate that in simplified environments described by their models, natural selection leads to the evolution and maintenance of a certain family of preferences, convex combinations of selfish and social preferences. Their result is about the evolution and stability of preference types rather than behavior. A particular type of preference with a Kantian flavor called "Homo Moralis" has an evolutionary

advantage. “Homo moralis” gives some weight to the principle of “acting according to that maxim whereby you can, at the same time, will that others should do likewise with some probability.” (Alger and Weibull 2016, p. 57).

On a more technical note, in the IEA a large heterogeneous population consists of players with different preference types, modeled as parameters in utility functions. In accord with standard economic theory, players make decisions and calculate best responses, i.e. they are rational in that they maximize a utility function whenever called upon to make a choice. These players are matched in pairs and on the basis of the payoff structure of the game, preference types are inherited to the next generation. In the next period their offspring are matched to play the same game again. An evolutionary process based on fitness determines the long-run distribution of preference types in the population.

Some clarifications are necessary. In economics, the interpretation of utility and preferences is matter of dispute. Preferences and utility are typically given either a psychological (mental) or behavioral interpretation (e.g. Okasha 2016; Dietrich and List 2016). For some economists, preferences refer to psychological states inside the brain of decision makers. Others take preferences to be patterns of observable behavior. What must be emphasized for the IEA is that preferences and behavior are distinct theoretical constructs. In these models, a utility function is assigned to an individual player and a player’s flexible choice of behavioral strategies depends on preferences that mutate and transfer from one generation to another. In other words, in the IEA preferences refer to psychological traits manifested in behavior and these traits are acquired through genetic inheritance.

Moreover, fitness maximization is not an assumption of the IEA. If we take an IEA perspective and examine whether natural selection will favor or oppose psychological traits that guide human behavior, there is no need for a principle like (inclusive) fitness maximization. Models of the IEA distinguish between utility maximization driving the choice of behaviors and the fitness consequences of behavior. Based on economic theory, the assumption is that players will choose strategies to maximize their utility. However, utility differs from evolutionary fitness that determines the long-run survival of preferences in a population. While individuals may act rationally by choosing behavioral strategies that maximize utility, the behaviors chosen do not need to maximize reproductive success.

While there is much more that can be said about the IEA¹⁰, for present purposes it is enough to emphasize that IEA is different from EGT in the following way. In contrast to models that examine pre-programmed behavior, the IEA ascribes utility functions to players and models the evolution of preference types. The focus on preferences as the primary target of analysis and the assumption that decisions are made based on rational utility maximization is what distinguishes the IEA from

¹⁰ An important feature of these models is that preference types are assumed to be (at least partially) observable. The degree to which these preferences are observed influences the outcome of preference evolution (e.g. Ok and Vega-Redondo 2001). In their review of evolutionary models in economics, Robson and Samuelson (2011) consider this as a shortcoming of the IEA. For a recent model in which preferences and the cognitive capacity to conceal and reveal them co-evolves, see Heller and Mohlin (2019).

EGT. However, EGT is not different from the IEA in other respects. What they share in common is theorizing based on tractable mathematical models, solution concepts like evolutionary stable strategies and stable fixed points as well as equations that describe evolutionary dynamics (like the replicator equation).

How can this type of evolutionary models refine Burnham's method and contribute to our understanding of human psychology? One way is by providing an evolutionary explanation. Applying models of the IEA can help economists explain psychological traits that are already known to be present in human populations. If we already have a good understanding of our psychology through laboratory and field experiments, evolutionary theorizing can help us determine the selection pressures that led to the evolution of psychological traits. Moreover, these models can contribute to economics even if we do not have yet a good grasp of human psychology. Determining the selection pressures might suggest hypotheses about additional features or details of the preferences included in the model. In both cases, the IEA can help economists to find out more about traits that they are not completely ignorant about and are worth exploring further.

