Euro Jnl Phil Sci (2016) 6:343–360
DOI 10.1007/s13194-016-0139-z

ORIGINAL PAPER IN PHILOSOPHY OF SCIENCE

Interdisciplinary success without integration

Till Grüne-Yanoff[1,2]

Received: 22 June 2015 / Accepted: 17 February 2016 / Published online: 7 March 2016

Abstract Some scholars see interdisciplinarity as a special case of a broader unificationist program. They accept the unification of the sciences as a regulative ideal, and derive from this the normative justification of interdisciplinary research practices. The crucial link for this position is the notion of integration: integration increases the cohesion of concepts and practices, and more specifically of explanations, ontologies, methods and data. Interdisciplinary success then consists in the integration of fields or disciplines, and this constitutes success in the sense that unification is epistemically desirable. In contrast to this account, I defend the thesis that successful interdisciplinary interaction does not necessarily imply the integration of these disciplines. I show this at the hand of two cases. In both the case of evolutionary game theory and the case of hyperbolic discounting, genuine interdisciplinary exchange took place. From both exchanges, the respective economic fields emerged substantially altered – it wasn't just a juxtaposition of disciplines in which disciplinary identities remained unchanged. Yet in neither case did the disciplines integrate. Rather, they developed their own concepts and methods, their own explanations, own ontologies, and their own views of what proper data standards were. Furthermore, the fields that emerged from these exchanges were very successful, if measured at the hand of properties like explanatory success, increase of control, bibliometrics and grant yields. Thus, I argue, there are cases of interdisciplinary success without integration.

Keywords Interdisciplinarity · Integration · Success

1 Introduction

Interdisciplinarity is a regulative ideal. One position, and perhaps the dominant one, found in the literature justifies this ideal by identifying interdisciplinarity with

✉ Till Grüne-Yanoff
gryne@kth.se; http://people.kth.se/~gryne/index.html

[1] Royal Institute of Technology (KTH), Brinellvägen 32, 10044 Stockholm, Sweden

[2] Helsinki University, Unionkatu 40A, 00014 Helsinki, Finland

integration: interdisciplinary success consists in the integration of disciplines or fields (Klein 2010; Holbrook 2013; Lattuca 2001). In this paper, I argue to the contrary that integration of disciplines, concepts or methods is not a necessary precondition for interdisciplinary success. I discuss two cases of interdisciplinary model exchange – between biology and economics, and between psychology and economics, respectively – that exhibited successful interdisciplinary interaction, but did not lead to integration of disciplines, concepts or methods. In particular, these interdisciplinary exchange episodes had important epistemic and institutional effects on the respective disciplines, so that one can justifiably attribute interdisciplinary success, but these effects led the respective disciplines into different directions, so that one cannot justifiably identify integration. My conclusion that integration is not a necessary condition of interdisciplinary success is relevant for current science policy, which seems strongly influenced by the integrationist view.

The paper is structured as follows. Section 2 discusses the normative content inherent in the notion of interdisciplinarity. Section 3 reviews some of the prominent accounts of interdisciplinary success as integration. Section 4 provides two cases of interdisciplinary interaction without integration. Section 5 concludes.

2 Interdisciplinarity success

The terms "interdisciplinary or "interdisciplinarity", as used in the current literature, are ambiguous with respect to their normative content.[1] These terms are either used in a descriptive sense, referring to actual states and events of discipline-crossing. Alternatively, they are used prescriptively, referring to ideal states or events that should be reached. This prescriptive aspect was already pointed out by the early authors on interdisciplinarity:

> "interdisciplinarity has to be understood as a teleological and normative concept." (Jantsch 1972, 7)

Interdisciplinarity in this sense of a prescriptive ideal thus describes non-actual states or events that are worthy to be realized. This raises the question of *success*: when and how is the prescriptive ideal of interdisciplinarity accomplished?

The answer to this question is complicated by the following ambiguity. "Interdisciplinary success" might be understood either as "the achievement of interdisciplinarity" or as the achievement of relevant aims through interdisciplinary means. In the first case, interdisciplinarity is an end in itself, and obtaining that end constitutes success. Current views in some science funding bodies seem to suggest such an understanding of interdisciplinary success. For

[1] The body of literature that I am referring to here is itself inter-disciplinary. Philosophers of science have contributed at a comparatively late stage (e.g. van der Steen 1990. Grantham 2004; D'Agostino 2012; Brigandt 2013; O'Malley 2013). Earlier contributions came from author located in very different disciplines (Klein from English literature, Jansch from engineering, Heckhausen from psychology, Gibbons and many of his co-authors from science and technology studies, to give just some examples of authors who began publishing in the 1970s and 80s).

example, the NIH in 2007 invested $210 million over 5 years in the funding of 6 Interdisciplinary Research Consortia. These funding programs are designed, according to the then NIH Director Elias A. Zerhouni

> "to encourage and enable change in academic research culture to make interdisciplinary research easier to conduct for scientists who wish to collaborate in unconventional ways." (NIH News 2007).

According to this statement, funding aims to establish and facilitate interdisciplinarity as agoal in itself. Interdisciplinary success according here consists in the achievement of a state of sufficient interdisciplinarity: for example when such unconventional collaborations across disciplines are indeed easier to conduct.

In contrast to this interdisciplinarity-as-an-end-in-itself approach, interdisciplinarity is also often understood as a means for better accomplishing certain scientific goals. The UK ESRC, for example, states that

> "many of the most pressing research challenges are interdisciplinary in nature, both within the social sciences and between the social sciences and other areas of research." (ESRC 2013)

According to this statement, funding aims to solve research challenges by interdisciplinary means. Because of the interdisciplinary nature of the research challenge, the way how to tackle these challenges must be interdisciplinary as well – and hence interdisciplinary proposals merit funding for instrumental reasons. Interdisciplinary success according to this account consists in meeting the identified challenges through the means of interdisciplinary research strategies.[2]

We can now distinguish different instrumental justifications of interdisciplinarity by the kinds of goals that this interdisciplinarity purportedly supports. For example, one could distinguish non-academic from academic goals. Accordingly, interdisciplinary research strategies might be *non-academically successful* if they contribute to reaching goals like economic efficiency, aesthetic ideals or ideological constraints; while interdisciplinary research strategies might be *academically successful* if they contribute to reaching goals like more detailed explanations, more accurate predictions or more effective control.

