Studies in History and Philosophy of Science 42 (2011) 386–397

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Models as products of interdisciplinary exchange: Evidence from evolutionary game theory

Till Grüne-Yanoff

Helsinki Collegium for Advanced Studies, University of Helsinki, Fabianinkatu 24 (P.O. Box 4), 00014, Finland

ARTICLE INFO

Article history:
Available online 5 February 2011

Keywords:
Models
Evolutionary game theory
Interdisciplinarity
Unification
Unity of science
Theory import
Biology
Economics

ABSTRACT

The development of evolutionary game theory (EGT) is closely linked with two interdisciplinary exchanges: the import of game theory into biology, and the import of biologists' version of game theory into economics. This paper traces the history of these two import episodes. In each case the investigation covers what exactly was imported, what the motives for the import were, how the imported elements were put to use, and how they related to existing practices in the respective disciplines. Two conclusions emerged from this study. First, concepts derived from the unity of science discussion or the unification accounts of explanation are too strong and too narrow to be useful for analysing these interdisciplinary exchanges. Secondly, biology and economics—at least in relation to EGT—show significant differences in modelling practices: biologists seek to link EGT models to concrete empirical situations, whereas economists pursue conceptual exploration and possible explanation.

1. Introduction

Modelling is a practice of scientific theorising in which many disciplines engage. Philosophers of science have recently been focusing more attention on this practice and its products. One aspect that rarely has been discussed, however, is how much modelling practice differs across scientific disciplines. In this paper I explore these differences at the example of evolutionary game theory (EGT). EGT's roots in game theory may suggest that its modelling practices both in biology and economics are somewhat similar. Contrary to this expectation, I find the modelling practices to significantly differ in a number of dimensions.

Evolutionary game theory is the product of two interdisciplinary transfers. First, in the late 1960s and early 1970s biologists adopted game theory, which had been developed for the social sciences, and in particular economics. Then, roughly twenty years later, economists adopted what biologists had made of it. This makes EGT an interesting case for the purpose of this paper. Both biologists and economists claim that their EGT models have their roots in classical game theory, the implication being that the two are related through reference to a common 'toolbox' for constructing and solving models. Beyond these obvious relations, some authors have claimed that EGT in biology was largely an import from economics (Odenbaugh, 2009), or at least that it is very similar to economic modelling (Sugden, 2009). However, as I will show, the two transfers differ in terms, firstly, of what was transferred, and secondly of the use to which the thus established models were put. This sheds light on the differences in modelling practices in biology and economics.

The paper proceeds as follows. Section 2 sets the conceptual framework for the analysis. Section 3 sketches the history of EGT in biology, and then investigates in four subsections how it relates to classical game theory. Section 4 sketches the three main uses of EGT in economics, and investigates how these uses relate to the biologists' version. Section 5 compares the two episodes and concludes.

2. Conceptual framework

Interdisciplinarity is a relative new topic in the philosophy of science. Exchange between disciplines has been largely addressed

E-mail addresses: t.grune@helsinki.fi, till.grune@helsinki.fi

0039-3681/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsa.2010.12.004

by authors interested in the unity of science hypothesis or in unification accounts of explanation. These accounts have offered tools of analysis for interdisciplinary exchanges—for example, theory reduction or derivational unification—that are too strong and too narrow to be useful for analysing many exchanges observed in actual scientific practice. As the following case studies will show, they are too strong in that they postulate goals and processes that are often not part of interdisciplinary exchanges; and they are too narrow in that they neglect important dimensions along which such exchanges take place.

In this paper, I propose to analyse the exchanges between economics and biology along four dimensions. First, I investigate the relations of the researchers involved in the exchange to the respective disciplines. Two extreme cases can be distinguished. In cases of imperialism, researchers of one discipline apply their theories and methods to types of phenomena that have traditionally been treated by another discipline.[1] In the contrasting case of cooperation, researchers of both disciplines work together to exchange knowledge, by creating hybrid contact languages and local cultures of interaction.[2] Neither of these extreme cases is applicable here. Transfers were instigated by members of the importing discipline, and cooperation was notably absent. Rather, it seems that the distance of importing researchers to the discipline they imported has a significant function.

Secondly, I examine what formal aspects of theories or models are transferred. Standard philosophy of science tends to describe model transfers as reduction or derivational subsumption. From a Nagelian point of view, for example, a theory is reduced to another if its laws can be logically derived from the latter, under appropriate conditions of connectability (Nagel, 1961). Alternatively, descriptions of phenomena of one discipline may be derived from argument patterns of another discipline (Kitcher, 1989). In both cases, the reducing model or subsumptive argument pattern would be transferred into the discipline due to its reductive or subsumptive success. Yet neither of these accounts seems fully applicable here. Models are adjusted or restructured for the specific purposes of the importing discipline, in such a way that their identity is affected. A usefully weaker notion here is that of a computational template—a computationally tractable equation that describes constraints on relationships between variables (Humphreys, 2004, p. 60). Thus, a transfer may only import certain templates into a new discipline, leaving its previous theoretical context behind.

Thirdly, I explore what aspects of a model's ontology are transferred into another discipline. Scientists have certain beliefs about what their model presupposes about the nature of things. Commonly, these beliefs are shared within a discipline or at least within a paradigm of that discipline. When the formal aspects of models are transferred, so the question, what happens to its associated ontology? The extreme case here is ontological unification (Mäki, 2009), where phenomena of the importing discipline are redescribed as manifestations of the ontology of that discipline from which the model is imported. Yet the curious cases are those in which ontological unification is intentionally not sought, but a model transfer is nevertheless considered an epistemic success.

Fourthly, I study how methodology and style associated with models are transferred into another discipline. It seems that if a model is introduced into another discipline, including its ontology, with imperialist intention, then the methodology and style associated with that model is part of the package.[3] But if transfers are instigated by members of the importing discipline, if formal aspects of models are adjusted or restructured, and if models are interpreted under a different ontology, then the import of methodology or style does not seem a necessary part of the transfer anymore. Conversely, the interesting possibility arises that only a methodology or style is adopted, while the formal aspects and the ontology of the associated model are changed beyond recognition.

These four dimensions structure the analysis of the two case studies of sections 3 and 4. In section 5, I will use this framework to compare the two transfer episodes, and draw conclusions about the modelling practices of the respective disciplines.

3. From economics to biology

The explicit use of game theory in biology began with a little-known paper written by R. A. Fisher (1958). Fisher, of course, is well known for his work on the sex ratio and the related notion of population-dependent fitness (Fisher, 1930). That discussion was to become the conceptual basis for applying game theory in biology, but in his 1958 paper he proposed employing it for rather different purposes. He sought an explanation of polymorphism in a species, and he considered the game-theoretic notion of a *mixed strategy* an appropriate model. According to classical game theory, a single player employs a mixed strategy when randomising among a number of pure strategies. Fisher, however, suggested that whole species played against each other, and that for a species to adopt a mixed strategy meant that pure strategies would be played with certain levels of frequency in the population. Showing that a mixed strategy was indeed optimal in certain contexts would thus explain a polymorphism of corresponding phenotype frequencies in that population.

Fisher's proposal was little more than a paragraph in length, and offered neither formal treatment nor solution. Lewontin (1961), although apparently not aware of Fisher's proposal, developed a similar idea in formal detail. Modelling games of a whole population against nature, he suggested Maximin as the optimality criterion, and solved a theoretical example as a mixed Maximin strategy. He then offered to interpret that mixed optimal strategy as the frequency of different genotypes (with corresponding phenotypes) in the population.

Lewontin proposed the use of game theory as a general 'calculus of population evolution' (Lewontin, 1961, p. 384).[4] However, few population biologists used it in the way he suggested. Many of them may have considered competition between populations to be a group-selection-type argument, which began to lose favour among biologists in the 1960s. Instead, they took up game theory in a different area.

Fisher (1930) had offered a frequency-dependent fitness explanation for why sex ratios had to be equal. Although he did not couch his argument in game-theoretic language, his ideas were later found to be most clearly expressible through EGT. The first person to do this was W. D. Hamilton. He both clarified and extended Fisher's approach in arguing that models of local competition had a 'close similarity to certain kinds of situations considered

[1] In Lazear (2000), for instances, all the illustrative cases he discusses are economists who apply economic models and methods to topics like crime, law, the family, prejudice, tastes, politics, religion and war.

[2] Galison (1997) suggested the "trading zone" metaphor for the case in which physicists from different paradigms and engineers collaborated in the development of particle detectors and radar.

[3] Indeed, Lazear (2000, 103) largely justifies his version of economic imperialism with reference to the superior style of economics—by which he means high degrees of formalisation and idealization, testability and counterfactual analysis.

