Erkenn (2013) 78:1143–1175
DOI 10.1007/s10670-012-9414-3

ORIGINAL ARTICLE

Who Got What Wrong? Fodor and Piattelli on Darwin: Guiding Principles and Explanatory Models in Natural Selection

José Díez · Pablo Lorenzano

Received: 26 January 2012 / Accepted: 14 November 2012 / Published online: 18 December 2012

Abstract The purpose of this paper is to defend, contra Fodor and Piattelli-Palmarini (F&PP), that the theory of natural selection (NS) is a perfectly bona fide empirical unified explanatory theory. F&PP claim there is nothing non-truistic, counterfactual-supporting, of an "adaptive" character and common to different explanations of trait evolution. In his debate with Fodor, and in other works, Sober defends NS but claims that, compared with classical mechanics (CM) and other standard theories, NS is peculiar in that its explanatory models are a priori (a trait shared with few other theories). We argue that NS provides perfectly bona fide adaptive explanations of phenotype evolution, unified by a common natural-selection guiding principle. First, we introduce the debate and reply to F&PP's main argument against NS. Then, by reviewing different examples and analyzing Fisher's model in detail, we show that NS explanations of phenotypic evolution share a General Natural Selection Principle. Third, by elaborating an analogy with CM, we argue against F&PP's claim that such a principle would be a mere truism and thus explanatorily useless, and against Sober's thesis that NS models/explanations have a priori components that are not present in CM and other common empirical theories. Irrespective of differences in other respects, the NS guiding principle has the same epistemic status as other guiding principles in other highly unified theories such as CM. We argue that only by pointing to the guiding principle-driven nature that it shares with CM and other highly unified theories, something no-one has done yet in this debate, one can definitively show that NS is not defective in F&PP's sense:

The title refers to the Sober–Fodor discussion in i-net (Sober and Fodor 2010).

J. Díez (✉)
LOGOS Research Group, Universidad de Barcelona, Montalegre 6, 08001 Barcelona, Spain
e-mail: jose.diez@ub.edu

P. Lorenzano
Universidad Nacional de Quilmes, CONICET, Quilmes, Argentina
e-mail: pablol@unq.edu.ar

in the respects relevant to the debate, Natural Selection is as defective and as epistemically peculiar as Classical Mechanics and other never questioned theories.

1 The Debate and F&PP's Main Strategy

> *Adaptationism simply cannot do what an evolutionary theory is supposed to do—explain how phenotypic traits are distributed in a population of organisms*. (Fodor and Piattelli-Palmarini, 2010a, p. 110)
> *The circulation of the proof copy of What Darwin Got Wrong […] has resulted in a volume of critical comment from biologists and philosophers that has not been seen since 1859* (Lewontin, 2010, fn. 1)

As Lewontin's quote shows, the case made by Fodor (2008a, 2008b) and Fodor and Piattelli-Palmarini (2010a-F&PP henceforth-, 2010b) against Darwinism has had an enormous impact among philosophers in general, and among philosophers of biology in particular, and has provoked one of the biggest philosophical controversies in recent decades. Their charge is, in brief, that

> (*) there is nothing non-truistic, counterfactual-supporting (i.e. that grounds the distinction between selection for and selection of), of "adaptive" character, and common to different explanations of traits evolution; the evolution of traits is in fact explained by a kind of "disconnected" natural history.

Sober, among others, defends NS from this charge, but in his debate with F&PP, and in other works, he also claims that NS has a priori specific components which are not present in other standard theories such as Classical Mechanics (CM) (although they are present also in some other evolutionary theories and perhaps in theories in other fields, e.g. economics; cf. Sober 2011). Although we agree with the main lines of Sober's response to F&PP, we think that his attribution of special a priori components to NS is misleading and provides some room for F&PP to continue arguing. Our objective is to show why F&PP's charge is unsound, and to qualify Sober's concerns on aprioricity, so that no alleged specificity of NS enables F&PP to continue arguing. We think that Sober's and other responses to F&PP,[1] although correct, do not put the focus on what in our view is the core of the issue, namely, the guiding-principle-driven explanatory structure of the theory. We argue that analyzing this guiding-principle structure is the best, if not the only way, to show that (*) is false. Our response focuses on these guiding-principle-driven explanations in NS, and, using a Kuhnian-structuralist framework, elaborates a detailed analogy with classical mechanics (CM) and its second law which, we claim, definitively closes the door on F&PP's case. Leaving other (strong, yet irrelevant for our case) differences between NS and CM aside, (*) is not more true of NS than it is

[1] Cf. e.g. Block and Kitcher (2010a, b), Coyne (2010), Dennet (2008), Futuyma (2010), Godfrey-Smith (2008, 2010), Lewontin (2010), Midgley (2010), Okasha (2010), Papineau (2010), Shapin (2010), and Sober (2008b, 2010).

of CM, which F&PP take as a perfectly bona fide empirical theory. The analogy so elaborated enables us to explicate, at the same time:

(i) the way in which NS accounts for the relevant counterfactuals;
(ii) where the unification lies and what its function is; and
(iii) the *apparent* explanatory emptiness of its general unifying guiding principle.

As far as we know, no one in the vast literature generated by this debate has explicated these three features together; all are necessary, and jointly sufficient, for a complete answer to F&PP's challenge.

The comparison with CM is not new. Sober refers to different aspects of NS in comparison with different mechanical laws, for example the law of inertia (Sober 1993, p. 14) or the law of gravitation (Sober 1984, pp. 50–51; Sober 2010); the closest position we find to our own is his reading of the second law as a *consequence law* (Sober 1984, p. 51), an interpretation which, although "combinatorial" like ours, does not elaborate the guiding principle aspects which are essential for our response, and which in any case Sober does not use in his responses to F&PP. Rosenberg (Rosenberg 1985, p. 128; Rosenberg 1994, p. 122) points to the similar role that implicit definitions play in CM and NS, but without distinguishing general from specific principles, which is also essential for our strategy. Dorato (2005, p. 121) makes a very brief, Kuhnianly-inspired remark on the schematic character of Newton's second law and the principle of natural selection, which we take as similar to the position we elaborate in detail in Sects. 4 and 5. Ginnobili (2010) proposes a reconstruction of NS as analogous to CM along the Kuhnian-structuralist lines we defend here. Our response to F&PP elaborates on ideas partially similar to these, and develops them in detail to face the particular problems that F&PP's challenge raises. This, we claim, is a novel and essential contribution to the debate because it enables us to explicate 1–3 above and thereby leave no room for F&PP to continue arguing.

As usual, we understand "natural selection" as the "mechanism" constituted by random variation, heredity, and differential reproduction; and NS as the theory that makes essential use of this mechanism in its explanations. In this regard, NS can be taken as intending to explain three different, progressively inclusive things:

1 why some traits/phenotypes present in a population in an environment and a period of time spread, stabilize, or disappear;
2 how this process plays an essential role in an account of "the tree of life", i.e. how species themselves evolve, appear, and disappear;
3 how analogous "evolutionary" explanations may account for the dynamics of other non-(directly/immediately) biological entities, for example markets, societies, scientific theories, and others.

Although almost everybody agrees that the objectives of NS are 1 and 2 (and some think that 3 is also a legitimate extension), all parties accept that the current debate is *only* about 1. It is not about the overgeneralization 3. It is not even about 2, which has specific problems even if we grant 1. The attack launched by F&PP is addressed at 1 as the very core of the whole adaptive program. Our debate is thus confined to 1.