In this paper, I show that interdisciplinary success – understood as instrumental academic success of interdisciplinary research strategies – does not require the integration of disciplines. The next section discusses what is meant by integration of disciplines.

[2] The exact role that interdisciplinary research strategies must play is somewhat unclear. Obviously, they must play at least some actual causal role in meeting these challenges for it to be counted as interdisciplinary success. But must that role be necessary (or could it have been achieved otherwise)? And must it be the dominant cause (or is a contributing role enough)? I will not discuss these difficult questions here, but rather focus on the claim that any interdisciplinary success necessitates integration of disciplines.

3 Interdisciplinary success as integration

What is interdisciplinarity? A now widely accepted distinction (judging by encyclopaedia and handbook articles like Klein 2010 or Krott 2003) goes back to the system theorists Erich Jantsch. He distinguished *Interdisciplinarity* from *Multi*- and *Pluridisciplinarity* through the elements of coordination. In interdisciplinarity, "a common axiomatics for a group of related disciplines is defined at the next higher hierarchical level, thereby introducing a sense of purpose" (Jantsch 1972, 10), while in pluridisciplinarity, the various involved disciplines, although cooperation, remain "juxtaposed" (ibid.) and in multidisciplinarity, they work towards a goal simultaneously, but without cooperation. A widely cited[3] current formulation is this:

> "[Multidisciplinary] juxtaposition fosters wider knowledge, information, and methods. Yet, disciplines remain separate, disciplinary elements retain their original identity, and the existing structure of knowledge is not questioned." (Klein 2010, 3)

Interdisciplinarity, in contrast, is characterized as an integration of different disciplines. It consists of "integration of two or more disciplinary languages with the aim of generating a common understanding." (Holbrook 2013, 4) it involves "inter-communication" between disciplines (Klein 2010, 3), "joint definition of variables or categories" (Klein 2010, 5) and ultimately "integrating the approaches of all the participants into the research design" (Klein 2010, 5). Crucially, integration is more than "collaboration" (Klein 2010, 5) and more than collecting elements of knowledge from different disciplines (Klein 2010, 3). What emerges from analysing this literature is that interdisciplinarity differs from multi- or pluridisciplinarity in two aspects. First, the disciplines involved in interdisciplinary interaction change their identity in some relevant way, while those in multi- or pluridisciplinarity do not. Second, the change that disciplines undergo in successful interdisciplinary exchanges leads them to integrate in a relevant way. While accepting the first as a condition of interdisciplinary success, I will argue against the second condition in this paper.

But let me first specify what I mean by integration. Integration has been proposed as a less constraining concept than theory reduction, which still captures a notion of unity between different theories or scientific practices. Historically, integration has therefore replaced theory reduction as a position against the disunity of science (Brigandt 2010). At the same time, it has been argued that that integration concerns multiple dimensions, going beyond the relations of theories (Brigandt 2013). O'Malley (2013) for example distinguishes between integrating explanations, integrating methods (inferential, modelling and experimental methods) and integrating data. Grantham (2004, 143) distinguishes theoretical and practical integration, specifying theoretical integration to consist of relations between explanations, ontological relations, or other conceptual relationships; while specifying practical integration as consisting of heuristic dependence, conformational dependence or methodological integration.

Notably, these accounts move away from exclusively focussing on theories as the object of either unification or integration, and instead broadening the concept to include

[3] 177 citations according to Google scholar in 2015.

various scientific practices. Grantham (2004), following Darden and Maull (1977) proposes scientific *fields* as objects of integration. O'Malley (2013) speaks about integrating practices more generally. I will use the terms (sub-) discipline and field interchangeably, following both Darden and Maull (1977) and Toulmin (1972) in characterising them as institutionally and historically characterised clusters of scientific practices.

To summarize, integration of fields or disciplines affects both their conceptual and the practical aspects. It affects the concepts they use, both in their explanations, as well as in their ontological content. It also affects their practices, specifically their terminology, their methods and their data. In the following discussion, by an *increase in integration* between two or more disciplines I mean an increased overlap between these disciplines in at least one of these categories.

This understanding of interdisciplinarity as an integration of disciplines has been accepted by many authors working in this area today:

> "Most definitions treat the integration of disciplines as the 'litmus test' of interdisciplinarity." (Lattuca 2001, 78).

> "The notion of 'integration' is so widespread in the ID literature that to question whether ID involves integration is almost heretical." (Holbrook 2013, 13)

Consequently, the integrationist position portraits interdisciplinarity, in contrast to multidisciplinarity, as the attempt to integrate multiple disciplinary approaches to a problem, where integration consist in the creation of a common language, joint categories, methods, and a common research design across disciplines.

This apparent consensus on interdisciplinary success also had a great influence on science policy. Funding agencies, when evaluating proposals, often stress the importance of integration as a success criterion for interdisciplinary research:

> "interdisciplinary research *integrates* elements of a wide range of disciplines, often including basic research, clinical research, behavioral biology, and social sciences so that all of the scientists approach the problem in a new way." (NIH News 2007, my emphasis)

> Measures of the value of interdisciplinary research and its impact can be framed as short-term (research breakthroughs, *development of new academic programs*); intermediate- term (effects on industry, public policy, the workforce); and long-term (*creation of new disciplines*). (NSF 2008, 11, my emphasis)

According to these views, interdisciplinary research is successful if it integrates disciplines, creates new academic programs and ultimately new disciplines. This position is often understood in the strong sense that integration is not only a contributing factor to interdisciplinary success, but also a necessary condition for it. It is against this strong claim of interdisciplinary success as integration that I will argue in the following section.