[4] Lewontin (personal communication) became sceptical of game theory soon after he published his paper because (i) real-world preferences are often cyclical, and (ii) he could not make his model fit with the phenomena of genetic separation and re-combination. Later he would become a vocal critic of Maynard Smith's ESS concept (Lewontin, 1982).

in the ''theory of games''' (Hamilton, 1967, 477). Using models of non-cooperative, non-zero sum games he modelled sex-ratio choice in parasites, either 'knowing' or 'expecting' double parasitism. He concluded that an extraordinary sex ratio could, under the envisioned circumstances, be an *unexploitable strategy*, and hence stable under natural selection.

Maynard Smith (1972) proposed game theory for yet another purpose, namely in adaptive explanations of conventional fighting. Animal contests are often fought in ways that limit the chance of injury or death, thus seemingly also reducing the chances of winning. Maynard Smith (1972) found that conventional fighting behaviour was stable against a number of alternative behaviours in a population of players. Although deriving his inspiration from classical game theory, he did not employ its solution concepts for these purposes, but developed his own notion, the *Evolutionary Stable Strategy* (ESS). He considered the link between his models and the classic models to be tentative, referring to a

> '*logical similarity* between the role of human reason in optimizing the outcome of a conflict between men, and the role of natural selection in optimizing the outcome of a fight between two animals' (Maynard Smith, 1972, p. 26, my emphasis)

Maynard Smith and Price (1973) extended the search for a stable strategy by using computer simulations; an analytic solution was also sketched, but was fully developed only in Maynard Smith (1974). Nevertheless, Maynard Smith and Price (1973) has become the *locus classicus* of EGT, and Maynard Smith's (1982) *Evolution and the Theory of Games* was to be the book that indicated the maturation of the field.

Commencing from ESS, Taylor and Jonker (1978) introduced a dynamic into evolutionary games. They assumed that the growth rate of those playing each strategy was proportional to its advantage. Although they stressed the many ways in which such a dynamic could be constructed, their assumption of exponential growth or decay was further developed by Zeeman (1980) and, once baptised the *replicator dynamics* by Schuster and Sigmund (1983), became the default dynamics for EGT.[5]

Taylor and Jonker thereby developed a new stability notion, which they showed was implied by ESS but did not imply ESS itself. They thus generalised the study of evolutionary games both by offering a stability notion weaker than the ESS, and by providing a model of out-of-equilibrium behaviour of a game. Nevertheless, they believed that ESS was the more important notion, for three reasons: ESS was easy to use, the dynamics required 'an assumption about the way in which fitness is translated into growth' (Taylor & Jonker, 1978, p. 154), and ESS seemed to be explanatorily sufficient for the games that had arisen in the biological literature. Given the later development of EGT, this claim exhibits an extraordinary amount of modesty.

In addition to the biologists who cast whole populations as players, and those who considered populations sets of players with fixed but inheritable strategies, a third strand used game models, and in particular the prisoner's dilemma (PD), in order to explain reciprocal altruism in animals. Here, again, Hamilton had a pioneering role. He suggested in a little-known paper (published in the proceedings of a 1969 conference) that 'situations of the PD type are probably quite common in nature' (Hamilton, 1971, 63). Better known is a paper by Robert Trivers, who acknowledged Hamilton's inspiration in showing how theoretical treatments of altruism 'can be reformulated in terms of game theory' (Trivers, 1971, p. 39), and applied the model to a wide range of behavioural phenomena. This strand of research was picked up ten years later by the political scientist Robert Axelrod, who collaborated with Hamilton on a paper on the evolution of cooperation. Axelrod and Hamilton (1981) combine the game-theoretic treatment of altruism with Maynard Smith's ESS notion, and offer a number of analytical approaches.

This brief overview of the early years of EGT is summarised graphically in Fig. 1.

The figure distinguishes between a (retrospectively diagnosed) theoretical relation (dotted line), a citation relation (dashed line) and a combination of the two (continuous line). The three strands of game-theory use in evolutionary biology are visible. The first one ended early in the period under discussion, having had no major conceptual influence on later development. The second strand developed into the backbone of EGT, even if the subject changed from the sex ratio to conventional fighting. The third strand lay idle for many years, but was finally synthesised with the second one. As will be shown in section 4, this synthesising paper and its authors played an important role in the transfer from biology to economics.

The following subsections discuss how the development sketched out above related to classical game theory. By classical game theory I mean the theory of rational human interaction, as developed by von Neumann and Morgenstern (1944), refined by John Nash in the early 1950s, and presented in textbooks such as Luce and Raiffa (1957).

3.1. Interdisciplinarity

Interdisciplinary transfer relies to a significant extent on the interdisciplinarity of the respective authors. Indicators of their positions include their contributions to the other discipline, their proficiency with the theory and their willingness to cite relevant works from the other discipline. With respect to early EGT, it is interesting to observe that authors' proficiency in classical game theory *decreased* as their work on EGT progressed.

When Fisher first introduced game theory into biology he spent most of his energy on establishing his record as a game theorist. We learn that he developed the notion of a randomised (or mixed) strategy as early as in 1934, and suggested Maximin as the criterion for selecting the optimal mixed strategy in the game *Le Her*. These results were thus published (in the *Mathematical Gazette)* before the *Theory of Games and Economic Behavior* and, as Fisher claims, without knowledge of von Neumann's preceding 1928 article in German. In suggesting mixed strategies to biologists, it seems, Fisher had simply found a new area for his *own* models.[6]

Although Lewontin did not claim involvement in the development of classical game theory, his knowledge of it is clear from his writings. A good part of Lewontin (1961) is devoted to a general discussion on the merits of different optimality criteria, and its application to the specific biological context is practically an annex. He clearly studied the cited von Neumann and Morgenstern (1944) and Luce and Raiffa (1957) in depth, and felt at home with this theory.

Hamilton's relation to game theory is more difficult to assess. He does not cite any game-theoretical texts in his 1967 work, nor does he discuss other game models or solution concepts. Nevertheless, he exhibits clear abilities to solve game models analytically. He cites Luce and Raiffa (1957) in his 1971 work, and a book by the political scientist Anatol Rapoport. He reports in his autobiography that he first encountered game theory—in the form of von Neumann and Morgenstern (1944)—as 'idle read-

[5] Taylor and Jonker's first assumption, that the growth rate of those playing each strategy is proportional to its advantage, constitutes the class of *monotone* dynamics, of which the replicator dynamic is the most prominent member. Despite efforts to broaden EGT studies to all members of this class, replicator dynamics continues to draw most attention.

[6] Fisher credits his assistant Luigi Cavalli-Sforza with the original idea of applying game theory to biology. Professor Cavalli-Sforza (personal communication) recalls discussing with Fisher the usefulness of game theory in the years 1948-50.

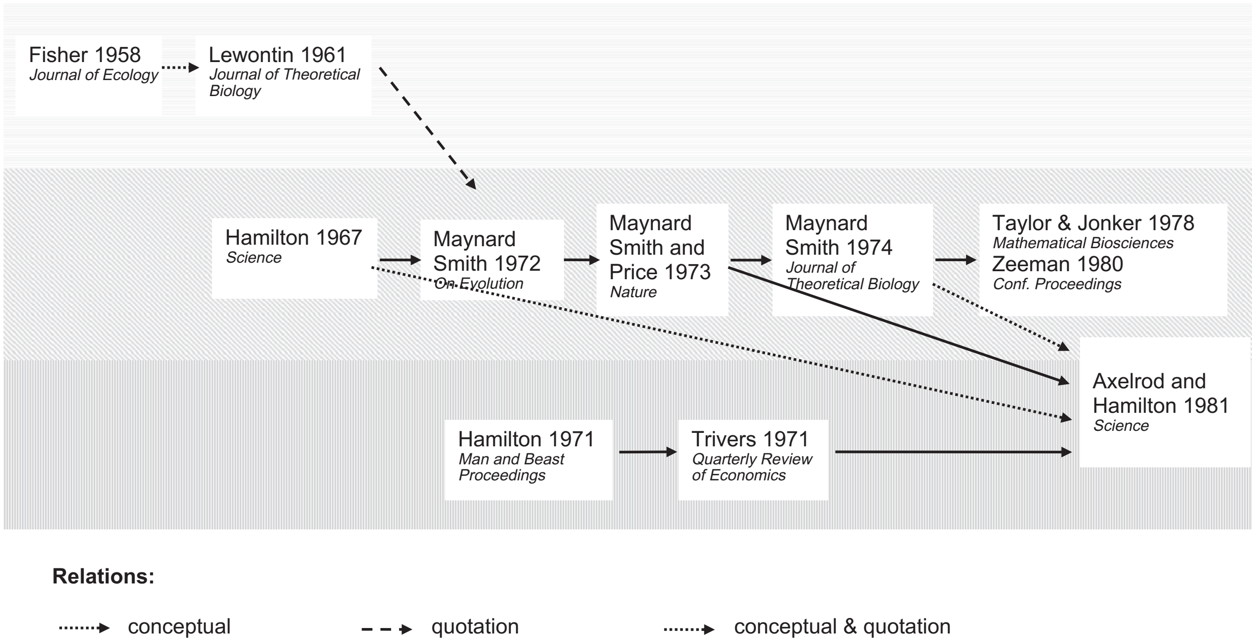

Fig. 1. The early development of EGT in biology.

ing at Cambridge' (during his undergraduate studies, 1957-60), but later found that it loomed 'with a stony insistence right in the centre of my chosen field' (Hamilton, 1996, p. 134).