So confined, two things are worth emphasizing and bearing in mind throughout the discussion: First, 1 is not taken by *anybody* in the debate as saying that natural selection is *the only* explanation of the evolution of phenotypes. Evolutionary biologists accepted, a long time ago, the action of different mechanisms besides artificial and natural selection (such as genetic drift and others, e.g. more recently the so-called horizontal gene transmission—cf. the 2010 special issue of *Biology and Philosophy*). The claim at stake is a more moderate one: natural selection is a *substantive* part of the explanation of the evolution of phenotypes in sufficiently varied and interesting cases. It is this moderate claim what F&PP do deny for, according to them, natural selection does not, and cannot, play *any* part in such explanations: natural selection cannot be *any* part of the story of an acceptable explanation of phenotype evolution. Therefore, according to them NS theory does not have *any* explanatory import. Second, F&PP want to make a case *specifically* against NS (and Skinner's behaviorism). What they claim is that there is something *specifically defective* in NS, that NS contains something substantially different from other bona fide empirical theories, for example CM, that makes it unacceptable as a scientific theory.

Our strategy will then consist in showing that the trait which according to F&PP makes NS *specifically* defective and thus void of *any* explanatory import, is in fact common to all highly unified theories, for example CM, and essential for their unified explanatory role. We take it that F&PP themselves accept that this, if true, would suffice to beat them.

F&PP claim to have two different complaints against NS: "its inability to solve problems about free-riding…was only one of the complaints we've had against natural selection. The other is that natural selection badly underestimates the significance of endogenous factors in the determination of phenotypes" (F&PP, p. 160). The two parts of the book correspond, roughly, to these two complaints. Part one focuses on empirical work. It summarizes empirical data and related theoretical results in biology which, according to them, challenge the very idea of natural selection: the existence of endogenous (Chaps. 2 and 3) and exogenous but not adaptive (Chap. 4) factors acknowledged today as explanatorily relevant; and (Chap. 5) specific mathematical regularities (Fibonacci series, potential laws,…) that also apply to other realms beyond biological phenomena (geology, psychology, or even economics and sociology). Yet, taken together, all that these facts show is, at most, that NS cannot, alone, explain *everything* about phenotype evolution. But, as we said, all the parties had already agreed on this. The data presented in this part are all compatible with the existence of bona fide adaptive explanations of the evolution of traits. Part two, which is mainly conceptual, argues that there are not, and there *cannot be*, such adaptive explanations. Here F&PP present their alleged knock-down argument against NS to the effect that NS has not just little scope but *no explanatory import at all*. The argument is based on NS's (alleged) inability to solve the problem of free-riding because of its (alleged) lack of counterfactual-supporting components. This argument is our target in the rest of this paper. Because it is independent of the empirical data summarized in Part I, we will not discuss these data. If we are right and F&PP's argument in Part II is wrong, and there are sufficiently many and varied bona fide adaptive explanations that share a

substantive explanatory mechanism, it *might* nevertheless be the case that, given the data presented in Part I, such explanations are fewer than was originally expected. Nothing in our contra-argument hinges on this.

Although their dialectics is not always easy to follow, we believe it is fair to take the core of F&PP's strategy here as being constituted by two steps:

(a) there is a valid argument with true premises that concludes that natural selection *cannot* distinguish between selection of/for two coextensive properties; and

(b) when such a distinction is made in biological practice as a bona fide explanation of the stabilization of a trait, biologists *do not use any alleged natural-selection principle* that could be rightly considered to be part of an alleged theory of natural selection: "the beef comes not from adaptationism but from the details of natural history" (F&PP, p. 148).

An essential part of their defense of (i) is their claim there are no laws of natural selection, in the minimal sense of *counterfactual-supporting regularities*. We are going to defend that:

(a′) a correct analysis of the structure of NS and its explanations makes it clear that the argument is flawed, which refutes their "in principle" case;

(b′) there actually are sufficiently rich and varied explanations in biological practice that *do* share principles that can rightly be considered both adaptive and explanatory, and thereby such explanations are not "disconnected natural history" (this refutes their "in fact" case); and

(c′) a core principle involved in NS explanations, what we call its "guiding principle", has a peculiar epistemic status, but the same happens in other bona fide scientific theories like CM.

Therefore, nothing epistemically specific, much less defective, characterizes NS in comparison to CM and other highly unified theories.

2 The Argument Based on Free-Riders

The problem of free-riding is the problem of coextensive properties that accompany each other but such that only one is adaptively efficacious while the other free-rides on the former. For instance, the heart both pumps blood and makes noise, yet it is selected for pumping blood, not for making noise (which is inseparable from pumping blood). The mouth/chin case is another well known example. In these cases we have selection-of both but selection-for just one.[2] F&PP are right in that, if we take into consideration

[2] F&PP trace the phenomenon back to the arch-spandrel example in Gould and Lewontin (1979). Although the arch-spandrel example is a case of trait free-riding (an adaptive trait correlated with another non-adaptive one, like the mouth-chin case, cf. Gould 1977; Lewontin 1978), the discussion moves to the blood-pumping vs noise-making kind of example, which is a case of function or effect free-riding (an organ with an adaptive effect that is accompanied by another non-adaptive effect, cf. Gould and Vrba 1982, for these cases). Though the two cases differ in some important aspects, we, like F&PP and others in this debate, will treat them here as equivalent in the respects relevant to this discussion, because both require appropriate counterfactuals: "if organism O had trait t_1 but not trait t_2, it would have had, in environment E, the same reproductive success" and "if trait/organ t performed action a_1 but not action a_2, individuals with t would have had, in environment E, the same reproductive success".

properties like "making noise", these cases are the norm and not the exception. A theory of natural selection must then ground the distinction between the truth-values of the following, and other analogous, counterfactuals[3]:

(PUMP) if the heart did not pump blood (but kept making noise), animals with hearts would lose fitness/differential reproduction,
(NOISE) if the heart did not make noise (but kept pumping blood) animals with hearts would lose fitness/differential reproduction.

Yet, F&PP contend, the alleged mechanism of natural selection *cannot* account for this distinction. Since without this distinction NS has no explanatory power at all, NS is unable to explain *anything*. Thus, it is not only that NS cannot provide the whole story, it cannot provide any part of the story. NS has *no explanatory import at all*.

But, why is NS unable to ground this distinction? F&PP shoot several bullets at NS's ability to support counterfactuals, the most serious argument being the last one (F&PP, Chap. 7, already presented in Fodor, 2008a, p. 11), which is the one on which, following Sober and others, we are going to focus.[4] In this argument they accept, as is actually accepted by everybody, that counterfactuals can be supported not only by designer minds but also by natural laws, and then they proceed by rejecting that there are laws in NS:

1 If t_1 and t_2 are coextensive traits, the distinction between selection for t_1 and selection for t_2 depends on counterfactuals about which of them *would be* selected in a *possible* world where the *actual* coextension is not valid.
2 The truth makers for such counterfactuals must be either (a) the intentions of the agent that affects the selection, or (b) laws about the relative fitness of having the traits.
3 Not (a) because there is no agent of natural selection.
4 Not (b) because considerations of contextual sensitivity make it unlikely there are laws of relative fitness ("laws of selection").
5 Therefore: Natural selection cannot explain the distribution of phenotypic traits in biological populations

[3] Godfrey-Smith (2008) seems not to agree that the difference of/for needs counterfactual differences: "we might make sense of the distinction between T_1 and T_2 using counterfactuals, but this is not the only way. An inspection of the character of the causal processes themselves may suffice to show that T_1 is causally salient while T_2 is not" (Godfrey-Smith 2008, p. 39). We believe that if T_1 is causally salient and T_2 is not then there is one counterfactual (e.g. PUMP) which "obtains" in nature and other counterfactual (e.g. NOISE) which does not. In this minimal sense we think that the distinction of/for amounts to counterfactual differences. There may be other, stronger sense in which Godfrey-Smith is correct.