4 Interdisciplinary success without integration

The above claims integration to be a necessary criterion for interdisciplinary success. Undoubtedly, integration is often considered a successful achievement of interdisciplinarity as an end in itself, as witnessed in many current disciplines that have grown out of interdisciplinary research. Examples include cognitive psychology, genomics, bioinformatics, neuroscience, and nanoscience.

Similarly, it would be pointless to deny that integration often also is an important means for epistemic success. For example, insights in molecular biology, specifically the interactions between the different types of DNA, RNA and protein biosynthesis, arose only because this new sub-discipline integrated important elements of genetics and biochemistry.

It might thus be tempting to justify interdisciplinarity through its purported necessary component of integration: only because it integrates, it serves an epistemic purpose. This would cast interdisciplinarity as part of the regulative ideal of science unification. Because science unification is epistemically desirable (as defended e.g. by Kitcher 1999 and Grantham 2004), seeking to increase interdisciplinarity is justified as long as it promotes integration. Unsurprisingly, defenders of the interdisciplinary-as-integration account follow such an unificationist line:

> "the roots of the concepts [of interdisciplinarity] lie in a number of ideas that resonate through modern discourse—the ideas of a unified science, general knowledge, synthesis and the integration of knowledge" (Klein 1990, 19)

According to the defendants of this integration account, unity of methods, concepts and ontologies lies at the heart of the ideal of interdisciplinarity. Failure in "weaving perspectives together", "reaching effective synthesis" or "promot[ing] communication and consensus" (Klein 2008, 7), implies failure of the interdisciplinary research strategy to reach scientific goals. Simply put, failure to integrate implies interdisciplinary failure.

Against this, some authors have pointed out that integration is overvalued as an epistemic objective:

> "science and philosophy alike often exaggerate the importance of interdisciplinary integration (apparently a remnant of the old unified science ideal)" (van der Steen 1990, 25)

Two reasons for such overvaluation have been proposed: first, integration might not be desirable for some epistemic reason, for example because it fosters terminological unity that is not backed up by conceptual unity and hence leads to undue generalisation (van der Steen 1993). Second, in many cases integration might be wrongly perceived as prescribing an outcome, while in fact it might only prescribe a process: the attempt of integration – with the possible outcome that two fields cannot be integrated (O'Malley 2013).

Furthermore, some authors have argued that interdisciplinary communication need not aim at consensus between disciplines (Holbrook 2013). Instead, cooperation between disciplines is problem-centred – when a discipline cannot tackle a problem alone, it seeks resources from other disciplines to help, without necessarily integrating these resources or integrating with that other discipline.

In this section, I seek to substantiate these theoretical contentions with the help of two case studies. Each case, I argue, is an example of interdisciplinary exchange, rather than multi- or pluridisciplinarity, because the involved disciplines are substantially affected. Furthermore, I show that in each case, the interdisciplinary exchange lead to epistemic success – to more detailed explanations, better control, and higher scientific activity. Crucially, the interdisciplinary exchange was an important causal factor in the production of this success. Yet, in contradiction to the interdisciplinarity-as-integration position, these cases do not satisfy the conditions for field or disciplinary integration discussed in section 3. Thus the two cases constitute examples of interdisciplinary success without integration.

4.1 Evolutionary game theory

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 for economics. Then, roughly 20 years later, economists adopted what biologists had made of it.[4] From either transfer, the importing discipline came out considerably affected. In biology, the game models allowed an improved representation of frequency dependent fitness – a theoretical concept that had been developed in biology many decades earlier, but which until now had been hampered by the lack of powerful modelling tools. When game theory provided these, biologists started a new wave of explanatory projects, applying game models to many different phenomena of cooperation and conflict between organisms. This considerably increased the number of mathematical biologists, thus supporting the rise of a new sub-discipline in biology.

In economics, the re-imported evolutionary game models offered new equilibrium selection and refinement strategies, and most importantly suggested new perspectives of justifying standard game theoretic solution concepts like the Nash Equilibrium. Through this justification, the transfer from biology affected a wide circle of (often applied) economists far beyond those who actually worked with the new models (Sugden 2001). This "evolutionary turn" in game theory generally, combined with those game theorists who began working on evolutionary game models themselves, considerably affected the discipline as a whole.

At first sight, it might seem that these interdisciplinary transfers led to the integration of these two disciplines. After all, both disciplines were affected by imports from the other, suggesting that they incorporated elements from each other and thus converged. In the following, I show that this was not the case. Although there were attempts at integration (at least from the side of economics), these attempts were soon frustrated through various ontological, conceptual and methodological obstacles. The real changes, instead, arose from attempts to deal with these obstacles. In trying to overcome them, scientists from both discipline worked out discipline-specific concepts and methods, and in that process moved their discipline away from the other. The

[4] Evolutionary game theory also has been adopted by other disciplines, notably philosophy. Although an interesting topic in its own right, I will disregard it here.

interdisciplinary transfer thus was a fruitful trigger of new modelling efforts in the respective disciplines – yet this triggered conceptual and methodological developments in different directions, rather than leading to an integration of the two disciplines. This dynamic is particularly obvious in the transfer of evolutionary game models from biology to economics, and I will focus on that episode in the following.

Transfer of game models from biology to economics began in the early 1980s. Two phases can be identified: a pioneering phase, where the goal of integration seemed possible and desirable, and a consolidation phase, where the ontological, conceptual and methodological obstacles came to the fore and the goal of integration progressively lost its appeal.