According to Trivers (1971), reciprocal situations between animals is 'exactly analogous' to the Prisoner's Dilemma. He is referring here to the payoff matrix, with its identification of players, their strategies and evaluations of consequences, and makes no mention of classical solution concepts at all. Trivers refers to Luce and Raiffa (1957), but in particular to Rapoport and Chammah's (1965) *Prisoner's Dilemma*. This book, written by two social psychologists, is somewhat of an outlier in early game theory in that it focuses on one particular game, starts with empirical observations, and offers non-standard solution concepts.

As Maynard Smith acknowledges in his early papers, the concrete inspiration for his engagement with game theory stemmed from a manuscript of George Price, which he refereed for the journal *Nature*.[7] After reading this paper he went to the University of Chicago 'to learn about game theory' (Maynard Smith, 1976, p. 42).[8] Price never published the piece, and Maynard Smith thus suggested the collaboration that resulted in Maynard Smith and Price (1973) (Maynard Smith, 1972, pp. vi–viii; see also Frank, 1995). However, on account of the long refereeing process Maynard Smith (1972) was the first in print.

Maynard Smith's writing (especially in his 1972 paper) conveys a sense of indifference to game-theoretic detail. Many of his claims about classical game theory are inaccurate, or at least ambiguous. For example, he claimed (in 1972!) that game theory was most successful in zero-sum games, ignoring the recent literature and theoretical developments in non-cooperative, non-zero-sum games and sometimes also conflating them with cooperative games. He also claimed, incorrectly, that utility functions were full cardinal measures, and he failed to clarify whether the agents in his model chose their moves simultaneously (without knowing of the other's move) or sequentially (after gaining such knowledge). These distinctions are clearly presented in Luce and Raiffa (1957), for example, which Maynard Smith had in his library (Sigmund, 2005), but he seemed not to have paid close attention to them.

All these claims suggest negligence rather than serious error. He considered game theory a tool, and as he found many aspects of the classical theory not useful for his purposes he presumably saw no need to pursue them further. What mattered, in his view, were some basic concepts and techniques, 'mainly the concept of payoff-matrix' (Maynard Smith, 1972, p. 15). Thus, the most influential founder of evolutionary game theory was apparently relatively unaffected by the details of game theory's concepts and formal results.

3.2. Formal concepts

The increasingly wide separation between biologists' and classical game theory is also evident in the diverging employment of formal concepts, techniques and theorems. All authors make use of basic concepts such as players, strategies and payoff matrices,[9] but whereas earlier authors tended to adopt solution concepts and techniques from classical game theory, later writers increasingly substituted them with their own constructions.

Both Fisher and Lewontin exclusively focused their efforts on zero-sum games, which they justified with reference to the phenomenon in which they were interested, namely pure conflict between whole populations. They then argued that the notion of a mixed strategy and the Maximin solution concept provided a structure that enhanced understanding of these phenomena. Thus they directly took over parts of classical game theory: their *reinter-*

[7] His manuscript, entitled 'Antlers, intraspecific combat and altruism' was accepted by *Nature* on 7 February 1969, provided that it was shortened (Frank, 1995, p. 381). Price never resubmitted it, and it is now lost. It is not known what inspired him to employ game theory. It may have been through Hamilton, with whom he had been in contact since 1964 (Hamilton, 1996, pp. 172–176).

[8] It remains unclear whether or not he contacted the economics department at Chicago, which was dominated by neoclassical economists who had little time for 'alternative' approaches such as game theory.

[9] Yet there are subtle differences of interpretation in the way these concepts are used. Cf. Laubichler, Hagen, & Hammerstein (2005) for different interpretations of the strategy concept.

pretation of these elements did all work in adjusting them to biology.

Hamilton appeared to be more wary of directly importing classical solution concepts into biology. This shows in his terminology. He referred to 'unbeatable strategies' in zero sum games, although he later identified these unbeatable strategies with Maximin strategies. He further referred to 'unexploitable strategies' as the optimal strategies in (non-cooperative) non-zero sum games. He did not identify them with any classical concepts, but it is evident that he solved the payoff matrix for the unique pure Nash equilibrium. His way of representing the game, however, seemed designed to deflect any claim to the direct importation of game-theoretic ideas.[10] Although his solution concepts remained anchored in the perspective of the individual player, Hamilton also offered a brief and informal hint at a population perspective. Discussing an arithmetical illustration of a population in which players only play two strategies, he argued that

> 'the weights [in the average fitness result] depend on the frequencies of the different types of pairs [in the population]' (Hamilton, 1967, p. 486).

Although this did not amount to a full population perspective, it foreshadowed the main idea of evolutionary stability. Trivers gave further evidence that Hamilton seriously considered this idea. Avoiding any reference to classical solutions of the Prisoner's Dilemma, he reported a personal communication with Hamilton, who offered a population perspective on the PD: players are endowed with one strategy, and their payoffs are interpreted as individual fitness. In a population with many altruists, Trivers argued, it is plausible that an altruist fares better than an egoist:

> '... the nonaltruistic type, when rare, cannot start to spread. But there is also a barrier to the spread of altruism when altruists are rare.' (Trivers, 1971, p. 39)

This, again, did not amount to a formal solution of the game, remaining informal and vague in terms of evolutionary stability. Yet it is clear that authors such as Trivers preferred chipping away at this new idea (and ending up without a formal solution) to taking over formal solutions from classical game theory.

Maynard Smith (1972) increased this distance from classical game theory further. As already discussed, he limited the transfer of game-theoretic concepts to the very basics, especially the payoff function. He then developed the population perspective head-on, constructing the notion of an Evolutionary Stable Strategy (ESS):

> 'S is ESS if, in a population in which most individuals adopt it, there is no alternative strategy which will pay better' (Maynard Smith, 1972, p. 21).

This is not the full definition developed in Maynard Smith and Price (1973) and Maynard Smith (1974), but it constitutes the basis of the latter notion. What is noteworthy is that here, as well as later, the solution concept is developed in explicit contrast to the Maximin notion, and without mention of any equilibrium concept applicable to non-zero sum games. Thus, Selten's later assertion that

> 'the basis for this extremely successful transfer [of game theory from economics to biology] is the concept of Nash equilibrium' (Kuhn et al., 1996, p. 165)

is not literally correct. There is no evidence that Hamilton, Price or Maynard Smith ever read Nash's papers, or learned about them in the game-theoretic literature of their day. Instead, they claimed to be developing new concepts for types of games about which classical game theorists had little of interest to say.

This is even more pronounced in Maynard Smith and Price (1973), who proposed a computer simulation as a solution for their game model.[11] Classical game theorists never employed such techniques, but rather insisted on analytic solutions. Maynard Smith, however, exhibited a certain distance to analytic solution techniques. For example, the formal and analytical treatment of equilibrium existence in Maynard Smith (1974) was left to a mathematician, John Haigh, and published as an appendix to the paper.

Hence, although classical game theory clearly constituted an *inspiration* for EGT, it seems that biologists developed most of its more sophisticated concepts themselves, in fact replacing existing classical solution concepts. The diagnosis of such a growing bifurcation between classical and biological game-theoretic concepts puts some doubt on the notion of a *transfer* or *import* of models or theory.[12]

3.3. Ontologies

Classical game theory is also characterised by a somewhat unified ontology, i.e. scientists' beliefs about what the formal apparatus presupposes about the nature of things in the world. Classical game theory, for example, is commonly seen as presupposing that human agents have certain epistemic and reasoning capacities, that they have well-formed preferences, and that these preferences are sufficiently stable to form meaningful plans.[13]

When biologists became interested in game theory they were attracted by the formal elements of the classical theory rather than its ontology. It is evident in the first articles to be published that they were attempting to reinterpret classical game theory, and later to adjust the formal elements to match useful biological interpretations. As Fisher notes, although learning may occur in higher animals, 'the most important aspect' of game theory for biological applications is that 'each species is ... *evolving* weapons, sense organs, and innate drives ... as to progressively improve its chances of success in these encounters (Fisher, 1958, p. 292, emphasis added). He also offers a population-level interpretation of mixed strategies, moving further away from individual agents and their chosen plans of action.

Lewontin notes more explicitly that various aspects of classical game theory—agents being purposive, optimality depending on preferences, selection through choice processes—fit badly with the intended biological application. Instead, he offers the probability of the survival of a population as an alternative to the expected utility measure, thus eliminating the reference to preferences; he distinguishes between his polymorphic interpretation

[10] Hamilton represents a sequential-move game as a payoff matrix. He writes the matrix with the origin in the lower left corner. He then indicates the sequentiality of the game by drawing best-response curves in the matrix.