[4] The previous three are, we think, clearly flawed, and, more importantly, they do not *specifically* involve NS. If they were valid they would undermine *any* theory that studies non-intentional objects and nevertheless uses counterfactuals (i.e. at least the whole of physics, chemistry, and biology): (1) "How *could* selection be sensitive to the consequences of *counterfactually* removing t but not t' if, *in point of fact, neither t nor t' actually is removed*? The answer is that it couldn't" (F&PP, p. 112); (2) "Actual causal relationships are not sensitive to counterfactual states of affairs: if it wasn't the case that A, then the fact that its being A would have caused its being B does not explain its [actual] being the case that B" (F&PP, p. 114); (3) "only minds are sensitive to distinctions among counterfactuals […] counterfactuals have their effects…only via the meditation of minds […] and Mother Nature has no mind" (F&PP, p. 116).

The argument is formally valid (yet note that only if in premise 4 we clean away the rhetoric qualification "it is very unlikely", for otherwise the conclusion should be just the weaker "it is very unlikely that NS can explain…", and they want to argue for the stronger conclusion 5; in many other passages they openly say that there are no laws of selection, e.g. Fodor 2008a, pp. 8–11; F&PP, Chap. 7; Afterwords, 2nd ed., pp. 183–185). We, with Sober and others, agree with premises 1–3 (at least for the sake of the present dialectics).[5] Premise 4 is the problem. True, *if there were no laws*, that is, *if there were no counterfactual-supporting* generalizations/principles/regularities/models of selection, there would be no counterfactual differences between selection of and selection for. But, is it the case that there are no laws of selection, *in this minimal counterfactual-supporting sense*? It must be emphasized that this minimal sense is the one that is relevant to the debate. There is consensus among philosophers of biology that there are no laws in NS in a stronger sense, i.e. laws similar to those in physics (maximally universal, non-context-dependent, etc.). But there is no consensus at all (quite the contrary, as this debate and the replies to F&PP show) that there are no laws in this relevant, minimal sense, i.e. that there are no adaptive counterfactual-supporting regularities.[6] All one needs to reply to F&PP, then, is to show that there are non-accidental (i.e. counterfactual supporting) general facts (though maybe each one applying to quite restricted domains) that can be correctly considered "adaptive".

We have thus reached the very core of F&PP's case. Everything depends on whether there are laws of selection in this relevant, minimal sense of *counterfactual-supporting general facts* (no matter how domain-restricted each one is) with a substantive, i.e. non-trivial, adaptive nature in common. Why do F&PP say that there are no such adaptive modal general facts? The main reason they give against the existence of laws of natural selection is the extreme locality and context-sensitivity of the value of traits for fitness. Whether a trait favors fitness or not is always strongly dependent on local conditions: "who wins a t_1 versus t_2 competition is massively context sensitive … Whether a trait militates for creature's

[5] Somebody could disagree regarding 1, e.g. no need of counterfactuals for the distinction of/for (cf. Godfrey-Smith 2010, last note); or regarding 2, e.g. if one buys counterfactual-supporting singular causation without laws. Although we believe that the difference of/for is related to counterfactual differences, and that counterfactual differences are grounded in nomic generalizations in the minimal sense of "counterfactual-supporting generalizations", we do not want to enter into these issues now. In any case, we concede F&PP 1 and 2 so that, if our response works with these concessions, it would also work without them.

[6] Part of the huge debate on laws in biology depends on what one understands by "law" (cf. e.g. Brandon 1997, Elgin 2003, Sober 1997, Rosenberg 1994). Here we make a minimal reading, i.e. principles or regularities with *counterfactual force*, which is all we need here. In the F&PP debate, almost all parties agree that the distinction selection of/for is causal, because properties selected for are causally efficacious for differential reproduction but free riders are not, and also that this causal distinction needs counterfactual differences. To avoid opening another front, we will not emphasize this causal reading of the counterfactual differences [with which we agree in *this* case, cf. Martínez and Moya (2011), for a recent discussion]; all we need for the point at stake here is the acceptance of regularities with counterfactual force in NS, irrespective of whether this modal force is, in turn, explained in causal or other terms. This minimal characterization of laws as counterfactual-supporting facts is similar to the one defended in Dorato (2012), and it is also compatible with some current proposals about laws in biology in particular, such as the "paradigmatic" (Carrier 1995) and "pragmatic" (Mitchell 1997) ones.

fitness…always depends on what else is going on in the neighborhood" (F&PP, pp. 123–124). That is, being big, or small, or green, or gray is not good or bad for fitness *per se*; it depends on the conditions of the ecological niche.

Well, of course whether a trait is beneficial for differential reproduction depends on the environment. This is the whole point of NS: it is only relative to "environmental pressures" that traits facilitate/impede reproduction-beneficial functions, for example eating, escaping predators, mating, and others.[7] Yet, all this fact implies is not that there are no adaptive regularities with counterfactual force, but just that NS laws are highly relational and essentially dependent on the variable "environment" and therefore applicable to restricted domains. This, though, does not make NS *specifically problematic*. First, it is not intrinsically problematic even if NS were the only case, because context sensitivity and local domain restriction do not cancel the modal force of nomic regularities. Second, NS is not the only case, we find relational laws/counterfactual-supporting regularities in other bona fide scientific theories in different fields (e.g. geology, meteorology, cosmology, or economics).

On some occasions F&PP seem to appeal to other objection against NS laws, namely that explanations of traits are sometimes, or often, extremely complex and multilevel: the explanation of traits involves many things and often conceptual/ontological machinery from different levels (cf. their analogy with the Napoleon losing Waterloo example, F&PP, p. 133; cf also their "Afterword" to the 2nd ed., pp. 183–185). Again, when true, all this shows is that in these multilevel cases we have explanations that involve inter-theoretical relationships. And again, these cases by no means show that in these explanations there is no use of *any* NS principle/law at all, which is the point at stake here.

All the reasons that F&PP offer against the existence of laws of selection are thus compatible with the existence of (context-relative and domain-restricted) adaptive counterfactual-supporting regularities. To clarify F&PP's case, it is worth emphasizing that their denial of laws of selection, and consequently of *adaptive* explanations, does not imply that according to them the evolution of traits is not explainable. F&PP's case is *not* that trait evolution cannot be explained. They insist that traits are explained in bona fide biological practice, and counterfactually (and indeed, according to them, causally) explained. Their point is, rather, that such explanations are not provided by an alleged unified NS theory but by a kind of natural history, i.e., a collection of particular *independent* explanations that have *nothing* (non-truistic) in common: there is no mechanism common to different trait-explanations that provides the kind of unifying explanatory resource that deserves the label "theory":

"When [explanations of the evolution of heritable traits] work it's because they provide plausible historical narratives, not because they cite covering laws. […] adaptationism does not articulate the mechanisms of the selection of heritable phenotypic traits; it couldn't because *there aren't any mechanisms of the selection*

[7] That natural selection always operates through environmental pressures, what Darwin labeled "the struggle of life", is correctly emphasized by Lennox and Bradley (1994) against more liberal readings, e.g. Lewontin's, which include in natural selection any (non-random drift) cause of differential reproduction.

of heritable phenotypic traits (as such)" (F&PP, p. 136, our emphasis; note that no "only very unlikely" caution is included here).