The *pioneering phase* began with the political scientist Robert Axelrod, who in 1980 published in a social-science journal (the *Journal of Conflict Resolution*) a paper applying the concept of an Evolutionary Stable Strategy (ESS) to a game involving humans. The ESS was first developed by the biologist Maynard Smith in 1972. It describes an equilibrium state in a population, in which most individuals have adopted a strategy S. This S is an ESS if it does not pay for any individual to change strategy when playing with randomly matched members of that population. Axelrod employed the ESS concept in order to analyse a computer simulation tournament between different strategies for the iterated Prisoners' Dilemma. He had called for the submission of optimal strategies for this game, and authors from many disciplines had contributed their best bet. The winning strategy of this tournament was TIT-FOR-TAT, which started out cooperatively and then reciprocated the previous move of the opponent. In order to show its stability, Axelrod 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 not only sought to integrate his social science study conceptually, but also ontologically and methodologically. Ontologically, he switched from a 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)

Mainstream game theorists – then as today – interpret game payoffs as utility, viz. as a numerical representation of the players' preferences. By instead interpreting payoffs as a measure of fitness, Axelrod sought to introduce a biological ontology into economics, thus integrating the two disciplines.

Methodologically, Axelrod employed a computer simulation – a method that was pioneered by evolutionary biologists, amongst others by Maynard Smith and Price.[5] Mainstream game theorists up to then rarely employed such techniques, but rather insisted on analytic solutions. Thus also methodologically, Axelrod sought to integrate the two disciplines.

[5] For an account of the pioneering role of Maynard Smith for agent-based modelling, see Sigmund (2005, 9).

Perhaps because he introduced so many foreign elements into the social sciences at the same time, Axelrod was rather critically received in economics. Nevertheless, his employment of evolutionary game theory was widely noted, and the subsequent literature eventually established his 1980 paper as a key reference. In contrast to this gradual acceptance in economics, his institutional connections to biology were more robust from the start. Notably, he co-authored a biology paper – on the evolution of cooperation – with one of the early developers of evolutionary game theory, the biologist William Hamilton (Axelrod and Hamilton 1981).

Other pioneers of evolutionary game theory showed similar integrative tendencies as Axelrod. The economist Selten (winner of the Nobel Memorial Prize in Economics 1994), for example, contributed to evolutionary game theory in biological journals (e.g. Selten 1980). Furthermore, he supervised the doctoral thesis of Peter Hammerstein, a biologist, who became an important author in biological evolutionary game theory. Similarly, the mathematical economist Immanuel Bomze proved the identity or implication relation between classical equilibrium notions and biologists' stability notions, which allowed consecutive authors to incorporate elements of evolutionary game theory in their proposals for game theoretic equilibrium selection (Bomze 1986).

Yet this pioneering stage with its integrative impulse was not to last. Those who applied evolutionary game theory to social phenomena in earnest soon found that the concepts, methods and ontology that came with evolutionary game theory from biology were often not well-suited for their purposes. This started the *consolidation phase*. The ontological differences were perhaps the first that were explicitly noted. Robert Sugden, for example, when presenting his evolutionary game model to explain the development of social conventions, wrote already in 1986:

> '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)

Instead of pursuing the biologists' path of modelling reproduction of genetically endowed traits, economists now increasingly sought analogical construction to gene reproduction (like meme replication), selection mechanisms different from natural selection (including various concepts of social learning) and interpretations of payoff functions that did not rely on fitness (viz. various evaluative functions).

Soon after, these ontological differences also led to the insight that the solution concepts of the evolutionary models needed adjustment. Instead of relying on the ESS and the Replicator Dynamic (RD) alone, economists searched for more flexible dynamics, which still exhibited the desirable mathematical properties of the RD, but that fitted better with the ontological assumptions about social learning. Samuelson's warning from 1997 is noteworthy here:

> 'although we have much to learn from biological evolutionary models, we must do more than simply borrow techniques from biologists' (Samuelson 1997, p. 37).

Instead of borrowing techniques, economists now expanded them to concepts useful for their purposes (e.g. monotone learning dynamics as a more abstract class of which the RD is a special case) or devised completely new ones (e.g. like stochastic dynamics, which introduced an aspect of mutation in the otherwise static RD).

This is only a brief sketch of a complex and multifaceted development (for a more detailed account, see Grüne-Yanoff 2011). Yet it suffices to support my argument that there can be successful interdisciplinary exchange without integration. Let me clarify the following questions, so that one can see more clearly why it supports my argument.

First, why is this an interdisciplinary exchange, rather than a pluri- or multidisciplinary one? Because there was a genuine exchange, which was the (albeit indirect) cause of substantial change in certain areas of game theory (what then became evolutionary game theory). One can see this with respect to the development of concepts: in the pioneering phase, concepts like the ESS, the RD, Darwinian fitness and differential reproduction were imported into economics, to then be later – in the consolidation phase – adjusted and developed into economic concepts. One can also see this with respect to methods like the use of computer simulations and phase diagrams, which was first introduced from biology and manipulated later. This left the area of economic game theory substantially changed, and the impulse for this change clearly came from the exchange with biology. Thus this exchange is a case of interdisciplinarity, which relevantly changed the field, rather that a case of pluri- or multidisciplinarity, in which the disciplines are "juxtaposed" and retain their original identity.

The second question I want to address is why this is not a case of integration. The answer is that the identity of the field of evolutionary game theory as it exists today was shaped in the consolidation phase, not the pioneering phase. Thus, it was shaped by a counter-reaction to the imported concepts and methods, not directly by those inputs themselves. This can be observed in conceptual, methodological and institutional aspects. Conceptually, I have already described the adjustment of imported concepts (e.g. monotone dynamics) and the generation of new ones (e.g. stochastic dynamics).