[11] Maynard Smith thus pioneered not only a new solution concept, but also a new solution technique, viz. agent-based modelling (Sigmund, 2005, p. 9). It is worth remarking that today biologists often make use of simulation techniques, while economists rarely use it. This constitutes an enduring methodological divide between the two disciplines.

[12] Contrast this with the way game theory was transferred into sociology: sociologists did not develop their own GT concepts, but rather acquired them directly from economists (Swedberg, 2000). A comparable case of direct transfer is found, perhaps ironically, in economists' adoption of EGT (see section 4).

[13] It is, of course, a difficult task to ascertain whether the theory indeed commits a theorist to these claims. One could argue that a theorist only needs to *assume* the truth of these claims, without *accepting* them. However, the history of game theory shows that most economists were greatly discomfited by the strength of these claims, and sought to soften them (cf. sections 4.1 and 4.2 of this paper). This suggests that they indeed saw themselves committed to *accepting* these claims.

of a mixed strategy and the classical version, thus reducing the purposive aspect; and he replaces the idea that players choose optimal plans with the claim that 'surviving groups . . . simply by chance, acquired optimal strategies' (Lewontin, 1961, pp. 401–2).

Hamilton offers a two-pronged strategy. He models the specific double-parasitism situation with standard assumptions, presupposing that a second parasite would be capable of knowing the number of eggs a first parasite had laid, as well as their sex ratio. He admits that these assumptions may be unwarranted, but claims that their incorrectness does 'not lessen the gamelike character of the situation' (Hamilton, 1967, p. 487). He goes on to give the outline of an alternative interpretation, in which an unexploitable sex ratio *gene* maximises individual fitness, and in which convergence to the optimal point is interpreted as the evolution of a species through natural selection. For example, he suggests that equilibrium could be found without foresight or sophisticated reasoning abilities: 'through trial and error, two naïve players would quickly learn that the constant playing of 1/4 was the optimum-yielding strategy' (Hamilton, 1967, p. 486).

This shows that biologists often rejected classical game-theoretic ontology. Beyond this negative conclusion, however, it must be recognised that they often used EGT to express their previously conceived ideas in more efficient terminology. A good example of this is Fisher's theory of the equal sex ratio. Fisher (1930) argued that when measuring individual fitness in terms of the expected number of *grandchildren*, individual fitness depended on the distribution of males and females in the population: if the number of males in a population is greater than the number of females, parents who produce more female than male offspring will have more grandchildren and hence a higher fitness level; if on the other hand the number of females is larger, parents who produce more males than females have higher individual fitness. Hence, he concluded, the sex ratio is in equilibrium when the total effort spent producing the two sexes is equal. Despite his apparent proficiency (cf. section 3.1) however, he did not employ game theory to express this notion. Later authors, such as Hamilton, found that expressing Fisher's theory in terms of game theory made it much clearer, and also showed its limitations (Hamilton, 1996, pp. 131–132). One could therefore conclude that when biologists adopted game theory, they often did so *merely* because they found it a formal framework within which they could express their own ideas more clearly.[14]

Maynard Smith (1972) and Maynard Smith and Price (1973) interpreted their own concepts in evolutionary terms, rather than reinterpreting classical game theory.[15] Strategies were now fixed in players and inheritable, and the advantages of interpreting payoffs as fitness rather than utilities were advertised. The population perspective was already built into the ESS concept, and did not require an added-on interpretation. Biologists thus constructed a formal framework that fitted their own ontology, but continued using the same terminology that was originally associated with the social-science interpretation of classical game theory.[16]

3.4. Phenomena-directedness

Classical game theory often is intended as a descriptive theory. Yet there were few attempts in theoretical papers to link game-theoretic models with concrete empirical accounts of the phenomena they were supposed to describe. Indeed, Luce and Raiffa felt compelled to remark on the relative lack of application of the theory to particular empirical problems (Luce & Raiffa, 1957, pp. 10–11).

True, from the 1980s onwards game theory has been applied to a number of areas, including industrial organisation, auctions or bargaining. But a strong division of labour remains between 'pure' and 'applied' game theory. Pure game theory only applies to intuitive but imaginary stories (Grüne-Yanoff & Schweinzer, 2008; Morgan, 2007). Typical examples of such stylised stories include the narrative of the two criminals in the Prisoners' Dilemma and the story of the hunters in the stag-hunt game. Until the 1980s, most economists specialising in game theory have been interested in such pure theory questions (Geçkil & Anderson, 2010). Applied game theory, in contrast, uses simple tools from pure game theory for the analysis of actual situations. Most of the effort in such applications is put into correctly describing the features necessary to model the situation as a game. Rarely do results from applications feed back into the pure theory.

There is no such division of labour and hence no distance from concrete empirical situations in biologists' use of game theory. All the papers concern one of the following phenomena: population polymorphism, sex ratios, conventional conflicts and reciprocal altruism.[17] They are specified with reference to studies of concrete empirical systems rather than through stylised or imaginary narratives. Fisher, for example, referred to polymorphic butterflies that display mimicry. Hamilton drew on a list of 22 species involved in double parasitism, many of which he apparently observed himself (Hamilton, 1996, p. 140). Maynard Smith (1972) discussed fighting behaviour among bighorn rams, fence lizards, native hens and sword-tail fish, and Maynard Smith and Price covered fights among some species of male poisonous snakes, mule deer, Arabian onyx and elephants in musth. Trivers extensively discussed cleaning symbioses in fish, warning calls in birds and reciprocation in humans. Axelrod and Hamilton referred to studies of the fig wasp, to the shift from symbiosis to parasitism with lessening recognition abilities, and to the relation between prosopagnosia ('face-blindness') and non-cooperation in humans.[18]

It is significant that these empirical cases not only served to illustrate the game models, but were also meant to test their empirical adequacy. Maynard Smith (1972), for example, considers the case of the Siamese fighting fish, which he admits is a counterexample to conventional fighting. Trivers offers a list of empirically testable criteria that validate his reciprocity model, including:

> '. . . that hosts suffer from ectoparasites; that finding a new cleaner may be difficult or dangerous; that if one does not eat one's cleaner, the same cleaner can be found and used a second time; that cleaners live long enough to be used repeatedly by the

[14] Cf. Dawkins (1989, p. 287): 'You don't actually have to use ESS language [to express ideas about the evolution of behaviour], provided you think clearly enough. But it is a great aid to thinking clearly'.

[15] At least that is the authors' self-understanding. Mary Morgan (personal communication) recalls Maynard Smith insisting that although he learned about game theory from a classical textbook (Luce & Raiffa, 1957), he 'only took the mathematics'.

[16] This continued use of social-science terminology led at times to serious confusion, which in turn contributed to the scepticism that the application of game theory to biology sometimes met. On this entanglement of EGT in the so-called sociobiology debate, see Segerstråle, 2000, chap. 3.

[17] There is also a difference in the *type* of phenomena on which the two theories focus. Classical game theory describes socially and temporally isolated encounters, whereas evolutionary game theory describes macro-social behavioural regularities (Gintis, 2000, p. 300).

[18] Admittedly, measuring concrete fitness values of strategy profiles proved difficult. But a lot of effort was put into this research area, and significant progress ensued. A good example of this is Dawkins' (1976, p. 81) initial caveat—'we know too little at present to assign realistic numbers'—and his later confident correction (Dawkins, 1989, pp. 283, 284) in the light of new results on gold digger wasps.

same host; and if possible, that individual hosts do, in fact, reuse the same cleaner.' (Trivers, 1971, p. 41)

The tradition of using empirical cases, as exhibited in these early papers, was a hallmark of game theory in biology. Today there are hundreds of papers on animal behaviour that make use of game theory. Across all taxa, attempts have been made to identify the chances of continued interaction, the strategies played, and the payoffs involved. Further, the early models developed by Trivers, Axelrod and Hamilton have been subjected to numerous adjustments in the light of empirical observations. They have been adapted for studying the effects on cooperation of, for example, group size, population structure, player mobility, encounter probabilities, mistakes, random payoff functions and many other factors (for a more systematic overview, see Dugatkin, 1997, pp. 25–28).

As Dawkins remarked, 'simple models can be elaborated and gradually made more complex. If all goes well, as they get more complex they come to resemble the world more' (Dawkins, 1976, p. 79). As I have shown in this section, this quest for de-idealisation was not only rhetorical—biologists were strongly committed to linking their idealised models back to concrete empirical situations.

3.5. Summary

Reviewing the economics-to-biology import episodes yielded the following results. First, knowledge of, involvement with and reference to classical game theory seemed to be inversely proportional to the influence the respective authors had on the course of EGT.