In particular, evolutionary models such as those of the IEA can contribute to the economic literature and the empirical study of preferences in the following ways.

First, when models based on the IEA suggest novel utility functions, like in the research of Alger and Weibull, the theoretical investigation into the evolutionary foundations of human sociality can be complemented by empirical research to find out whether participants in controlled laboratory experiments and the field actually have the psychological traits attributed to players by utility functions included in the model¹¹.

Second, behavioral economics have developed mainly with the aim to build simple utility-based models to explain a wide variety of evidence in experimental research. These models expand the argument in utility functions to include a desire or preference for social outcomes. Inequity aversion and fairness utility functions are examples of such models (Fehr and Schmidt 1999; Bolton and Ockenfels 2000). In an experimental setting, it is possible to empirically examine how utility functions of evolutionary models fare against those typically suggested by other behavioral economists.

Third, by looking at a utility function it is often not immediately clear what kind of psychological processes underlie human decision-making. Preference in the IEA is an abstract concept that precedes choice and refers to a vast array of motivating factors (e.g. emotional commitments, aversions, desires). If, however, the explanations of these evolutionary models are found plausible, hypotheses can be formed about the details and properties that these preferences have. These hypotheses can be tested in the laboratory and the field. For example, neuroeconomics might provide a

¹¹ An interesting application of the IEA can be found in the work of Clavier and Chapuisat (2016). They investigate the details of the argument contained in a family of utility functions advanced by Alger and Weibull and conclude that genuine altruistic motivations can be found in long term reciprocal interaction and kin and romantic relationships.

deeper understanding of human preferences by examining corresponding arrangements of neurons within the nervous system.¹²

There is, however, an important issue that needs to be addressed. Economists that apply the IEA do not make an attempt to investigate whether the selection pressures described by their models reflect conditions present in past environments. In evolutionary theory the relevant context is historical, namely the ARE that impinged upon particular animal and human traits. The emphasis however of the IEA (and many models of EGT) is on the mathematical aspects of evolutionary theory. In these models the result “springs out from the mathematics” (Alger and Weibull 2013, p. 2271) and “mathematics show the way from evolutionary stability conditions to the preferences that evolution favors” (Alger and Weibull 2016, p. 57). What is problematic is that evolutionary scientists do not draw upon historical information to build their evolutionary models. Moreover, they do not examine empirical data that could confirm or invalidate the assumptions postulated by their models. In short, these models are ahistorical in the sense that relevant historical facts are completely absent from the evolutionary explanations they provide.

The point here is that researchers not solely concerned with purely theoretical problems internal to evolutionary theory have to provide evidence that support the model construction. By disregarding historical data in studying the workings of selection processes, theorists run the risk of building evolutionary models that leave out important variables, include irrelevant variables or fail to account for the actual functional relationship between independent and dependent variables. It is not only important that their evolutionary models are mathematically coherent, but they must also be empirically adequate in that the selection pressures advanced must have actually operated over a stretch of human evolutionary past¹³. If the assumptions of evolutionary models are largely inaccurate, the explanations they provide will not represent the historical processes they are supposed to examine.

A serious attempt to check whether there are historical studies that support evolutionary models of cooperation can be found in the work of Bowles and Gintis (Bowles and Gintis 2013: Chapter 6; see also Sterenly 2014). They examine in detail historical evidence to make clear that their models are consistent with the actual process that they purportedly study. For example, they present evidence of inter-group conflict and high rates of violent death in the late Pleistocene, to advance the view that human cooperation evolved by group selection. Moreover, they examine the size of nomadic hunter-gather bands and conclude that in contrast to what is typically assumed, they were not small but included around 90 members. Based on such evidence, they claim that simple direct reciprocity models do not adequately explain the evolution of human cooperation. It is unfortunate that advocates of the IEA have not carried out a similar analysis of the conditions allegedly obtained during preference evolution.

¹² For a discussion of different project in neuroeconomics, see Vromen (2011).