Economists similarly distanced themselves from the methodological premises under which biologists had developed and employed evolutionary game theory. They employed different modelling methods, largely steering clear of computer simulations and instead insisting on analytic modelling techniques. Furthermore, they used models for different purposes. While biologists employed evolutionary game models predominantly to explain adaptive properties of behavioural patterns found in specific biological phenomena (e.g. cleaner-fish - host cooperation, sex ratio in mammals, shape of antlers in deer), economists often employed evolutionary game models for much more abstract purposes. These include – as discussed above – equilibrium refinement and selection, as well as a providing an evolutionary justification for standard concepts like the Nash equilibrium. Only a minority of economists working with evolutionary game theory used them for explanatory purposes similar to those of the biologists (for example, evolutionary explanation of trust, reciprocity and norms). These changes in concepts and methods clearly require the originally imported concepts and methods as preconditions. Thus the exchange indirectly produced these changes. One must therefore consider this a case of interdisciplinary exchange, albeit without integration.

These conceptual and methodological differences were also reflected in various forms of continuing and even increasing institutional distances. All economists engaged in evolutionary game theory had detailed knowledge of the biological literature, and regularly cited biological publications. However, in the consolidation phase, the apparent need arose to develop economic publication channels for their papers, separate from the biological ones. For example, the *Journal of Evolutionary Economics* was founded 1991. It lists "evolutionary games" as part of its Aims and Scope description, but these are clearly reserved for evolutionary games *in economics*, not biology. Similarly, when in 1992 a special issue on evolutionary game theory was published in the *Journal of Economic Theory*, contributors were exclusively economists. When D. Friedman in 1998 declared, in an article that reviewed the theory of learning in games, that "Evolutionary economics goes mainstream" it was thus clear that evolutionary game theory was moving towards mainstream economics (and hence away from biology) rather than converging on an integration of the two fields.

This does not mean that evolutionary game theory had *become* the mainstream. Even in game theory, it is practiced today only by a small part of the community. Yet where this relatively small group is institutionalized, it exhibits a decidedly economic identity, distinct from biology.[6] No new discipline or sub-discipline has arisen that would integrate insights and practices from biology and economics. Serious collaboration seems rare even on the individual level.[7] Nor do I know of any joint teaching programs that would instruct both economists and biologists together in evolutionary game theory. Consequently, it is not surprising that the number of "defectors" (students who take their PhD in one discipline but are then employed in the other) is low. In short, the institutional separation between biology and economics is considerable and seems to even have increased since the 1990s.[8] Consequently, one cannot identify an increase in integration between the respective fields in biology and economics, either with respect to explanation or ontological commitments, nor with respect to their methods or their data.

The third question I want to address is why this case was one of interdisciplinary *success*. Despite the fact that the interdisciplinary exchanges between economics and biology have not led to their integration, they have led to very successful developments in the importing disciplines, in particular in economics. In the first instance, the

[6] For example, the *Canada Research Chair in Economic Theory and Evolution*, founded in 2002 is supposed to deliver "Theoretical work on the relationship between evolutionary biology and human economic choice, attitudes toward risk, and strategic interactions." Yet in practice, this chair is occupied by an economist, who has published only 2 papers in biology outlets.

[7] When searching the database of NSF-funded projects over the last 20 years, from the 36 funded projects relating to evolutionary game theory, not a single involved a collaboration between economists and biologists.

[8] Recently, there seems to be a development of evolutionary game theory as a "neutralized tool" to be used across different disciplines: problem-oriented, but without disciplinary identity. This development might have been heralded already by Jörgen Weibull's textbook from 1995, which stressed mathematical analysis, but had little to offer in terms of applications. Two new journals are the result of this development: *Dynamic Games and Applications*, founded 2011, "is devoted to the development of all classes of dynamic games, [including] evolutionary games" and its coverage includes "applications to economics and management science, biology and ecology,…" (Website). Similarly, the *Journal of Dynamics and Games* (founded 2014) is an applied mathematics journal, which includes papers on evolutionary game theory and aims at applications to economics as well as biology. The need for such a neutral identity, in my view, arises from the increased disciplinary distances of those parts of biology and economics where this neutral theory is supposed to be applied.

exchange has led to the development of new tools that allow the representation of evolutionary and learning dynamics not possible before. Second, it led to the elaboration of new solution concepts and to the devising of novel equilibrium selection tools, some of which were crucial for the acceptability of game theory at a critical phase (Sugden 2001). Third, the relatively small number of practitioners has had highly visible publications in top-ranking journals, and they were highly successful in acquiring external funding: Since, 1991, the NSF has granted 24 large funding projects for evolutionary game theory projects in the Social, Behavioural & Economic Sciences.[9] Thus, the development of evolutionary game theory in economics is an example of scientific success through interdisciplinary means, yet without integration with the transferring discipline of biology.

4.2 Hyperbolic discounting

The hyperbolic discounting function is the product of the intermittent collaboration between two disciplines, economics and psychology. Modern microeconomics began using an inter-temporal discounting function (of exponential shape) as a standard model assumption in the 1930s. Cognitive psychology from the 1960s on sought to experimentally determine the temporal discounting of reward effects first in animals and then in humans. In the early 1970s, these two efforts were brought together through the work of George Ainslie and others. This synthesis maintained that the hyperbolic shape of the discounting function was a universal model assumption that could nevertheless be empirically measured. The consensus began to crumble in the early 1990s, when the consequent measurement attempts gave widely diverging results. Subsequently, economists focused increasingly on an axiomatised and tractable representation of the discounting function that allowed modelling a rational response to inter-temporal inconsistency. Psychologists, in contrast, focused more on empirically observable strategies that people use in order to deal with inter-temporal inconsistency, either leaving the concept of an inter-temporal discounting function behind or referring to it merely as a theoretical schema (for a detailed account, see Grüne-Yanoff 2015).