"Natural selection theory is often said to provide a mechanism for the evolution of phenotypes. That, however, is precisely what it doesn't do. What explains why there are the phenotypes there are is not natural selection but natural history […] the beef comes not from adaptationism but from the details of natural history" (F&PP, p. 148).

The book, however, does not contain any detailed reconstruction of some such natural history-like explanation of trait-evolution. So we can not assess whether what they consider acceptable trait-evolution explanations contain adaptive principles playing an essential role. In the absence of this, let us turn to the practice in biology and review some simple, paradigmatic examples, and analyze in full detail another less simple one, and see whether there is some explanatory component in common that is of an adaptive nature and carries modal force.

In the next two sections we will show that:

(i) there are (enough, paradigmatic, and varied) explanations of trait-evolution that do have something nomic, i.e. counterfactual-supporting, in common;
(ii) this common explanatory mechanism can rightly be considered to be of an "adaptive" nature;
(iii) the sense in which all NS explanations share a common explanatory resource is *exactly analogous* to the sense in which explanations in other bona fide empirical theories, for example classical mechanics, share an explanatory resource;
(iv) this common NS explanatory resource has a subtle epistemic status, but this is not specific to NS, for the same happens with analogous principles in other theories that have never been challenged (and F&PP do not want to challenge).

We take it that 1–4 are sufficient against F&PP's case contra NS. Note that 3 and 4 are essential for our strategy: they block any possibility of continuing arguing against NS on the basis of an alleged epistemic failure *specific* to NS.

3 NS—Explanations of Phenotypic Evolution and the General NS Principle

We are familiar with the kind of explanations NS provides for some simple examples, for example the long neck of giraffes in their habitat, the colorful tail of peacocks, or the change in color of moths in industrial Sheffield. In these and other standard examples the explanation of the evolution of a trait relies on its utility, in a given context, for improving the performance of a specific behavior or function[8] which is beneficial for reproduction. In the case of giraffes, the length of their neck improves, in the context, the performance of a function, food supply, which is good for differential reproduction. In the still simple but more interesting case of black

[8] Here, and henceforth, we use "function" in a neutral way, simply as a common label for things like eating, escaping predators, attracting sexual partners, etc. We do not want to engage in the debate regarding the use of a technical notion of function in biology in general and in evolutionary biology in particular. Nothing in what follows hinges on which position one sides with in the function-debate.

moths, after the change in the environment the blackness improves the performance of a function, escaping predators, which is good for differential reproduction. And so on. Here "function/behavior good for differential reproduction" means, roughly, that groups of individuals that perform the function/behavior better tend, *ceteris paribus*, to increase their relative ratio (without necessarily becoming the majority) in the environment compared with others that perform the function less efficiently.

Let us now reconstruct a less simple example of a NS explanation, namely Fisher's model/explanation of the (approx.) 50–50 offspring sex ratio, and check whether this explanation also makes use of an adaptive principle, although this use may not be immediately apparent. We use here the simplified version given by Sober,[9] for it suffices for our present concerns and is also relevant to our discussion of Sober's position on NS a priori models in the next section.

"Fisher's model considers three generations—parents produce offspring who then produce grand offspring. What mix of sons and daughters should a parent produce if she is to maximize the number of grand-offspring she has? If there are N individuals in the grand offspring generation, and if the offspring generation contains m males and f females, then the average son has N/m offspring and the average daughter has N/f offspring. A mother thereby gains a benefit of N/m from each of her sons and a benefit of N/f from each of her daughters—these benefits being the number of grand offspring they give her. So individuals in the offspring generation who are in the minority sex on average have more offspring. Hence, the best strategy for a mother is to produce offspring solely of the minority sex. On the other hand, if the sex ratio in the offspring generation is 1:1, a mother cannot do better than the other mothers in the population by having an uneven mix of sons and daughters." (Sober 1993, p. 16)

Let us analyze this model/explanation [here, also, we closely follow Sober (1993, p. 17)].

- There are three generations: G_1, the mother; G_2, offspring; G_3, grand offspring. The mother has m sons and f daughters (G_2).
- Upon random mating, the M sons and the F daughters generate a total of N grand offspring (G_3). The average son then has N/m offspring, and the average daughter has N/f.
- E is the quantity of energy that the mother has to invest in offspring, and p is the percentage invested in sons, so pE is the energy invested in male offspring and $(1 - p)E$ is the total energy invested in female offspring. p and $(1 - p)$ indicate then the energy distribution ratio for sons/daughters.
- Let the energy costs for each son and daughter be, respectively, C_m and C_f. Then the number of sons equals $p \cdot E/C_m$ (total energy at the disposal for sons over the energy cost of each son) and the number of daughters equals $(1 - p) \cdot E/C_f$.
- Let the benefit received by the mother for each son and daughter be, respectively, B_m and B_f. Then mother's total benefit $= B_m \cdot [p \cdot E/C_m] + B_f \cdot [(1 - p) \cdot E/C_f]$

[9] Cf. also Sober (2010). For the historical version cf. Fisher (1930); for Fisher's sources, such as Darwin (1871) and Düsing (1884), cf. Edwards (1998, 2000).

- A different energy distribution p' does better than p if:

$$B_m \cdot [p^* \cdot E/C_m] + B_f \cdot [(1 - p^*) \cdot E/C_f] > B_m \cdot [p \cdot E/C_m] + B_f[(1 - p) \cdot E/C_f]$$

  which simplifies to: $[(B_m/C_m) - (B_f/C_f)] \cdot (p^* - p) > 0$

- *Because doing better is doing reproductively better* (more on this important condition later), then $B_m = N/M$ and $B_f = N/F$. Substituting we obtain that p^* does better than p if:

$$[(N/m \cdot C_m) - (N/f \cdot C_f)] \cdot (p^* - p) > 0$$

- Finally, at equal costs, i.e. when $C_m = C_f$, and simplifying, this reduces to:

$$(1/m - 1/f) \cdot (p^* - p) > 0$$

Which means that p is optimum when $f = m$, i.e. when there is a tendency to equal sex ratio, which is what we wanted to explain.

Thus, upon random mating and equal costs of sons/daughters, the birth sex ratio approximates 50–50. Up to here, this is more or less standard analysis of Fisher's model. In order to make our point, let us schematize this explanation, emphasizing its main components:

SexRat : *IF* (1) *random mating*
(2) *cost per son* = *cost per daughter*
(3) *benefit provided per son/daughter is its average reproductive contribution*
THEN (4) *birth sex ratio approximates* 50 − 50

Different conditions may give rise to different outcomes (cf. Hamilton 1967). For instance, if we change the initial condition (1) to (1′) brother/sister mating, we obtain, (4′), a female sex ratio bias. If (2) is substituted by (2″) "cost is proportional to mortality" then we would obtain (4″) a bias in the birth sex ratio in favor of the mortality ratio. Other substitutions of (1) and/or (2) would give rise to other explananda.