Secondly, biologists constructed the more sophisticated formal EGT concepts themselves. One could speak of the import of formal concepts only with respect to very basic notions such as strategies or pay-off matrices, and it may be more appropriate to refer to formal inspirations rather than imports or transfers in these contexts. A bon mot attributed to Kissinger is as follows: game theory's major contribution was the payoff matrix, in that it made people realize that their adversaries were looking at the conflict in a similar way as they were (Aumann, 1985, p. 188). Just like the cold warriors, biologists seemed to take this simple device from classical game theory - and to have introduced it into the context of natural selection via their *own* formal models.

Thirdly, biologists were never interested in importing the ontology of classical game theory into their discipline, despite the social-scientific slant of the terminology. In many cases the EGT framework was used to express existing ideas in a more efficient way.

Fourthly, the purpose of the majority of EGT models in biology was to represent specific, empirically documented phenomena. This phenomena-directedness was supported in terms of applying the models to concrete empirical situations, and making modifications accordingly.

4. From biology to economics

It was a political scientist who introduced evolutionary game theory into economics. Robert Axelrod (1980), writing in the *Journal of Conflict Resolution*, seemed to be the first person to apply ESS to a game involving humans in a social-science journal.[19] He staged a computer tournament between submitted strategies for the iterated Prisoners' Dilemma. In order to ascertain the robustness of the winning TIT-FOR-TAT strategy, he made the presence of a strategy in the tournament dependent on its success in past rounds, yielding an evolutionary convergence on TIT-FOR-TAT. In order to show its stability, he then analytically investigated whether other strategies could invade a population in which agents played TIT-FOR-TAT with high probability. Axelrod interpreted this investigation in an evolutionary way, and suggested that TIT-FOR-TAT could be an ESS.

Axelrod cited Trivers (1971), Dawkins (1976) and a popular science article of Maynard Smith from 1978 in this paper. Curiously, his application of EGT to social contexts preceded his contribution to biology: only after writing on the evolution of human strategies did he initiate the collaboration with Hamilton that led to Axelrod and Hamilton (1981).[20]

Another interesting feature of Axelrod's 1980 article is his explicit adoption of a biological ontology. He discusses how to simulate 'future generations of a tournament', and then switches from an ontological neutral perspective of replication of strategies to a biological perspective of reproduction of strategy bearers:

> 'we simply have to interpret the average payoff received by an individual as proportional to that individual's expected number of [truly-bred] offspring (Axelrod, 1980, p. 398)

For a number of reasons, not least his reliance on computer simulation, Axelrod received a rather critical reception in economics, or was simply ignored. Nevertheless, his employment of EGT was widely noted, and references to his papers are to be found in many pioneering EGT publications in economics (e.g., Bomze, 1986; Sugden, 1986).

From then on, economists and other social scientists adopted BGT for a number of purposes. It is noteworthy that all the 'early adopters' had perceived some deficiency in classical game theory that led them to follow the evolutionary perspective. I have therefore structured this section according to the various deficiencies that the respective authors perceived.

4.1. The recovery project

The first adoption motive is inherently conservative, the aim being to *recover* or maintain the standard theoretical framework of classical game theory by giving it a new, evolutionary interpretation. It had long been felt that game theory imposed too strong rationality requirements on individual players. The implausibility of these assumptions became a pressing issue in the late 1980s when game-theoretic rationality was explicated through its epistemic prerequisites. It became clear just how strong and implausible the epistemic requirements for Nash equilibrium really were. EGT seemed to offer a way out of this problem:

> 'Maynard Smith's book *Evolution and the Theory of Games* directed game theorists' attention away from their increasingly elaborate definitions of rationality. After all, insects can hardly be said to think at all, and so rationality cannot be so crucial if game theory somehow manages to predict their behavior under appropriate conditions' (Binmore, foreword in Weibull, 1995, p. x).

From this point of view, EGT should therefore prop up the solution concepts of classical game theory, rather than replace them. The basis for this project lay in various formal results published in the 1980s. In particular, in an article in the *International Journal of Game Theory* Bomze (1986) proved the identity or implication relation between classical equilibrium notions and biologists' stability notions. He showed specifically that (i) if *S* is an ESS, then (*S*,*S*) is a Nash

[19] I should note that some social scientists, like Selten (1980), had earlier contributed to EGT in biological journals. Thus economists had noted EGT earlier than Axelrod.
[20] The British biologist Richard Dawkins established their contact (Dawkins, 1989, p. 214).

equilibrium; (ii) (*S*,*S*) is a strict Nash equilibrium if *S* is a regular pure ESS (i.e. *S* is a pure strategy and a strict best reply to itself); (iii) if *S* is an asymptotically stable dynamic equilibrium, then (*S*,*S*) is a Nash equilibrium. It was on this article that significant works on EGT in economics, including van Damme (1987), Weibull (1995) and Samuelson (1997), were based.

The formal implication results made it possible to perceive the equilibria of individual strategic encounters as based on evolved behavioural regularities among large populations. As Binmore, for example, claimed, 'as regards positive economics, it would seem to be evolutive processes that matter most' (Binmore, 1988, p. 185, fn. 16). Later he more explicitly rejected the notion of classical rationality: if it was valid to model people as maximisers, this could only be because 'evolutionary forces, biological, social and economic, [are] responsible for getting things maximised' (Binmore, 1994, p. 11). Crucially, the formal implication results allowed the continued use of standard game-theoretic notions. Although there is a more fundamental story 'behind' human behaviour, it is perfectly justifiable to treat it 'as if' it was indeed driven by cognitive maximisation efforts. In a curious twist, the assertion of evolutionary ideas thus led to the continuation of standard game theory—in particular to the continued use of the Nash equilibrium, minus all the bothersome epistemic conditions.

> 'Even if strategically interacting agents do not meet these epistemic conditions, their long-run aggregate behavior will nevertheless conform with them because of the workings of biological or social selection processes.' (Weibull, 1994b, p. 868)

Just as Friedman had used an evolutionary idea in his famous methodology essay to defend standard microeconomic assumptions, so evolutionary ideas were now used in game theory to prop up the classical theory (for a discussion on Friedman's evolutionary programme see Nelson & Winter, 2002, pp. 25–28; on the relation between EGT and Friedman's argument, see Vromen, 2009). The fine distinction now is that formal identity proofs for results from EGT and classical theory seem to offer a much more precise foundation.

Thus, on the formal level economists did *not* import biologists' EGT for the purpose of recovery. They did not replace the Nash equilibrium with ESS or with stability results from replicator dynamics. Nor did they—as the biologists had some 20 years earlier—develop new EGT-inspired solution concepts. They rather concluded from the formal implication results that they were licensed to stick with their formal framework, but needed to replace its difficult epistemic interpretation with an evolutionary one.[21]

Indeed, economists interested in EGT professed a much closer affinity with biology than biologists interested in game theory had. For example, they freely quoted from biologists' works, in particular Maynard Smith and Price (1973) and Maynard Smith (1982). They also often adopted the terminology, describing social events as driven by evolution or selection processes (e.g., Samuelson, 1997, p. 51: Mailath, 1998, p. 1355). Yet this professed affinity was based on an analogy construction: the *claim* that biological and social processes shared the same core evolutionary features.

Most economists took this claim for granted without investigating the social processes further. When confronted with the question of how strategies replicated they would answer in the same way as Axelrod had: through the reproduction of strategy bearers. Only a few attempted to investigate alternative transmission mechanisms,[22] and if they did they focused on the question of whether such an alternative mechanism would give rise to the same stability properties as the biological models. Little effort was expended on investigating the true mechanisms of cultural evolution, and concerns that learning could not easily be subsumed under natural selection were ignored (cf. Grüne-Yanoff, 2011; Vromen, 2006). Instead, in their attempt to make the formal implication results useful for economics, economists more or less reluctantly accepted important parts of the biologists' ontology.