The development of the hyperbolic discounting function began in the 1970s. Two phases can be distinguished: the collaboration phase, and the later separation phase. The *collaboration phase* began when George Ainslie, a clinical psychiatrist by training, joined Richard Herrnstein's pigeon lab at Harvard's psychology department in 1967. With his interdisciplinary background between psychiatry, psychology and economics, Ainslie was able to interpret the pigeon lab's results in a way that the psychologists had not done before: "When I pointed out [to Herrnstein] that the matching formula implied a hyperbolic discount curve … he set me up in his laboratory … then he waited patiently for the 6 years it took to show that pigeons have the expectable intertemporal conflict" (Ainslie 2001, x).

Herrnstein's matching law described a correlation between the relative rates of response and the relative rates of reinforcement in concurrent schedules of reinforcement. Ainslie's perspective shifted the focus away from experimenting with concurrent schedules at variable time intervals and towards discrete trials in which sooner smaller

[9] Economists were also comparatively more successful than other disciplines. While they acquired 24 NSF grants for evolutionary game projects since 1991, mathematicians only obtained 7 and biologists only 5.

rewards are compared to larger later ones. This new experimental design allowed plotting the effect of reward as a function of delay – a functional representation close to the economists' representation of an inter-temporally discounted utility function. In his 1975 paper, Ainslie for the first time drew these discounting curves. In this paper, he also coined the term "hyperbolic discounting".

Thus the hyperbolic discounting function was born in psychology. It offered a functional representation similar in form but different in content from those proposed by earlier economists. At the same time, it was consistent with the psychologists' findings from the pigeon lab, but represented this content in a novel form. Through this new representational form, novel research questions arose, including the observability of intertemporal inconsistency, the measurability of the discounting rate, and the nature of self-control.

These questions had already arisen in the economic literature in the 1950s. But it remained largely unappreciated until the transfer between psychology and economics happened in the late 1970s, fuelled by Ainslie's synthesising role.[10] Once this basis had been laid, economists adopted both model concepts as well as experimental methods from psychology.

Through this conceptual and methodological transfer, the collaboration phase also saw a strengthening of institutional ties between psychologists and economists. Ainslie himself contributed to this in various ways. He introduced psychologists to the economic literature. In the 1980s, he moved intertemporal experiments from animals to humans. Notably, the title of a 1981 paper is "The application of economic concepts to the motivational conflict in alcoholism", referring to hyperbolic discounting as an "economic concept". This institutionally bridging function cumulated in his 1991 article in the *American Economic Review*, one of the leading economic journals, on "Intertemporal choice. Derivation from 'rational' economic behaviour from hyperbolic discounting curves." Other interactions included Drazen Prelec, who, after working in Herrnstein's lab as an undergraduate and getting his PhD at Harvard in Experimental Psychology in 1983, became a leading contributor to inter-temporal choice research in economics. Similarly David Laibson, another leading economist working on intertemporal choice, in the 1980s sat in undergraduate classes of Herrnstein and Ainslie and considers the two as "two of my most important influences" (personal communication). The 1980s also saw a number of conferences (e.g. 1982 organised by John Elster and 1985 organised by the Sloan foundation), in which many leading researchers from both fields - amongst others Thaler, Kahneman, Festinger, Loewenstein, Mischel, Ainslie, Summers, Schelling and Baumol – participated.

These increased collaborations soon yielded new research results in economics. Thaler (1981) showed with human experimental data that discounting functions are often highly non-exponential. Loewenstein (1987) showed that implied discount rates are much higher for questions about monetary gains than for losses. Benzion et al. (1989) showed that larger outcomes are discounted at a lower level than smaller ones. Many of these papers also quote some of the psychological literature. From this, a

[10] Strotz (1956) paper, the most visible expression of these early economic views, was cited only 24 times in economic journals until 1980, as a search of the *Social Science Citation Index* [SSCI] reveals. Since the 1980s, however, it has been cited more than 420 times. Thus, it required the mediation through behavioural economics that made Strotz' paper a classic in the economics literature.

veritable measurement tradition developed. Frederick et al. (2002) lists 42 such measurement papers until 2002 alone – both from psychologist and economists.

This increased research activity however showed that the idea of a hyperbolic discounting function as a universal model assumption that could be empirically measured was unfounded. The measurements yielded fantastically divergent results. Discount rates varied with income and other personal characteristics, and also with the kind of goods or events that were discounted. Numerically, measured rates ranged from 3 to 96,000 %. In some cases, the discounting factor turned out to be negative. Consequently, economists in the late 1980s had to scale back their ambitions and realise that their measurement efforts would not yield a universal discounting function – and not even a universal functional *form* – pertinent to their explanatory and predictive needs. This led to multiple conceptual and methodological adjustments, which brought with it the institutional separation from psychology, starting the *separation phase.*

Conceptually, economists gave up the idea of a directly measurable discount rate, and instead interpreted the discounted utility model as an idealised model that described the evaluations of temporally distributed *multiple selves* and their relation to each other. This model would then serve as a plausible assumption in a model that explained how people would rationally engage in pre-commitment behaviour or exhibit overspending and impulsive behaviour in the absence of such commitment devices.

This conceptual change had a number of methodological implications. First, an axiomatisation of inter-temporal utility was formulated (Prelec 1989; Loewenstein and Prelec 1992). Notably, this axiomatisation very much stayed within the framework of the standard economic model, in fact including the exponential discounting form as a special case. Second, a more tractable functional form was proposed for the discounting function, which "mimics the qualitative property of the hyperbolic discounting function, while maintaining most of the analytical tractability of the exponential discount function" (Laibson 1997, 450). This *quasi-hyperbolic* discounting function clearly is not suitable for measurement (as it only preserves the hyperbolic function's qualitative property) but it is enough for the new explanatory purposes. Third, where economists experimented, they could now impose their own, more stringent requirements. The direct measurement project had burdened economists with "enormous tactical problems" of providing real incentives: "would subjects believe that they get paid in 5 years?" (Thaler 1981, 207). Now that his project had been dropped, real monetary incentives could be arranged for all relevant experimental designs. Finally, the concept of temporally distributed multiple selves allowed an integration of the theory of inter-temporal choice into game theory and hence into mainstream economic theory.