What about (3)? (3) remains unchanged in all sex ratio explanations. This, we claim, is the crucial NS component of sex ratio explanations: (3) is the specific version that a general NS principle takes for these specific explananda. Although (3) is sometimes not explicitly mentioned (e.g. Sober 2008a, p. 45), and other times only mentioned in passing (e.g. Sober 1993, p. 17), it is crucial, for it is because of (3) that *SexRat* is an adaptive model/explanation of NS. It is because (3) is a particular application to sex ratio cases of a general NS principle that Fisher's model counts as a NS explanation. Very roughly, the NS general principle may be taken as stating something like[10]:

NSGP: Phenotypic (heritable) trait t increases (/decreases) its chances for spreading and stabilizing in environment E if it facilitates (/impedes) the

[10] This version suffices for our present concerns. For a detailed discussion of a NS general principle, see, e.g., Sober (1984, 1993), Brandon (1982, 1996), Kitcher (1993, § 2.4), Rosenberg and McShea (2008), Ginnobili (2010); we will discuss NSGP in more detail in the next section.

performance of a function or behavior *B* that enhances reproductive success

In short: *environmentally adaptive (heritable) traits increase their chances of spreading and stabilizing* ("adaptive" summarizes the idea that the trait facilitates a function that improves differential reproduction). This NSGP, together with a specification of the function *B* that a given trait *t* improves in *E*, explains that such *t* spreads in *E*. Note that, although, because of the existence of other non-adaptive mechanisms of trait-change (e.g. genetic drift), the conditional is just "if" and not "if and only if", the aim is that this adaptive mechanism applies to sufficiently many, varied and interesting cases of trait-evolution. Yet there is no need to include this as a claim in the formulation of the law; it is "stated" not in the letter of the principle but in the *practice of its application*, i.e. successfully applying this general principle to more and more varied and interesting trait phenomena (more on this crucial point later).

In our example, 3 includes the specific application of NSGP for the specific case in which the trait *t* is *having a particular m/f offspring ratio*, and the function/behavior *B* is just *fecundity* (an obvious reproduction-"enhancer"). 3 states that, in this context, *B*'s benefit for reproductive success increases with the average number of offspring. In other explanations, NSGP adopts other specific forms specifying the reproductively beneficial function that, in the given environment, is performed better with than without the trait:

- in the giraffe case, the length of the neck (*t*) through improving food supply (*B*)
- in the black moth case, their blackness (*t*) through improving escape from predators (*B*)
- in the peacock case, the colorful tail (*t*), through improving sexual attraction (*B*).

Other more complex examples would show other less apparent functions/behavior that enhance better reproductive success with than without the trait in point. For example, a detailed reconstruction of the Lotka–Volterra predator–pray model[11] would show that it makes implicit use of a specific application of NSGP, namely, that the predating ratio stabilizes when it maximizes differential reproduction. And the same applies to other more complex explanations involving two or more reproductively beneficial functions, e.g. the explanation for color patterns in *poeciliid* fishes (cf. Endler 1983). Space constraints do not enable us to reconstruct these and other cases in detail, but it can be shown that all have a structure in which the explanans includes, together with initial and boundary conditions, another condition with counterfactual force specifying the environmentally beneficial function relevant to the particular trait-explanandum in point.

Thus, *every* explanatory NS model/conditional specifies, in its antecedent, initial and boundary conditions, *and* another component which is the specific application of NSGP for the particular explanandum stating the function which is beneficial for differential reproduction and does better, in the relevant context, with than without

[11] For our present exemplification concerns, it does not matter that this model's simplicity makes it not very successful when applied to real populations. The analysis of other more complex, and empirically better suited models (e.g. the Rosenzweig–MacArthur model) would lead to a similar structure.

the trait: eating, escaping predators, attracting sexual partners, reproducing with a specific ratio, predating in a specific proportion. It is by means of these concrete functions as specifications of the NSGP general mechanism that we can explain the evolution of the trait in the specific environmental context. So, here we have the hidden "mechanism" that F&PP deny: *NSGP is what is common to/presupposed in all natural selection explanations*, and what makes NS a bona fide unified explanatory theory. Every NS explanation makes use of a specific application of NSGP to the specific explanandum in point.

Note that F&PP cannot object that NSGP is not a "mechanism" in the *mechanistic* sense. If "mechanism" is so narrowly understood, then classical mechanics would not qualify as a bona fide explanatory theory either, for not gravity, nor elasticity, nor friction, … are present in all mechanical explanations, and the only "explanatory resort" common to all mechanical explanations is Newton's second law, which obviously is not a mechanism in the narrow, mechanistic sense. It is our claim that the same happens in NS: there is a common explanatory mechanism, though not a mechanism "mechanistically" understood. What F&PP might, and do, complain about is that NSGP is empty, truistic, or trivial and therefore explanatorily useless and scientifically unacceptable. This is their final bullet, to block which it is worth comparing NSGP with Newton's second law and its role in CM. We conclude this section by clarifying F&PP's charge of emptiness (and its relation with Sober's thesis on the apriority of adaptive models). In the next sections we elaborate our comparison with CM and extract the consequences that allow us to definitively close F&PP's case against NS.

After arguing against the existence of laws of selection, F&PP take into consideration a possible response by the friend of NS offering a general nomological adaptive principle, actually very similar to our NSGP:

> How about treating the theory of natural selection as a theory schema, perhaps along the following lines: adaptationism makes the empirical claim that, for each phenotypic trait (or, for each phenotypic trait that is an adaptation) there is an ecological problem of which the trait selected-for was the solution. *Adaptationism per se* does not say, in any particular case, either which phenotypic trait was selected-for or which problem it was selected-for solving. But it does say that, in any bona fide case of adaptation, there always is such a trait and such a problem. This claim constitutes the basic empirical commitment of the theory. (F&PP, p. 131, our emphasis).

F&PP seem to concede here that a NS unifying principle could be enough to block their case. Yet, they immediately reply that such a principle does not work, because it is "merely definitional", or empty, hence explanatorily useless, hence scientifically unacceptable:

> We think that's fine if, *but only if*, "adaptation", "selection-for", etc., are independently defined, so that (for example) "adaptations are traits that are selected for" is a contingent truth rather than a definition. (F&PP, p. 131)

And they object in a related vein in several places, e.g.:

> the theory of natural selection reduces to a *banal truth*: "If a kind of creature flourishes in a kind of situation, then there must be something about such creatures (or such situations, or both) in virtue of which it does so." Well, of course there must; even a creationist could agree with that. (F&PP, p. 137, our emphasis)
>
> "If, in the ecology they occupy, birds with wings are better off than birds without them, there must be something about the birds, or about the ecology, or about the two together, in virtue of which birds with wings are better off in that ecology than birds without them. That's just *a routine application of the principle of sufficient reason*." (F&PP, p. 148, our emphasis)

In the "Afterword" to the 2nd edition they insist that the only thing that different "adaptive" explanations may have in common is an explanatorily empty, and therefore useless, "principle of sufficient reason", and that "[a]ll there is, is natural history" (F&PP, 2nd ed, pp. 186–187, i.e. *disconnected* natural history). Thus, according to F&PP the only allegedly adaptive mechanism that could be proposed as common to different explanations of traits evolution is merely definitional, a banal truth or a routine application of the principle of sufficient reason. This is then their ultimate main objection: the only candidate for a counterfactual-supporting adaptive regularity common to different traits explanations is actually merely definitional, truistic or empty (cf Lewontin 2010, for a similar diagnosis of their case), it thereby has no explanatory import and is unacceptable as the unifying factor of a bona fide empirical theory.