Symptomatic of such analogy constructions is the appearance of the meme concept in EGT, which was originally proposed by Dawkins in order to leave a place for cultural evolution without reducing it to a genetic level. As an operative concept it remained rather vague (as Dawkins himself notes, cf. Dawkins, 1976, pp. 209–210), in particular with respect to its identity conditions and range, but this did not deter economists from employing the meme in order to provide a social ontology that fitted the formal EGT structure. Binmore (1994) is an example. He proposed that memes determine human behaviour in ways that yield their optimal replication, and that optimal meme replication necessitates maximising behaviour among individuals. Maximising behaviour is thus an epiphenomenon of meme evolution:

> 'People who are inconsistent [in their preferences] will necessarily be sometimes wrong and hence will be at a disadvantage compared to those who are always right. And evolution is not kind to memes that inhibit their own replication.' (Binmore, 1994, p. 27)

The meme ontology is supposed to match the formal structure of EGT on the one hand, yet on the other it is supposed to fit in with the social sciences. Attempting this double-match, the recovery project largely avoids reference to specific phenomena or empirical data. Authors refer to 'social' or 'cultural' selection, without spelling out what the underlying replicators, replicator bearers, or replication mechanisms are. Instead, they seek maximal generality through proof of formal equivalence or implication. This results in the selective import of a truncated EGT ontology.[23]

4.2. Equilibrium refinements and selection

The second adoption motive was the desire to resolve some of the issues arising with *equilibrium refinements* and *equilibrium selection*. In many games the identification of Nash equilibria did not eliminate all 'unreasonable' equilibria, or more generally did not yield a unique Nash Equilibrium. Therefore, the late 1970s and 1980s saw the proliferation of additional criteria that were supposed to select only 'credible' or 'reasonable' equilibria. The problem here was an embarrassment of riches. More and more competing refinements were developed, which imposed massive demands on the agents' ability to reason (and enormous faith that other agents would follow similar reasoning paths). These were difficult to work with, and the predictions were not always consistent with intuition, common sense or experimental evidence. What

[21] It should be noted that Nash suggested in his 1950 dissertation that his notion of formal equilibrium could be interpreted as a population equilibrium. This so-called Mass-Action Interpretation shares surprising similarities with the EGT interpretation on which economists were working in the 1980s and 90s. However, this part of Nash's thesis was only published in Nash (1996, pp. 32–33), and the earliest published references are Leonard (1994) and Weibull (1994a). Jörgen Weibull (personal communication) recalls that he cited the relevant passage from Nash at a meeting of the International Institute for Applied System Analysis (IIASA) in Schloss Laxenburg in 1993. Although many among the audience of about a hundred were game theorists—including Josef Hofbauer, Karl Sigmund and Peyton Young—no one knew this was Nash's work, and many seemed greatly surprised. So it seems that the pioneers of EGT in economics did not know about the Mass-Action Interpretation until around 1993 or '94. The reason for this delayed publication seems to have been an editorial decision of *Annals of Mathematics*, although Nash now cannot remember why this happened (cf. Kuhn et al., 1996, p. 181).

[22] For example, Sugden (1986), Young (1993), Björnerstedt & Weibull (1996) and Schlag (1998).

[23] This re-interpretive use of EGT may be the main reason for the *bloodlessness* of the evolutionary turn that some authors have diagnosed (e.g., Sugden, 2001): despite years of practising game theory with epistemic foundations, no one protested against the replacement of these foundations with evolutionary ones.

was even more troubling, there was no proper basis on which to interpret these refinements or to choose from among them.

Against this growing problem, the notion of evolutionary stability seemed like a potential solution. In particular, economists saw value in Axelrod (1980): in an early review Milgrom argued that Axelrod's

> '... extension of the idea of stability beyond Nash equilibrium provides a counterpoint to traditional game theory and a challenge to game theorists to expand their views.' (Milgrom, 1984, p. 308).

In a similar vein, despite his otherwise strong reservations about Axelrod's work, Binmore looked back to the 1980s and admitted:

> 'I believe that [Axelrod] did make an important contribution to game theory. [...] He did us the service of focusing our attention on the importance of evolution in *selecting* an equilibrium from the infinitude of possibilities whose existence is demonstrated by the folk theorem.' (Binmore, 1998, emphasis added)

Most equilibrium refinements required equilibria to survive perturbances created by opponents playing out-of-equilibrium strategies. Both replicator dynamics and ESS model the stability of strategy populations in the face of out-of-equilibrium perturbations. It thus appeared to be closely analogous to the refinement project—but, crucially, it worked without the problematic rationality and epistemic assumptions. Furthermore, biological EGT seemed very unified, with few but powerful solution concepts. Thus, it was hoped that its import would have both a simplifying and a unifying effect.

Again, formal identity proofs played a crucial role in this project. Inspired by the formal proofs in Bomze (1986) and others, economists became interested in notions of evolutionary stability in the late 1980s. The implications always go from evolutionary stability to refinement—evidently the stability conditions are more restrictive than the refinements. Thus, it was also hoped that EGT would have selective virtues beyond those of the standard refinements. van Damme had written his PhD thesis on equilibrium refinement and selection in 1983, and included Bomze's results in the first published version of it in 1987.

As in the recovery project, authors involved with refinements and selection made direct reference to biological EGT. However, they sought to directly import certain formal concepts into classical theory, first in order to replace some of the existing refinements (which are based on dubious epistemic and rationality assumptions), secondly to unify the refinement approach (by providing a common, evolutionary basis), and thirdly to move it forward (as the EGT concepts tend to be stricter than most existing refinements). Thus the import of formal concepts here took precedence over the import of an existing ontology, in particular in the first and third purposes. The ontological dimension largely seems to have been an afterthought. Authors sometimes point out that they are modelling social or cultural rather than natural selection, but these claims are not concretised, nor investigated empirically.

Only when it was realised that EGT would not meet most of the expectations initially put upon it did the writers issue warnings about the need to develop a fully-fledged theory of social evolution. As Samuelson wrote in his book *Evolutionary Games and Equilibrium Selection*: 'although we have much to learn from biological evolutionary models, we must do more than simply borrow techniques from biologists' (Samuelson, 1997, p. 37). By that time it had turned out that both evolutionary stable states and ESS too often failed to exist in economically interesting applications, and more suitable alternative notions of evolutionary stability needed to be developed.[24] This required, he argued, investigation of the underlying dynamic processes, and only then could the relation between evolutionary arguments and equilibrium selection be clarified.

It is informative to compare this interim result of EGT in economics with EGT in biology. It took economists ten years, from 1987 to 1997, and many papers on ESS and asymptotic stability, to strike a cautionary note like Samuelson's (and then in a book rather than a journal article). In contrast, as shown in section 3, biologists immediately saw the need to develop their own concepts, and to keep their distance both formally and ontologically from classical game theory.

4.3. Evolutionary explanations

Experiments in psychology and economics conducted the late 1970s and early '80s identified behavioural and economic phenomena for which no appropriate economic explanation was available. These experiments led to an increased demand for explanations of social institutions, in particular of fairness norms. There were some early attempts to provide evolutionary explanations, yet without reference to biology (Schotter, 1981). The first to address this issue from the perspective of EGT was Robert Sugden in his 1986 book. He saw himself in the tradition of Schotter, but unlike him drew explicitly on biology:

> 'In many ways I learned most from the biological literature, since it provided a theoretical framework that neo-classical economics, with its emphasis on rational maximising in a static environment, is ill-equipped to supply' (Sugden, 1986, p. vii)

Using the tools of evolutionary game theory, in particular ESS, he showed how self-enforcing conventions of property and reciprocity can evolve spontaneously out of the interactions of self-interested individuals. He went on to argue that such conventions tend to become norms even if they arbitrarily favour some people relative to others, and even if they do not maximise social welfare. Sugden was thus the first author to use EGT in order to explain the evolution of human social institutions. In his wake, various authors proposed evolutionary explanations based on EGT (although they did not necessarily follow Sugden's approach): from the evolution of conventions (Young, 1993) and norms (Binmore & Samuelson, 1994) to the 'indirect evolutionary approach', the aim of which was to account for the development of fairness preferences in the Ultimatum Game (Güth & Yaari, 1992).

Sugden's account is noteworthy because, unlike many that followed, it maintains a marked distance from biologists' EGT.

> 'my concept of utility is quite different from the Darwinian concept of fitness, and [my concept of] learning from experience is quite different from natural selection. I am concerned with social evolution and not with genetic evolution, with economics and not with sociobiology.' (Sugden, 1986, p. 26)

He reinterpreted ESS as a comparison of one player's expectations between different options. Expectations start as mere guesses but are updated as the player repeatedly interacts in similar situations, experimenting with different strategies. Thus, no inter-player utility comparisons are required (as in biological ESS), and strategy replication is a matter of the player's choice (not of strategy-bearer reproduction, as in biological EGT).[25]

[24] At around the same time Friedman (1998, p. 18) expressed a similar concern that was not limited to the refinement project: 'Economists must-re-adapt evolutionary theory to economics before evolutionary game models can become routine and widespread. We must shed some of the biological adaptations and develop some new adaptations for economics'.

[25] It is questionable whether Sugden's approach can be subsumed under theories of natural selection (cf. Grüne-Yanoff, 2011).

Table 1
A 2×2 game with fitness payoffs.

	L	R
T	6,2	4,4
B	5,1	2,0

However, Sugden's theory and its distance from biology remain an exception in economics. Other authors seeking to explain social phenomena through EGT have been far more willing to adopt it formally and ontologically. The example that I will briefly discuss here is the indirect evolutionary approach (IEA).

IEA models the evolution of preferences in a population of agents who rationally choose their strategies accordingly (for a brief overview, see Grüne-Yanoff & Hansson, 2009, pp. 19–22). The basic idea is that preferences induce behaviour, behaviour determines 'success', and success regulates the evolution of preferences. What is meant here is *reproductive* success: the ability of a preference to increase the (true) reproduction of its holder through the behaviour it induces.