> "On one level, the idea of multiple selves…is a radical departure from the utility-maximising framework. But because this conceptualisation of intertemporal choice uses a familiar tool – dynamic game theory – it is ready made for adoption by economists interested in improving the behavioural realism of our models" (Rabin 1998, 40).

This conceptual shift, combined with the methodological implications of axiomatisation, tractability, experimental design and theory unification pushed economic research away from the psychological perspectives that had been dominant in the pigeon lab, and also from current psychological trends (which focus more on procedural models and attempt to empirically identify coping strategies). This separation also expresses itself institutionally: publications by economists working in this field have largely stopped citing psychological literature, and they now typically appear in mainstream economic journals (like the *Quarterly Journal of Economics*, *American Economic Review*, *Journal of Economic Behaviour and Organisation*). Of a sample of 172 NSF grants on topics involving "discounting" or "self-control", granted since 1988, not a single one was identifiable as a collaboration between psychologists and economists. In short, the institutional separation between psychology and economics is considerable in the area of inter-temporal choice.[11]

Thus, the two disciplines did not integrate, neither with respect to explanation or ontological commitments, nor with respect to their methods or their data. However, from either transfer, the importing discipline came out considerable affected. In psychology, the reception of the economic discounting model yielded a functional representation for the experimental data; it helped psychologists focus their research interest on inter-temporal motivational conflict; and it offered a normative benchmark against which the notion of self-control could be made more precise. In economics, grappling with the psychological research introduced a new form of the discounting function, and provided empirical evidence for that form. Furthermore, it offered considerable experimental expertise and empirical evidence on a topic economist before had never empirically treated. Last, it introduced the notion that people could be systematically deviate from the normative ideal, and thus gave rise to the question how to rationally deal with this. Thus again, as in the previous case, the exchange was truly interdisciplinary, rather than multi- or pluridisciplinary. But it did not lead to an increase of integration between psychology and economics. The interdisciplinary transfer was a fruitful trigger of new modelling efforts in the respective disciplines – yet it triggered conceptual and methodological developments in different directions, rather than leading to an integration of the two disciplines.

Despite the fact that the interdisciplinary exchanges between economics and psychology have not led to their integration, they have led to very successful developments in the importing disciplines, in particular in economics. In the first instance, the exchange has lead to new (tractable) representations of motivation that explain self-control behaviour (or lack thereof) and its various related economic phenomena (as illustrated e.g. in Laibson 1997). Second, the exchange has introduced a very productive conflict between positive and normative models of behaviour, which led to a whole battery of policy proposals (e.g. Thaler and Sunstein 2008), many of which have now been implemented by administrations in different countries. Third, behavioural economics, which comprises work on inter-temporal choice as one of its main pillars, has rapidly ascended in size and prestige within the economics profession. Although not uncontroversial, it must these days be counted towards the economic mainstream, with

[11] A possible exception of this observation can be found in some business schools. There, it seems, scholars are encouraged to work on topics that conceptually, methodologically and institutionally straddle both disciplines. I give just two examples with quotations from their respective websites.

- "Professor Weber [Columbia Business School] works at the intersection of psychology and economics. She is an expert on behavioral models of judgment and decision making under risk and uncertainty."
- "Vladas Griskevicius [Carlson School of Management] has published over 40 articles in top business and psychology journals examining sustainability, green marketing, motivation, emotion, social influence, social norms, and conspicuous consumption."

its representatives winning some of the major awards in the discipline,[12] and them occupying numerous named chairs.[13] Behavioural Economics is now presented in numerous textbooks, and also in many popularizing monographs. It was highly successful in acquiring external funding: since 1985, the NSF has granted 144 large funding projects for discounting-related projects in the Social, Behavioural & Economic Sciences.[14] Thus, the development of inter-temporal choice models in economics is an example of scientific success through interdisciplinary means, even though this development did not lead to an integration of the disciplines of psychology and economics.

5 Conclusions

Some scholars see interdisciplinarity as a special case of a broader unificationist program. They accept the unification of the sciences as a regulative ideal, and derive from this the normative justification of interdisciplinary research practices. The crucial link for this position is the notion of integration: integration increases the coherence of concepts and practices, and more specifically of explanations, ontologies, methods and data between fields. Interdisciplinary success then consists in the integration of fields or disciplines, and this constitutes success in the sense that unification is epistemically desirable.

In contrast to this account, I defended the thesis that successful interdisciplinary interaction does not necessarily imply the integration of these disciplines. I showed this at the hand of two cases. In both the case of evolutionary game theory and the case of hyperbolic discounting, genuine interdisciplinary exchange took place. From both exchanges, the respective economic fields emerged substantially altered – it wasn't just a juxtaposition of disciplines in which disciplinary identities remained unchanged. Yet in neither case did the disciplines integrate. Rather, they developed their own concepts and methods, their own explanations, own ontologies, and (to the extent that they used data at all) they had very discipline-specific views of what proper data standards were. They could not have generated these novel concepts and methods without being exposed to the interdisciplinary exchange. But the development was an active counterreaction to this exposure, which led them into different directions from each other, rather than a convergence on to an integrated field.

The surprising thing, perhaps, is that despite this lack of integration, these interdisciplinary exchanges were not failures. To the contrary, they were very successful, if measured at the hand of properties like explanatory success, increase of the ability to control, bibliometrics and grant yields. This was possible because the interdisciplinary exchanges acted as inspiring impulses for the importing disciplines. That is, they first imported bits of another discipline because they hoped it would help them solve their

[12] One Nobel (Kahneman) and one Clark Medal (Rabin).

[13] To name but a few: the Chair for Behavioral and Experimental Economics at LMU Munich; the Slater Family Behavioral Economics Chair at Boston University; the Paul A. Volcker Chair in Behavioral Economics at Syracuse University; and the Research Chair of Decision Theory and Behavioral Game Theory at ETH Zürich.