¹³ An enduring criticism of explanations that evolutionary researchers typically give for traits is that they are analogous to Kipling’s Just So Stories (Gould and Lewontin 1979). This criticism refers to accepting evolutionary explanations without sufficient support by empirical evidence.

To sum up, the IEA points out to a direction that evolutionary minded economists can take, should they be interested in applying evolutionary method. Most of the work in EGT and sociobiology ignores human psychology and decision-making, focusing on behaviors and their fitness effects. In contrast, economists often apply the IEA and assume that players are utility maximizers. Moreover, they treat human preferences as a type of biological trait that is acquired, modified and retained by outcompeting other traits in the population. Based on the discussion of the IEA, my argument is that instead of evolutionary models that privilege patterns of behavior, economists can draw upon or build evolutionary models where preferences and choice take center stage. Such evolutionary models can constrain the hypotheses space because they screen off those preference types that are not evolutionary successful.

4 Decision-making and utility maximization from an evolutionary perspective

In the previous section I argued that Burnham's method will be more reliable if economists apply evolutionary models that include some assumptions about human psychology and decision-making. This refinement of Burnham's method was based on the IEA, a theory devoted to studying the evolution of preferences. Moreover, I argued that a model's specifications have to be borne out from or have to be supported by empirical evidence about past environments. Hence, the following problem emerges. By design, these models include a decision-making process, typically based on rational utility maximization. One might wonder whether this assumption is plausible, given what we know about our ancestral past. Moreover, there is no shortage of alternative approaches that can be found in the literature (e.g. dual process theory, heuristics and biases, social learning biases, massive modularity hypothesis). An economist might want to consider substituting the utility maximization component of the IEA with a different model of decision-making.

In the IEA, economists' main interest is not whether people are rational utility maximizers, as most economic models typically assume, but to investigate the empirical and theoretical relevance of self-regarding and social preferences, given that people are rational. These models take for granted that players act rationally and make choices that maximize a utility function. Although utility maximization has been successfully applied in economics, it is often criticized on theoretical, empirical and conceptual grounds. However, the economists discussed in the previous section apparently hold that the assumption of utility maximizing players can be retained when taking insights from evolutionary theory into account. What evolutionary theory necessitates is modifying the arguments included in utility functions (e.g. Robson and Samuelson 2011; Bowles and Gintis 2013).

Not every economist agrees with this view. Ben-Ner and Putterman (1998, 2000) for example, explore what evolutionary theory has to offer to economics. Tellingly, they write that "in the long-run, the economics discipline as a whole will recognize that the old assumption of rational, strictly self-interested individuals is not only an inexact and special approximation, but also inconsistent with a scientific view

of human nature as the product of an evolutionary process” (2000, p. 97). Roughly speaking, more radical economists think that evolutionary theory implies that the standard economic theory should be discarded altogether (Burnham et al. 2016, p. 139). And they often put forward alternative decision-making accounts such as the “adaptive toolbox” approach with fast and frugal heuristics (e.g. Gigerenzer and Selten 2001) or modularity-based theories (e.g. Kurzban 2012).¹⁴

What does evolutionary theory has to say about human decision-making? From an evolutionary perspective, is there a plausible account of decision-making? Unfortunately, there are no clear answers to these questions. It is worrying that very little work has been done to develop evolutionary arguments, build evolutionary models and test claims about human decision-making. Economists of the IEA, for example, have not provided arguments for the implicit claim that utility maximizers instead, for example, of those that apply heuristics and biases must have had an evolutionary advantage. What they examine is whether some features of utility functions (i.e. preference types) would outperform in terms of reproductive success competing features of utility functions. Hence, it remains an open question whether (and to what extent) such decision-making capacities evolved in response to challenges of past environments.