The mechanism that drives this reproductive advantage is the combined ability of an agent to *commit* to non-equilibrium strategies, and to *signal* this commitment to others. In certain games this kind of ability induces opponents to adjust their strategy choices in a way that enhances the fitness of the agent. Consider the following example (Table 1).

Strategy *T* strictly dominates *B*, and *R* is a strict best response to *T*. The payoffs for each strategy profile (notated first for player 1 and second for player 2) represent individual fitness results. The unique Nash equilibrium is thus (*T*, *R*). However, if player 1 could commit to playing *B*, and make this commitment known to player 2, then player 2 would respond—in order to maximise his or her utility—by choosing *L*. This would lead to result (*B*, *L*), which is better for player 1 than the Nash equilibrium (*T*, *R*).

How can player 1 make such a commitment? According to IEA, nature makes this commitment for the players by endowing them with preferences that distort their fitness values. In other words, a distortion of player 1's fitness values that results in a preference for (B,L) over (T,L) and (B,R) over (T,R) would increase his or her fitness in this game. IEA focuses on the question of which distortions are evolutionarily stable in which populations and for which games.

IEA refers directly to the biological literature. It imports central formal notions of EGT in terms of using its population perspective to construct its models and its stability notions to analyse them. It also imports major ontological elements of biological EGT in explicitly interpreting the underlying payoffs as 'fitness' and envisaging a selection process working on preference distortions. It offers two alternative interpretations of this process: a biological interpretation according to which the behaviour increases the number of the preference-carrier's offspring, who are genetically endowed with the same preference, and a social interpretation suggesting that the behaviour leads to the increased adoption of the preference by others, possibly through learning or imitation. As in the case discussed above, however, the social interpretations are the result of analogy construction, lacking in empirical support and rather derived from the biological interpretations.

4.4. Summary

The above review of biology-to-economics import episodes yielded the following results. First, all economists engaging in EGT have detailed knowledge of the biological literature, and regularly cite biological publications.

Second, those interested in EGT for the sake of equilibrium refinement and evolutionary explanation took over the central EGT concepts from biology. They constructed relatively few new notions, particularly in the early years, and instead explored the consequences of introducing existing formal concepts into the body of economic modelling.

Third, economists adopting EGT for recovery purposes more or less reluctantly imported parts of the biological ontology into their discipline. In some cases it was the ontology of classical game theory that troubled them, and they seemed eager to replace it with a more acceptable and less demanding interpretation of their own concepts from biology—at least to the extent that their previous formal concepts were not disturbed too much. But the ontological commitment of biology did not integrate well with the social outlook of economics. Because economists lacked resources to provide a more fitting re-interpretation, they often engaged in analogy construction, as for example illustrated by the meme concept.

Fourth, many EGT models in economics lack any direct reference to specific phenomena, but attain maximal generality by remaining on the level of abstract formal identity. Furthermore, the adoption of biological ontology is rarely supported in empirical studies of concrete evolutionary processes and dynamics. Instead, analogy constructions to biological processes more often than not replace any attempt at justifying social selection processes independently.

5. Conclusion

In this paper, I pursued three separate goals—to investigate the history of EGT, to learn from this history about the respective modelling practices in biology and economics, and to contribute to the study of interdisciplinarity more generally.

As to the third of these, I have shown that concepts derived from the unity of science discussion or the unification accounts of explanation are too strong and too narrow to be useful for analysing many interdisciplinary exchanges observed in actual scientific practice. Instead, I have suggested investigating such exchanges with respect to involved scientists, formal model elements, model ontologies, and modelling methods and styles. Applying these four dimensions gives a richer picture of interdisciplinary exchanges, and allows clearly distinguishing and contrasting the two exchange episodes under study.

As to the second goal, comparison of the two exchange episodes revealed clear differences between biology and economics. Table 2 contrasts the findings of sections 3 and 4.

These differences are identifiable in all four dimensions. Biologists involved in the development of EGT kept their distance to classical game theory. Those with the strongest impact on EGT

Table 2
Comparing the two import episodes in four dimensions.

	From Economics to Biology	From Biology to Economics
Interdisciplinarity of authors	Inversely proportional to the author's influence on EGT	Detailed knowledge of the biological literature on EGT
Transfer of formal concepts	Only the most basic; a matter of inspiration more than import	Recovery: no import Selection and Explanation: direct import from biology, few novel concepts in the first years
Transfer of ontologies	Complete reinterpretation, often with the help of existing biological theories	Evolutionary interpretation strongly based on biology. More or less reluctant commitment to biological ontology
Methodologies and styles	Empirically studied concrete situations	Intuitive but imaginary stories

seemed to be least connected to classical game theory. Economists who imported EGT, in contrast, had detailed knowledge of the literature in biology, and often interacted with the relevant biologists.

Furthermore, very different things were imported into the respective disciplines. Biology only took basic inspiration-giving ideas from economics. Neither formal results, nor more advanced formal concepts and techniques, nor ontological commitment were transferred from classical game theory. Economics, in contrast, accepted formal concepts and techniques, formal results, and ontological commitment.

Depending on whether one considers a payoff matrix to constitute a computational template, one may say that biologists at most imported rudimentary computational templates. Economics, in contrast, engaged in theory reduction—in the recovery and refinement projects—and in derivational unification in the evolutionary explanation project.

Moreover, economists sought or were forced into some degree of ontological unification in the sense that they re-described economic phenomena in terms of the newly imported evolutionary ideas, for example as strategy population dynamics, natural selection or strategy replication. This was at least their intention, even if they failed to show that these re-descriptions had a factual basis.

Finally, the modelling styles of these two disciplines differ. In biology, there was an evident need to develop a new formal framework, to bestow it with a new ontology and to justify these models and their use with respect to new insights into interesting phenomena and particular empirical cases. Model modifications were driven mainly by attempts to explain concrete phenomena better, which led to countless de-idealisations and re-idealisations. In their use of EGT models biologists expressed their pretension to Galilean science.[26]

Economists, in contrast, mainly used EGT to either recover or expand their formal framework. Ontological commitments were considered flexible, and were welcomed as long as the generation of formal results was not impeded. Phenomena, and in particular empirical studies, played a secondary role. Thereby they depicted a science that is mainly concerned with conceptual exploration and possible explanation (cf. Aydinonat, 2008, p. 112: Grüne-Yanoff, 2009, pp. 96–97). Thus, despite relating to the same theory, the modelling styles associated with EGT in the respective disciplines differ considerably.

As to the first goal of this paper, I hope that the history of EGT has not reached an endpoint, and that its future development continues to be influenced by interdisciplinary exchange. As shown in section 4, EGT in economics is often hampered by an ill-fitting ontology imported from biology. Replacing it with more suitable interpretations would require studying the underlying cognitive and social mechanisms of learning, and adjusting the theory accordingly—as biologists did when importing classical game theory forty years ago. In this regard, economists can still learn a lesson or two from biologists.

Acknowledgements

I thank Luigi Cavalli-Sforza, Richard Lewontin and Jörgen Weibull for patiently answering my questions about events that go back many decades, Joan Nordlund for correcting grammar and spelling, Kirsi Reyes for checking references and in particular Jaakko Kuorikoski, Aki Lehtinen, Uskali Mäki, Catarina Marchionni, Mary Morgan, Jack Vromen and an anonymous referee for extensive comments on earlier drafts of this paper. The usual disclaimers apply.