The charge of emptiness behind F&PP's denial of adaptive nomological regularities echoes previous similar concerns in the philosophy of biology literature.[12] And here is where Sober comes in: although he defends NS from F&PP's attack, in his discussion with Fodor and other works he claims that, compared with, e.g., CM, NS is epistemically peculiar in that its models are a priori (a trait shared with other theories in the field of evolution and, perhaps, economics also; cf. Sober 2011). This claim, despite Sober's protests, is taken by Fodor in his support. Although Sober insists that the a priori components do not make NS explanatorily empty, Fodor argues that this peculiarity, correctly understood, makes NS trivial, hence explanatorily useless and scientifically unacceptable (Sober and Fodor 2010). Everything, then, finally depends on whether NS contains certain components that are "empty" in a specific defective sense compared with other bona fide theories such as CM. If it does, F&PP win; if it does not, they lose. After clarifying what Sober takes to be a priori, we will show in the next section that, in the sense in which one can say that NS contains empirically "empty" or a priori components, the same applies to other bona fide theories, including CM. This, we claim, definitively eliminates the small space that, unwillingly, Sober leaves to F&PP to continue arguing.

[12] The first reference is probably Mivart (1898, p. 272). The charge has been recurrent, and made even by prominent philosophers of science, for example Popper (1963, "as tautological"; 1972, 1976, as "unfalsifiable metaphysical program"). Cf., e.g., Sober (1984, Chap. 2), Rosenberg (1985, Chap. 5.2), and Lennox (2001) for discussion and references on different aspects of this issue.

What does Sober take to be a priori in NS? What he *initially* (cf. Sober 1993, 2008a, b, 2011) takes to be a priori is contained in the following quote that takes Fisher's model as an example:

> I think that Fisher's theory is a mathematical truth; *the consequent follows mathematically from the antecedent* once everything is stated carefully. (Sober 2008a, p. 45, our emphasis; cf. also Sober 2011)

We take this to claim that conditionals like the above *SexRat: IF* 1, 2, and 3 *THEN* 4 are a priori. Thus, the first interpretation of Sober's claim is ($Sober_1$): "In NS explanatory conditionals, the antecedent a priori implies the consequent". $Sober_1$ is true, as Fisher's model, Hardy–Weinberg's "law", Lotka–Volterra's model and many other a priori conditionals exemplify. Yet, the immediate reaction here is to contend that this is not specific to NS compared with, say, CM: given certain initial conditions, together with (a specific application of) the second law and certain other empirical assumptions, we a priori infer, for instance, Galileo's free fall regularities, or Kepler's planetary movement, or the pendulum trajectory, or…. Therefore we have the same kind of a priori conditionals in CM and, in general, in any explanatory theory. Recently, Sober has accepted this objection and has qualified what he considers to be specifically a priori in NS:

> The objection I want to consider claims that the pattern I have described in evolutionary theory applies trivially to all scientific theories. For example, consider Newtonian mechanics, which is generally taken to be an empirical theory. Because this theory (*T*) allows one to deduce a prediction (*P*) from a specification of initial and boundary conditions (*IB*), we can construct a conditional of the form "If *IB* and *T*, then *P*" that is a priori true. The objection is correct that *T*'s being empirical does not prevent this construction from being carried out. But the situation in evolutionary theory is different. The models I have described are a priori; they have the form "if *IB*, then *P*"; *there is no empirical law in the antecedent*. (Sober 2011, our emphasis).

Thus, what Sober now takes as specifically a priori in NS is this: ($Sober_2$) "NS explanatory conditionals are of the form "if *IB*, then *P*" containing no empirical law in the antecedent" (Sober takes "non-empirical" and "a priori" as synonymous in this context). This new sense has two possible readings. First reading: "the antecedent contains no laws, and a fortiori no empirical laws, i.e. the antecedent contains *only I* and *B*". Yet upon this reading $Sober_2$ cannot be true. We have seen that the conditional that summarizes Fisher's model makes essential use of (3) in its antecedent, and (3) is neither in *I* nor in *B*. According to us, (3) is a substantive NS claim as a particular application of NSGP. But bracketing our specific proposal about (3) for a moment, it is uncontroversial that (3), which is common to all sex-ratio explanations, is essential for the conclusion obtained and is neither an initial nor a boundary condition. So the antecedent *cannot* contain only *I* and *B*. Note that if it contained only *I* and *B*, then F&PP would actually have a strong point: NS explanations would not have *anything in common,* because *I* and *B* change from one explanation to another. Sober, though, contends many times that NS explanations *do* have something in common. So, this first reading is neither sustainable nor

consistent with Sober's other claims. The only possible alternative reading is that the antecedent contains no *empirical* laws but contains an extra component other than *I* and *B* which, although explanatorily necessary, is not empirical but a priori. If this is the case, then, what all NS explanations have in common would be a priori. Hence, once Sober abandons Sober_1 as a characterization of the a priori components of NS, and *if* one considers that aprioricity implies empirical emptiness, the only acceptable reading of Sober_2 is to attribute to NS a peculiarity related to F&PP's emptiness objection. This might explain why, despite Sober's protests and insistence that aprioricity does not imply explanatory uselessness, Fodor finds some air in Sober's comments on the aprioricity of NS models. Thus, to definitely block F&PP's case, it is crucial to clarify whether, or in what sense, NS explanatory models make use of epistemically peculiar (empirically empty?, a priori?) theoretical tools, what the role of NSGP in NS explanations is, and whether NS is alone in this practice or in good company.

4 Natural Selection, Classical Mechanics, and Guiding Principle-Based Explanatory Theories

Recall F&PP's description of what a friend of NS could say quoted above: "*Adaptationism per se* does not say, in any particular case, either which phenotypic trait was selected-for or which problem it was selected-for solving. But it does say that, in any bona fide case of adaptation, there always is such a trait and such a problem" (our emphasis). We take F&PP's "adaptationism per se" to be pretty close to our NSGP, a principle telling us that in every specific case we have to find out the function that, in the specific environment, facilitates differential reproduction better with the trait than without it. Is this "merely definitional", as they complain? Or is it rather a perfectly acceptable explanatory unifying principle? It is our claim that NSGP is no more definitional than Newton's second law is. Although many decades ago there was some discussion as to whether the second law is a definition, for a long time now there has been a robust consensus that at least it is not an *explicit* definition, i.e. a dispensable one. It can be taken as an *implicit* definition of "mass" and "force", in the Ramsey–Carnap–Lewis sense in which the axioms are implicit definitions of theoretical terms (Díez 2005), which makes these terms *T*-theoretical in a theory *T*, i.e. *introduced* by the theory, in contrast to other terms that the theory borrows from other theories (Balzer et al. 1987). But, leaving some subtleties aside, all parties agree that in this respect the axioms involved in these implicit definitions (either all or only some special central ones in the theory) have empirical content and are essential for the explanatory power of the theory. Although it is difficult to fully clarify what this empirical content is, there is no discussion as to whether these implicit definitions are present in theories like CM, thermodynamics and others. Thus, again, if NSGP is "definitional" in *this* implicit-definition sense, then no *special* shame for NS, no more than for CM, thermodynamics and other never questioned theories. Let us see this in more detail by comparing NS and CM in this regard.