When economists present intuitive evolutionary arguments in favor of utility maximization, the result remains inconclusive. Take for example Robson’s (2001) claim that the evolutionary function of expected utility is that it allowed optimal, rapid and flexible responses to novel environments¹⁵. To the extent that humans make utility maximizing choices, they do so because they are endowed with some sort of a general-purpose cognitive device. In constantly changing environments, such a device to make relatively accurate, fast and flexible decisions among alternative options might have been beneficial to survival and reproduction.

There are, however, alternative decision-making accounts informed by evolutionary theory. To give one example, a group of researchers around Gerd Gigerenzer claim that in our decision-making process we are using simple satisficing, domain-specific heuristic rules (Gigerenzer and Selten 2001). Although this group does not engage in evolutionary modelling, it purports that heuristics have evolutionary origins. They write that “evolution would seize upon informative environmental dependencies [...] and exploit them with specific heuristics if they would give a decision-making organism an adaptive edge” (Gigerenzer and Todd 1999, p. 19). What Gigerenzer and the ABC group appear to emphasize is that in a relatively stable environment with recurrent but distinct evolutionary problems, relying on simple heuristics would have been adaptive because of the benefits in decision-making speed and low energy expenditure.

¹⁴ Some economists believe that the neuroscientific investigation of the human brain will eventually lead to new mechanistic theories of decision-making (e.g. Burnham et al. 2016). While I do not doubt that such research can materialize in the future, in this section I only examine decision-making accounts currently applied by economists in their research.

¹⁵ For an in-depth examination of Robson’s evolutionary project, see Vromen (2003).

This quick overview by no means exhausts all the accounts expounded about human decision-making in economics. It suffices, however, to make the following point: It is not always evolutionary advantageous to make decisions according to standard economic theory. In a novel environment, having a general-purpose device might have been adaptive since it would allow for flexible choices that match new conditions better than behavior led by fixed rules. A constantly changing environment would render fixed heritable rules obsolete after a few generations. That, though, is not enough to exclude heuristics from being adaptive in the ancestral environment. Under more stable conditions, decisions according to rigid but computationally simple heuristic rules might have outperformed in terms of survival and reproduction decisions based on utility maximization.

The issue here that economists cannot simply indicate what kind of environments might have favored a particular decision-making theory. And it is not possible to develop their line of argument any further without an insight about the actual selection pressures obtained in our ancestral past. To my knowledge, economists that build models based on the IEA have not made a serious attempt to provide empirical evidence that is relevant to assess whether our past speaks in favor of utility maximization, or against it. Without a detailed understanding of the environment in which alternative decision-making capacities evolved, it is not possible to argue as to whether decisions based on utility maximization outperformed simple heuristics. It is not even possible to make a particularly strong case in favor of the weaker claim that decision-making based on standard economic theory was significantly common in past environments. The upshot is that it remains unclear to the extent that humans behaved rationally because they were endowed with some sort of a general-purpose utility maximizing device or their behaviors availed a multitude of special-purpose heuristic rules.

How can one determine which decision-making account is the most likely candidate to be included in models that provide natural selection explanations? The best means of approaching this issue is to supplement evolutionary arguments with evidence from fields such as primatology that focuses on the characteristics of extant non-human primates and ancestral hominids and anthropology that examines specific hunter-gatherers groups (Marlowe 2005). Moreover, phylogenetic analysis for hominid evolution could provide a chronology of major evolutionary changes in the social environment (Foley 1996). Although evidence showing what kind of decision-making must have been operative in our ancestral past is by no means easy to come by, evolutionary minded economists could put some research effort into gathering data that would serve as the basis for exploring the fitness benefits and costs of alternative decision-making capacities.

An important clarification to this argument needs to be made. From an evolutionary perspective, preferences are not at issue here. The problem is that economists of the IEA take for granted that decisions are made by a cognitive device that leads to utility maximization. The same point will still apply if the same economists routinely assumed that human decision-making is based on domain-specific modules, heuristics and biases, social learning biases, a dual process system or any other human decision-making process. The difference would be that, in this alternative case, they would have ignored to consider in their

evolutionary models competing decision-making theories, which nevertheless could have found support by empirical evidence from anthropology, archaeology or primatology.