References

Aumann, R. J. (1985). What is game theory trying to accomplish? In K. Arrow & S. Honkapohja (Eds.), *Frontiers of economics* (pp. 28–76). Oxford: Blackwell.
Aydinonat, N. E. (2008). *The invisible hand in economics: How economists explain unintended social consequences*. London: Routledge.
Axelrod, R. (1980). More effective choice in the Prisoner's Dilemma. *Journal of Conflict Resolution, 24*, 379–403.
Axelrod, R. M., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*, 1390–1396.
Binmore, K. (1988). Modelling rational players I. *Economics and Philosophy, 3*, 179–241.
Binmore, K. (1994). *Game theory and the social contract. Vol. 1: Playing fair*. Cambridge, MA: MIT Press.
Binmore, K. (1998). Review of Axelrod's The complexity of cooperation. *Journal of Artificial Societies and Social Simulation, 1*(1). http://jasss.soc.surrey.ac.uk/1/1/review1.html.
Binmore, K., & Samuelson, L. (1994). An economic perspective on the evolution of norms. *Journal of Institutional and Theoretical Economics, 150*(1), 45–63.
Björnerstedt, J., & Weibull, J. W. (1996). Nash equilibrium and evolution by imitation. In K. J. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The rational foundations of economic behavior* (pp. 155–181). New York: St. Martin's Press.
Bomze, I. (1986). Non-cooperative two-person games in biology: A classification. *International Journal of Game Theory, 15*, 31–57.
Dawkins, R. (1976). *The selfish gene*. Oxford: Oxford University Press.
Dawkins, R. (1989). *The selfish gene* (2nd ed.). Oxford: Oxford University Press.
Dugatkin, L. A. (1997). *Cooperation among animals: An evolutionary perspective*. New York: Oxford University Press.
Fisher, R. A. (1930). *The genetic theory of natural selection*. Oxford: Clarendon Press.
Fisher, R. A. (1958). Polymorphism and natural selection. *Journal of Ecology, 46*(2), 289–293.
Frank, S. A. (1995). George Price's contributions to evolutionary genetics. *Journal of Theoretical Biology, 175*, 373–388.
Friedman, D. (1998). On economic applications of evolutionary game theory. *Journal of Evolutionary Economics, 8*(1), 15–43.
Galison, P. (1997). *Image and logic: A material culture of microphysics*. Chicago: The University of Chicago Press.
Geçkil, I. K., & Anderson, P. L. (2010). *Applied game theory and strategic behavior*. Boca Raton, London, New York: CRC Press.
Gintis, H. (2000). Classical versus evolutionary game theory. *Journal of Consciousness Studies, 7*(1–2), 300–304.
Grüne-Yanoff, T. (2009). Learning from minimal economic models. *Erkenntnis, 70*(1), 81–99.
Grüne-Yanoff, T. (2011). Evolutionary game theory between interpersonal comparisons and natural selection: A dilemma. *Biology & Philosophy*, in press.
Grüne-Yanoff, T., & Schweinzer, P. (2008). The roles of stories in applying game theory. *Journal of Economic Methodology, 15*(2), 131–146.
Grüne-Yanoff, T., & Hansson, S. O. (2009). Preference change: An introduction. In T. Grüne-Yanoff & S. O. Hansson (Eds.), *Preference change: Approaches from philosophy, economics and psychology* (pp. 1–26). Berlin, Heidelberg, New York: Springer.
Güth, W., & Yaari, M. (1992). Explaining reciprocal behavior in simple strategic games: An evolutionary approach. In U. Witt (Ed.), *Explaining process and change – Approaches to evolutionary economics* (pp. 23–34). Ann Arbor: University of Michigan Press.
Hamilton, W. D. (1967). Extraordinary sex ratios. *Science, 156*, 477–488.
Hamilton, W. D. (1971). Selection of selfish and altruistic behaviour in some extreme models. In J. F. Fisenberg & W. S. Dillon (Eds.), *Man and beast: Comparative social behavior* (pp. 57–91). Washington, DC: Smithsonian Press.
Hamilton, W. D. (1996). *Narrow roads of gene land. Vol. 1: Evolution of social behaviour*. Oxford: Oxford University Press.
Humphreys, P. (2004). *Extending ourselves: Computational science, empiricism, and scientific method*. Oxford: Oxford University Press.
Kitcher, P. (1989). Explanatory unification and the spatial structure of the world. In P. Kitcher & W. Salmon (Eds.), *Scientific explanation* (pp. 410–505). Minneapolis, MN: University of Minnesota Press.
Kuhn, H., Harsanyi, J., Selten, R., Weibull, J., Van Damme, E., Nash, J., & Hammerstein, P. (1996). The work of John Nash in game theory. *Journal of Economic Theory, 69*(1), 153–185.
Laubichler, M. D., Hagen, E. H., & Hammerstein, P. (2005). The strategy concept and John Maynard Smith's influence on theoretical biology. *Biology and Philosophy, 20*, 1041–1050.

[26] It is tempting to speculate about the impact game theory had on modelling methodology and style in biology. Clearly, formal modelling was present in biology long before the advent of EGT, as witnessed for example by the Lotka-Volterra equations. Yet the ascent of EGT seems to have boosted the stance of theoretical modelling in biology. Furthermore, it seems to have made popular the idea of "micro foundations" for macro (population-) dynamics into biology—a style of explanation that had firmly taken hold in economics in the 1950s and which greatly contributed to the growth of pure theory. Yet to substantiate these speculations would require an investigation into biology beyond the scope of this paper.

Lazear, E. P. (2000). Economic imperialism. *Quarterly Journal of Economics, 115*(1), 99–146.

Leonard, R. J. (1994). Reading Cournot, reading Nash: The creation and stabilisation of the Nash equilibrium. *Economic Journal, 104*, 492–511.

Lewontin, R. C. (1961). Evolution and the theory of games. *Journal of Theoretical Biology, 1*, 382–403.

Lewontin, R. C. (1982). Review of [Maynard Smith 1982]. *Nature, 300*, 113–114.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.

Mäki, U. (2009). Economics imperialism: Concept and constraints. *Philosophy of the Social Sciences, 39*(3), 351–380.

Mailath, G. J. (1998). Do people play Nash equilibrium? Lessons from evolutionary game theory. *Journal of Economic Literature, 36*(3), 1347–1374.

Maynard Smith, J. (1972). *On evolution*. Edinburgh: Edinburgh University Press.

Maynard Smith, J. (1974). The theory of games and the evolution of animal conflicts. *Journal of Theoretical Biology, 47*, 209–221.

Maynard Smith, J. (1976). Evolution and the theory of games. *American Scientist, 64*, 41–45.

Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.

Maynard Smith, J., & Price, G. (1973). The logic of animal conflict. *Nature, 146*, 15–18.

Morgan, M. (2007). The curious case of the Prisoner's Dilemma: Model situation? Exemplary narrative? In A. Creager, E. Lunbeck, & M. N. Wise (Eds.), *Science without laws* (pp. 157–185). Durham: Duke University Press.

Milgrom, P. (1984). Review of Axelrod's *The evolution of cooperation*. *Rand Journal of Economics, 15*, 305–309.

Nagel, E. (1961). *The structure of science. problems in the logic of scientific explanation*. London: Routledge and Kegan Paul.

Nash, J. F. Jr., (1996). *Essays on game theory*. Cheltenham, UK: Edward Elgar.

Nelson, R. R., & Winter, S. G. (2002). Evolutionary theorizing in economics. *Journal of Economic Perspectives, 16*(2), 23–46.

Odenbaugh, J. (2009). Models. In S. Sarkar & A. Plutynksi (Eds.), *A companion to the philosophy of biology* (pp. 502–524). New York: John Wiley and Sons.

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's Dilemma*. Ann Arbor: University of Michigan Press.

Samuelson, L. (1997). *Evolutionary games and equilibrium selection*. Cambridge, MA: MIT Press.

Segerstråle, U. (2000). *Defenders of the truth: The battle for science in the sociobiology debate and beyond*. Oxford: Oxford University Press.

Selten, R. (1980). A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology, 83*, 93–101.

Schlag, K. (1998). Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits. *Journal of Economic Theory, 78*(1), 130–156.

Schotter, A. (1981). *The economic theory of social institutions*. Cambridge: Cambridge University Press.

Schuster, P., & Sigmund, K. (1983). Replicator dynamics. *Journal of Theoretical Biology, 100*, 533–538.

Sigmund, K. (2005). John Maynard Smith and evolutionary game theory. *Theoretical Population Biology, 68*, 7–10.

Sugden, R. (1986). *The evolution of rights, cooperation, and welfare*. Oxford: Basil Blackwell.

Sugden, R. (2001). The evolutionary turn in game theory. *Journal of Economic Methodology, 8*, 113–130.

Sugden, R. (2009). Credible worlds, capacities and mechanisms. *Erkenntnis, 70*(1), 3–27.

Swedberg, R. (2000). Sociology and game theory: Contemporary and historical perspectives. *Theory and Society, 30*, 301–335.

Taylor, P., & Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences, 40*, 145–156.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology, 46*, 35–57.

van Damme, E. (1987). *Stability and perfection of Nash equilibria*. Berlin, Heidelberg, New York: Springer. Second edition 1991.

von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton: Princeton University Press.

Vromen, J. (2006). Routines, genes and program-based behavior. *Journal of Evolutionary Economics, 16*(5), 543–560.

Vromen, J. (2009). Friedman's selection argument revisited. In U. Mäki (Ed.), *The methodology of economics. Milton Friedman's essay at 50* (pp. 257–284). Cambridge: Cambridge University Press.

Weibull, J. W. (1994a). *The mass-action interpretation of Nash equilibrium. Working paper 427*. Stockholm: The Industrial Institute for Economic and Social Research.

Weibull, J. W. (1994b). The 'as if' approach to game-theory: Three positive results and four obstacles. *European Economic Review, 38*, 868–882.

Weibull, J. W. (1995). *Evolutionary game theory*. Cambridge, MA: M.I.T. Press.

Young, H. P. (1993). The evolution of conventions. *Econometrica, 61*(1), 57–84.

Zeeman, E. C. (Ed.). (1980). *Population dynamics from game theory* Nitecki, Z., & Robinson, C. (Eds.). *Global theory of dynamical systems. Lecture notes in mathematics* (Vol. 819). Berlin, Heidelberg, New York: Springer.