What we have identified in NS is a case of what has been called "guiding principles", already identified in theories like CM (Newton's second law), thermodynamics (Gibbs' equation), classical genetics (Lorenzano 2000) and other highly unified theories (Moulines 1984). As a guiding principle Newton's second law $\Sigma F = m \cdot a$ can be read as follows:

CMGP: When mass particles change their motion, look for forces that when added up account for the change in motion

In a similar vein, as a guiding principle NSGP could be read, roughly, as:

NSGP: When a trait changes in a population, look for its adaptive force, i.e. for some function or behavior that, in the given environment, enhances reproductive success better with than without the trait

Analogously to the way in which different particle-trajectories are explained appealing to different kinds of force (gravitational, frictional, elastic, etc.), different trait-evolutions are explained by appealing to different reproductive/adaptive needs: food supply, escaping predators, attracting sexual partners, fighting sexual rivals, having a specific proportion of male–female offspring, etc. To explain a specific particle-trajectory/trait-evolution is to find a specific version of the general nomological dependence. To consider a particle trajectory as a mechanical trajectory, and thus to accept it as a CM explanandum, is to trust that we are going to find specific forces, masses, and a specific form of CMGP that imply the trajectory. And not all trajectories are treated in this way: for instance, a pen in someone's fingers doing funny things at her will is not (at least not until we reduce psychology to CM[13]). More interestingly, light beams (in classical physics) were treated as a CM explanandum for a while, but ceased to be so considered later. Analogously, to consider a trait-evolution as a natural selection phenomenon, and thus accept it as a NS explanandum, is to trust that we are going to find a specific adaptive problem in the environment and a function that affects differential reproduction differently when performed with the trait than without it. And not every trait-trajectory is treated in this way. If we cut the tail of every male lion at birth for generations, this trait-trajectory would not count as a possible explanandum for NS. More interestingly, changes in traits due to random drift or horizontal transmission do not count as explananda for natural selection.[14]

[13] Of course CM aims to (and in fact does) explain the pen's trajectory in terms of the hand's trajectory (and some auxiliary assumptions). But it does not aim to explain the intentional movement of the hand/arm itself, or to be more precise, to explain the neural event (which bio-mechanically explains the movement of the hand/arm) in terms of subject's intentions. Whether materialism is true is an open issue; and even if, as we think, intentionality is ontologically reducible to, or at least supervenient on, physicality, it does not follow that intentional explanations reduce to physical explanations. Thus to exclude intentional movements from CM's explanatory scope does not require granting the existence of "magic powers" or the like (we thank an anonymous referee for asking for a clarification on this point).

[14] Trait-evolutions with direct human breeding intervention count as selection explananda, but of *artificial* selection rather than natural selection; one could include them in NS if, as we think we should, one includes intentional human intervention of this kind as a possible "natural" selective pressure among others, but we cannot discuss this independent issue here.

This point is worth emphasizing. Does the fact that CM does not aim to apply to particle-trajectories like the funny movement of a pen in someone's hand need to be said "explicitly in the laws", e.g. in the second law? No! CM "says" that a specific trajectory is mechanical by explaining it via a CM specific law. CM practitioners consider a specific trajectory as a candidate for a mechanical one by trying to find a specific CM law that explains it. There is no need to explicitly say in a law "a mechanical trajectory is…". A mechanical trajectory is any trajectory successfully explained by a particular application of CMGP. It is only in this implicit sense that the second law "says" what a mechanical trajectory is. The same applies to NS: there is no need to explicitly say "a naturally selected trait evolution is…". All we need is, when one considers that a specific trait evolution is naturally selected, to try to find a function that in the relevant environment affects differential reproduction differently with the trait than without it. And the belief that a vast majority of trait-evolutions are (at least partially) NS-explainable is implicitly shown by NS practice explaining more and more trait evolutions by specific applications of NSGP. Note that the fact that guiding principles are sometimes (as in NS and classical genetics, and contrary to what happens in CM and thermodynamics) "implicit", i.e. implicitly used or presupposed in explanations yet not explicitly formulated, does not imply that such principles are *true by definition* (so we are not guilty of the definitional strategy that F&PP attribute to Godfrey-Smith and Sober, and rejoin, in the "Afterword" to the 2nd ed., pp. 179–180), nor that there cannot be phenomena to which they are not designed to apply.

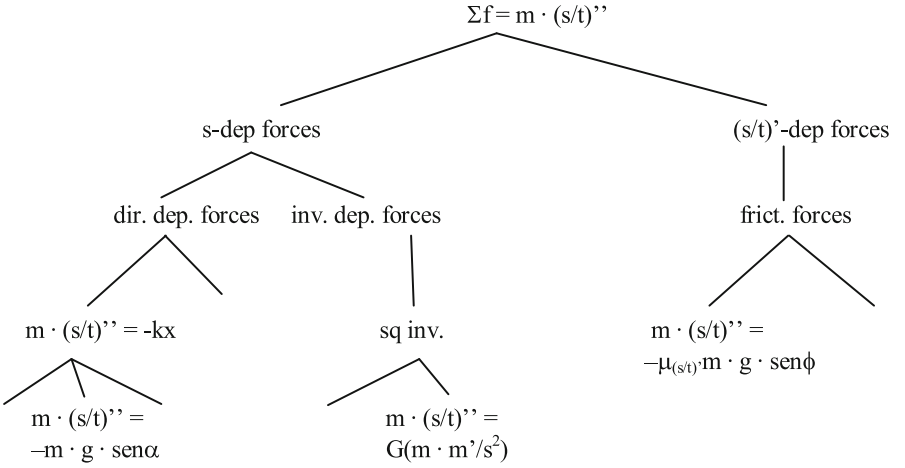
The picture that emerges here is very close to what Kuhn says about disciplinary matrices, nomic generalizations, and exemplars/applications. According to Kuhn, in highly unified theories like CM there are some generalizations that are not "specific laws" but rather "schemes" which take specific forms for specific problems/applications:

> generalizations [like $f = ma$…] are not so much generalizations as generalization-sketches, schematic forms whose detailed symbolic expression varies from one application to the next. For the problem of free fall, $f = ma$ becomes $mg = md^2s/dt^2$. For the simple pendulum, it becomes $mg\ sen\alpha = -md^2s/dt^2$. For coupled harmonic oscillators it becomes two equations, the first of which may be written $m_1d^2s_1/dt^2 + k_1s_1 = k_2(d + s_2 - s_1)$. More interesting mechanical problems, for example the motion of a gyroscope, would display still greater disparity between $f = ma$ and the actual symbolic generalization to which logic and mathematics are applied. (Kuhn 1970, p. 465)

It is striking how much this passage resembles F&PP's quote above discussing an analogous interpretation of NS (and rejecting it as defective in a sense in which theories such as CM are not!). This Kuhnian idea has been elaborated in detail by Sneedian structuralism with the notions of *specialization* and *theory-net*, and has been applied to several sufficiently robust and unified theories.[15] For instance, the

[15] For a standard and totally precise exposition, and application to CM, thermodynamics and other theories cf. Balzer et al. (1987). For a more informal presentation, see Moulines (2002). The program originates in Sneed (1971), and Kuhn (1976) acknowledges that it is the approach that captures his proposal best.