To conclude this section, it turns out that the criticism raised against the utility maximization assumption of the IEA also applies to evolutionary arguments for or against any particular account of decision-making. Whether evolution favors domain specific modules, simple heuristics, dual process theory or utility maximization depends on the ARE. From an evolutionary perspective, it is not possible to assess alternative decision-making theories without an empirical analysis of specific features present in multiple past environments. In addition to abstract evolutionary theorizing, more empirical work is necessary to unravel the ARE against which many of our traits evolved.

5 Conclusion

Let us recap. For Burnham it is possible to formulate hypotheses and make predictions about human psychological traits, on the basis of what patterns of behavior would have likely been adaptive in the ARE. I argued that Burnham's method is evolutionary in the sense that evolutionary models are applied in determining the ARE for a particular behavior, and subsequently in searching for and framing hypotheses about psychological traits that can be explored in economic experiments. As we have seen, there will routinely be several scientifically plausible psychological hypotheses compatible with any given evolutionary model of behavior, and this raises the problem of grounding specific hypotheses about human psychology into such models.

This however should not lead economists to question the value of evolutionary theory when applied to the social sciences. Economists' interest in the details of psychological traits and decision-making has brought their evolutionary theorizing far-off to that of EGT and sociobiology. Burnham's method would be more reliable if instead of applying evolutionary models that target behavior, economists build evolutionary models of their own that target particular features of human psychology and cognition. In particular, economists of (or close to) the mainstream might opt for evolutionary models of the IEA that allow for an enrichment of received economic theory with novel insights about human preferences. Such models would allow economists to make progress in elucidating traits related to human psychology.

The drawback is that economists have to provide sufficient reasons why their evolutionary models commit to a particular account of decision-making. We saw that it is still an open question which decision-making theory is plausible from an evolutionary perspective. The proliferation of evolutionary models in economics did not result in significant contributions to the collection and processing of data that can illuminate the environment in which many of our traits evolved. It is important that an evolutionary study of human preferences and decision-making should also make use of empirical facts about the actual selection pressures at work during human evolutionary history.

Acknowledgements Earlier versions benefited from the comments of Jack Vromen, Attilia Ruzzene, Marina Uzunova, James Grayot, Osman Caglar Dede, Sine Bagatur, Philippe Verreault-Julien, Conrad Heilmann, Oskar van Eeden and participants of seminars at Erasmus Institute for Philosophy and Economics. The usual caveat applies.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akçay, E., Van Cleve, J., Feldman, M. W., & Roughgarden, J. (2009). A theory for the evolution of other-regard integrating proximate and ultimate perspectives. *Proceedings of the National Academy of Sciences United States of America*, 106(45), 19061–19066. <https://doi.org/10.1073/pnas.0904357106>.
- Alger, I., & Weibull, J. W. (2013). Homo Moralis: Preference evolution under incomplete information and assortative matching. *Econometrica*, 81, 2269–2302.
- Alger, I., & Weibull, J. W. (2016). Evolution and Kantian morality. *Games and Economic Behavior*, 98, 56–67. <https://doi.org/10.1016/j.geb.2016.05.006>.
- Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401), 464. <https://doi.org/10.2307/2234133>.
- Axelrod, R. (1984). *The evolution of cooperation*. New York: Basic Books.
- Ben-Ner, A., & Putterman, L. (1998). Values and institutions in economic analysis. In A. Ben-Ner & L. Putterman (Eds.), *Economics, values and organization* (pp. 3–69). New York: Cambridge University Press.
- Ben-Ner, A., & Putterman, L. (2000). On some implications of evolutionary psychology for the study of preferences and institutions. *Journal of Economic Behavior & Organization*, 43(1), 91–99. [https://doi.org/10.1016/s0167-2681\(00\)00110-4](https://doi.org/10.1016/s0167-2681(00)00110-4).
- Binmore, K. (1998). *Review of R Axelrod's "The complexity of cooperation: Agent-based models of competition and collaboration"*. New York: Springer.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 90(1), 166–193. <https://doi.org/10.1257/aer.90.1.166>.
- Bowlby, J. (1969). *Attachment and Loss. Volume I. Attachment*. New York: Basic Books.
- Bowlby, J. (1973). *Attachment and Loss. Volume II: Separation, anxiety, anger*. New York: Basic Books.
- Bowles, S., & Gintis, H. (2013). *A cooperative species: Human reciprocity and its evolution*. Princeton: Princeton University Press.
- Burnham, T. C. (2013). Toward a neo-Darwinian synthesis of neoclassical and behavioral economics. *Journal of Economic Behavior & Organization*, 90, S113–S127. <https://doi.org/10.1016/j.jebo.2012.12.015>.
- Burnham, T. C., & Hare, B. (2007). Engineering human cooperation. *Human Nature*, 18(2), 88–108. <https://doi.org/10.1007/s12110-007-9012-2>.
- Burnham, T. C., & Johnson, D. (2005). The biological and evolutionary logic of human cooperation. *Analyse und Kritik*, 27, 113–135.
- Burnham, T. C., Stephen, E. G., Lea, S. E., Bell, A., Gintis, H., Glimcher, P. W., et al. (2016). Evolutionary behavioral economics. In D. S. Wilson & A. Kirman (Eds.), *Complexity and evolution: toward a new synthesis for economics. Strungmann Forum Report 19*. Cambridge, MA: The MIT Press.