[14] Economists were also comparatively more successful than other disciplines. While they acquired 144 NSF grants for "discounting" or "self-control"-related projects, psychologists only acquired 28.

problems. Then they learned that these imports could not really do the job they had hoped, and started expanding and revising these imports to suit their needs.

One might speculate that these extra-disciplinary impulses re-levelled the “balance between exploration and exploitation” (March 1991, 71) that characterises optimal organisational learning, and which in the importing disciplines perhaps had swung too far away from exploration and renewal. But instead of embarking on the high-risk adventure of field integration, these fields then reverted to the epistemic safety of their respective disciplines, where they would exploit these newly-won impulses through disciplined inquiry (for a discussion of this balance in a context different from the current paper, see D’Agostino 2012).

Such speculations aside, the result of these exchanges were successful and highly discipline-specific developments, which had been caused by interdisciplinary exchanges, but which did not lead to integration. Thus, there are cases of interdisciplinary success without integration, free from unificationist pretensions.

Acknowledgement Previous versions of this paper were presented at the workshop “What is Interdisciplinary Success?” in Lund and at the pre-EPSA symposium “Towards Philosophies of Interdisciplinarity” in Helsinki. I thank both of these audiences for very valuable comments. Financial support from TINT, the *Finnish Centre of Excellence in the Philosophy of Social Science*, made this research possible. Thank you!

References

Ainslie, G. (2001). *Breakdown of the will*. Cambridge: Cambridge University Press.

Axelrod, R. (1980). More effective choice in the Prisoner’s Dilemma. *Journal of Conflict Resolution, 24*, 379–403.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390–1396.

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: an experimental study. *Management Science, 35*, 270–284.

Bomze, I. (1986). Non-cooperative two-person games in biology: a classification. *International Journal of Games Theory, 15*, 31–57.

Brigandt, I. (2010). Beyond reduction and pluralism: toward an epistemology of explanatory integration in biology. *Erkenntnis, 73*(3), 295–311.

Brigandt, I. (2013). “Integration in biology: philosophical perspectives on the dynamics of interdisciplinarity” (introduction to the special section). *Studies in History and Philosophy of Biological and Biomedical Sciences, 44*, 461–465.

D’Agostino, F. (2012). Disciplinarity and the growth of knowledge. *Social Epistemology, 26*(3–4), 331–350.

Darden, L., & Maull, N. (1977). Interfield theories. *Philosophy of Science*, 43-64.

ESRC (2013). *Guidance for applicants*. http://www.esrc.ac.uk/funding-and-guidance/applicants/. Accessed 19.8.2013.

Frederick, S., Loewenstein, G., & O’Donoghue, T. (2002). Time discounting and time preference: a critical review. *Journal of Economic Literature, 40*(2), 351–401.

Grantham, T. A. (2004). Conceptualizing the (Dis) unity of science. *Philosophy of Science, 71*(2), 133–155.

Grüne-Yanoff, T. (2011). Models as products of interdisciplinary exchange: evidence from evolutionary game theory. *Studies in History and Philosophy of Science, 42*, 386–397.

Grüne-Yanoff, T. (2015). Models of temporal discounting 1937–2000: interdisciplinary exchanges between economics and psychology. *Science in Context, 28*(4), 675–713.

Holbrook, J. B. (2013). What is interdisciplinary communication? Reflections on the very idea of disciplinary integration. *Synthese, 190*(11), 1865–1879.

Jantsch, E. (1972). Inter-and transdisciplinary university: a systems approach to education and innovation. *Higher Education, 1*(1), 7–37.

Kitcher, P. (1999). Unification as a regulative ideal. *Perspectives on Science, 7*(3), 337–348.

Klein, J. T. (1990). *Interdisciplinarity: History, theory, and practice*. Detroit: Wayne State University.

Klein, J. T. (2008). Evaluation of interdisciplinary and transdisciplinary research: a literature review. *American Journal of Preventive Medicine, 35*(2), S116–S123.

Klein, J. T. (2010). A taxonomy of interdisciplinarity. In R. Frodeman, J. T. Klein, & C. Mitcham (eds.). *The Oxford handbook of interdisciplinarity* (pp. 15–30). Oxford University Press.

Krott, M. (2003). Evaluation of transdisciplinary research. In: *Encyclopedia of life-support systems*. Oxford: EOLSS Publishers.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112*, 443–477.

Lattuca, L. R. (2001). *Creating interdisciplinarity: Interdisciplinary research and teaching among College and University faculty*. Vanderbilt University Press.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal, 97*, 666–684.

Loewenstein, G., & Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics, 107*(2), 573–597.

NIH News (2007). http://www.nih.gov/news/pr/sep2007/od-06.htm. Accessed 19.8.2013.

NSF (2008). IGERT *Workshop Report*, NSF 09-33. http://www.nsf.gov/pubs/2009/nsf0933/index.jsp?govDel=USNSF_124. Accessed 19.8.2013.

O'Malley, M. A. (2013). When integration fails: prokaryote phylogeny and the tree of life. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 44*(4), 551–562.

Prelec, D. (1989). Decreasing impatience: Definition and consequences. *Harvard Business School Working Paper.*

Rabin, M. (1998). Psychology and economics. *Journal of Economic Literature, 36*(1), 11–46.

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

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

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

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies, 23*(3), 165–180.

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.

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters, 8*, 201–207.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Toulmin, S. (1972). *Human understanding: The collective use and evolution of concepts*. Princeton University Press.

van der Steen, W. J. (1990). Interdisciplinary integration in biology? An overview. *Acta Biotheoretica, 38*(1), 23–36.

Van Der Steen, W. J. (1993). Towards disciplinary disintegration in biology. *Biology and Philosophy, 8*(3), 259–275.