- Clavien, C., & Chapuisat, M. (2016). The evolution of utility functions and psychological altruism. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 56, 24–31. <https://doi.org/10.1016/j.shpsc.2015.10.008>.
- Collins, J., Baer, B., & Weber, E. J. (2016). Evolutionary biology in economics: A review. *Economic Record*, 92(297), 291–312. <https://doi.org/10.1111/1475-4932.12260>.
- D'Arms, J., Batterman, R., & Gorny, K. (1998). Game theoretic explanations and the evolution of justice. *Philosophy of Science*, 65(1), 76–102. <https://doi.org/10.1086/392627>.
- Dietrich, F., & List, C. (2016). Mentalism versus behaviourism in economics: A philosophy of science perspective. *Economics and Philosophy*, 32(2), 249–281. <https://doi.org/10.1017/s0266267115000462>.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817–868. <https://doi.org/10.1162/003355399556151>.
- Fehr, E., & Schneider, F. (2009). Eyes are on us, but nobody cares: Are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences*, 277(1686), 1315–1323. <https://doi.org/10.1098/rspb.2009.1900>.
- Foley, R. A. (1996). An evolutionary and chronological framework for human social behavior. *Proceedings of the British Academy*, 88, 95–117.
- Gigerenzer, G., & Selten, R. (2001). *Bounded rationality : The adaptive toolbox*. Cambridge: The MIT Press.
- Gigerenzer, G. & Todd, P. M. & the ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
- Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London B*, 205(1161), 581–598.
- Grafen, A. (2006). Optimization of inclusive fitness. *Journal of Theoretical Biology*, 238(3), 541–563. <https://doi.org/10.1016/j.jtbi.2005.06.009>.
- Grafen, A. (2014). The formal darwinism project in outline. *Biology & Philosophy*, 29(2), 155–174. <https://doi.org/10.1007/s10539-013-9414-y>.
- Griffiths, P. E. (1996). The Historical Turn in the Study of Adaptation. *The British Journal for the Philosophy of Science*, 47(4), 511–532. <https://doi.org/10.1093/bjps/47.4.511>.
- Güth, W., & Yaari, M. (1992). An evolutionary approach to explain reciprocal behavior in a simple strategic game. In U. Witt (Ed.), *Explaining process and change: approaches to evolutionary economics*. Ann Arbor: University of Michigan Press.
- Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: A critique of some recent interpretations of experimental games. *Theoretical Population Biology*, 69(3), 339–348. <https://doi.org/10.1016/j.tpb.2005.09.005>.
- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, 26, 245–256.
- Heller, Y., & Mohlin, E. (2019). Coevolution of deception and preferences: Darwin and Nash meet Machiavelli. *Games and Economic Behavior*, 113, 223–247. <https://doi.org/10.1016/j.geb.2018.09.011>.
- Hirshleifer, J. (1977). Economics from a biological viewpoint. *The Journal of Law and Economics*, 20(1), 1–52. <https://doi.org/10.1086/466891>.
- Irons, W. (1998). Adaptively relevant environments versus the environment of evolutionary adaptedness. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(6), 194–204. [https://doi.org/10.1002/\(sici\)1520-6505\(1998\)6:6<194:aid-evan2>3.0.co;2-b](https://doi.org/10.1002/(sici)1520-6505(1998)6:6<194:aid-evan2>3.0.co;2-b).
- Kanazawa, S. (2003). The Savanna principle. *Managerial and Decision Economics*, 25(1), 41–54. <https://doi.org/10.1002/mde.1130>.
- Kitcher, P. (1985). *Vaulting ambition: Sociobiology and the quest for human nature*. Cambridge: MIT Press.
- Kurzban, R. (2012). *Why everyone (else) is a hypocrite: Evolution and the modular mind*. Princeton: Princeton University Press.
- Lehmann, L., Alger, I., & Weibull, J. (2015). Does evolution lead to maximizing behavior? *Evolution*, 69(7), 1858–1873. <https://doi.org/10.1111/evo.12701>.
- Mailath, G. J. (1998). Do people play Nash equilibrium? Lessons from evolutionary game theory. *Journal of Economic Literature*, 36(3), 1347–1374.
- Marlowe, F. W. (2005). Hunter-gatherers and human evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 14(2), 54–67. <https://doi.org/10.1002/evan.20046>.

- Maynard, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.
- Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the Dictator Game: It's not how much you give, it's being seen to give something. *Evolution and Human Behavior*, *34*(1), 35–40. <https://doi.org/10.1016/j.evolhumbehav.2012.08.004>.
- Nowak, M. A., & Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature*, *437*(7063), 1291–1298. <https://doi.org/10.1038/nature04131>.
- Ok, E. A., & Vega-Redondo, F. (2001). On the evolution of individualistic preferences: An incomplete information scenario. *Journal of Economic Theory*, *97*(2), 231–254. <https://doi.org/10.1006/jeth.2000.2668>.
- Okasha, S. (2016). On the interpretation of decision theory. *Economics and Philosophy*, *32*(3), 409–433. <https://doi.org/10.1017/s0266267115000346>.
- Richerson, P. J., & Boyd, R. (2006). *Not by genes alone: How culture transformed human evolution*. Chicago: The University of Chicago Press.
- Robson, A. J. (2001). Why would nature give individuals utility functions? *Journal of Political Economy*, *109*(4), 900–914. <https://doi.org/10.1086/322083>.
- Robson, A. J., & Samuelson, L. (2011). The evolutionary foundations of preferences. In J. Benhabib, A. Bisin, & M. Jackson (Eds.), *The social economics handbook* (pp. 221–231). Amsterdam: North Holland.
- Sterelny, K. (2014). Cooperation, culture, and conflict. *The British Journal for the Philosophy of Science*, *67*(1), 1–31.
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35–57. <https://doi.org/10.1086/406755>.
- Vromen, J. (2003). Why the economic conception of human behaviour might lack a biological basis. *Theoria*, *18*(48), 297–323.
- Vromen, J. (2011). Neuroeconomics: Two camps gradually converging. What can economics gain from it? *International Review of Economics*, *58*, 267–285.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.