CM theory-net looks (at a certain historical moment) like follows (only some terminal nodes are shown here, and in a simplified version, but this suffices for our present exemplification concerns):



The net has CMGP, Newton's second law, as the top unifying nomic component, and opens down different branches for different phenomena/explananda, branches that can be reconstructed in different steps: first, space-dependent forces versus velocity-dependent ones; then the space-dependent branch specializes into direct and indirect space-dependent; direct space-dependent branch specializes in turn into linear negative space-dependent and…; inverse space-dependent branch specializes into square inverse and…; at the bottom of every branch we have a totally specific law that is the version of the guiding principle for a specific phenomenon: pendula, gravitation, inclined planes, etc. (Kuhn's "detailed symbolic expressions"). Note that the top–bottom relationship is not one of implication or derivation, but of *specialization* in the structuralist sense (Balzer et al. 1987, Chap. IV): bottom laws are specific versions of top ones, i.e. they specify some functional dependences that are left partially open in the laws above in the branch.

This brief sketch suffices to clarify the peculiar epistemic status of guiding principles like CMGP and NSGP. What is crucial, and this was already emphasized by Kuhn, is that these top general principles cannot be empirically tested "in isolation": *they can be tested, and eventually falsified, only through one of its specific versions for a specific phenomenon* (that is why, after a failed prediction, one *may* change the general principle—Kuhn's revolutions—but can also try to fix the anomaly by modifying only the specific law—Kuhn's normal science). In this sense guiding principles are "programmatic" or heuristic: they tell us the kind of things we should look for when we want to explain a specific phenomenon. But taken in isolation, without their specializations, they say empirically very little. They can be considered, when considered alone, "empirically non-restrict".[16]

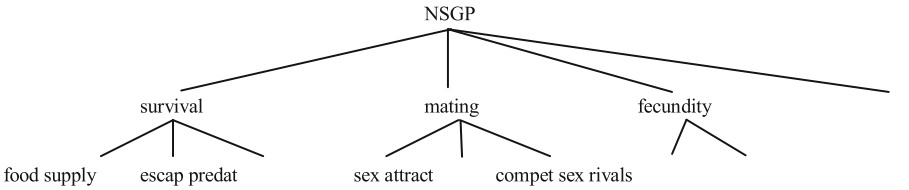
[16] This term is Moulines' (1984); Kuhn uses "quasi analytic" (Kuhn 1976), Díez (2002) "concept-constitutive" and Lorenzano (2006) "synthetic a priori" (also used in Kuhn 1990). If one thinks that

For instance, in the case of Newton's second law, it is always *formally* possible to fit *any* trajectory, no matter how rare or artificial, *if* one accepts *any* kind of function, no matter how artificial or crazy; such crazy functions, though, are not acceptable as "mechanical" specializations of CMGP for they are not relevantly similar to other, paradigmatic ones (this is part of the essential pragmatic role of Kuhnian paradigms). The epistemic status of such guiding principles, then, is subtle. It is not well captured by simply saying that they are "empty": though they (taken alone) are empirically non-restrict, they are not dispensable definitions, truisms or simple applications of the principle of sufficient reason. Once they are correctly understood, one can qualify them with the term one likes best, but this does not make them empty in the same way as "bachelors are unmarried" or "acceleration is the second derivative of space over time" are. One might perfectly well call them simply "a priori", but this does not make them analogous to "every closed trilateral polygon has three angles", or to a priori implications like *SexRat*.

It is our claim that NS is analogous to CM in these respects, which are the relevant respects in the current debate. NSGP is common to all NS-specific explanations in that it specializes down into different specific versions for different adaptive phenomena, every specialization specifying the kind of function beneficial for differential reproduction that is differently performed in the environment with than without the trait: selection by survival, by partner mating, by fecundity, etc.; the survival branch specializes, in turn, into food supply, escaping predators, etc.; the partner mating branch in turn into sexual attraction, rival competition, etc.[17] NS theory-net thus has a structure similar to CM, in which every node group a family of functions of the same "adaptive kind" at a specific level [we represent here only part of the net, which suffices for the purposes of exemplification's sake; cf. Ginnobili (2010), for a more complete, and slightly different, proposal within the same structuralist framework]

Footnote 16 continued
guiding principles are constitutive of the content of theoretical concepts, then such principles are obviously connected with the analytic/synthetic distinction. Important as this issue is, it is independent of F&PP's case, for it applies equally to NS and to CM; for reasons of space we cannot discuss it here (see references in this footnote for a discussion).

[17] We have avoided explicitly talking of "fitness" in the text (thus our use of the long winded "a trait performing better a function that is beneficial for reproduction"), for there is a huge discussion in the literature about different meanings of the term and its role in NS that is not essential for our debate and may cause confusion. But there is one important comment worth making here. This net-like guiding principle-driven picture of unified theories also helps to answer a recurrent question involving fitness in the philosophy of evolution, namely, how general fitness is related with specific physical traits and functions (cf. e.g. Rosenberg 1978, 1983; Sober 1993; Brandon 1990). We refer here to ecological fitness, not to statistical fitness characteristic of population genetics (this terminology is Rosenberg and Bouchard's 2002, but the distinction is made by other authors using different terms; cf. Sober 1993; Matthen and Ariew 2002; Ariew and Lewontin 2004; Pigliucci and Kaplan 2006). The answer is that, in this net-like picture of NS, general ecological fitness is a general concept for the fact, stated by NSGP, that there is some ecological problem that is better solved with certain trait than without it, but it is in the particular specializations/applications of the principle that the specific functions and physical traits are mentioned. This structural answer seems to us clearer than other metaphysical ones in terms of supervenience (Rosenberg 1978; Sober 1993) or propensity (Brandon 1990). We thank S. Ginnobili for pointing out the connection to us and for proposing the structural, or meta-theoretical, solution.



5 Guiding Principle-Driven Explanations and Counterfactual Import

We are now in a position to answer the original question raised by F&PP's challenge: How does NS, *as a unified explanatory practice*, ground the relevant counterfactual differences? Answer: in exactly the same way in which CM grounds the corresponding relevant counterfactual differences; differences, for instance, between: (DIST) "if Mercury's distance from the Sun were different, its velocity would be different" and (COL) "if Mercury had a different color, its velocity would be different" [both read with *ceteris paribus* clauses, like (PUMP) and (NOISE)].

Which is this way? Let us follow it step by step:

- First, explanations have counterfactual import. How does CM account for the difference between (DIST) and (COL)? (DIST) is, and (COL) is not, true according to CM because there is an explanatory model of planets' velocity at the bottom of one branch of the CM theory-net with distance to the Sun as relevant initial condition, but there is no explanatory model with color as a relevant initial condition for planet's velocity. We take that this is uncontroversial as an explication of how CM accounts for the relevant counterfactual differences in mechanics. Generalizing: a counterfactual "if c were the case, then e would be the case" is true according to a theory T if T has a branch with an explanatory model that includes c (or not c) as a relevant initial condition in its explanans and includes e (or not e) in its explanandum. Thus, when one is asked to show how a theory accounts for counterfactual differences it suffices to show that it includes explanatory models of one kind but not of the other. And this is precisely what happens in NS and its relevant counterfactual differences: (PUMP) is true for there is an explanatory model in a branch of the NS theory-net that specifies pumping as beneficial for survival, hence for reproduction; (NOISE) is not true for there is no branch that specifies making noise as beneficial etc. The same applies to other relevant counterfactuals in NS such as: "if trees in moths' environment had not become dark, moths would not have (ceteris paribus) become black", "if moths' predators had disappeared, moths would not have (ceteris paribus) become black", or "if cost per male were higher than cost per female, then the sex ratio would be male-biased", etc. All these counterfactuals are true according to NS because NS contains the corresponding explanatory models that ground them in the aforementioned sense, just as in CM and in any other explanatory theory.
- Second, the counterfactual import of explanations relies on the use they make of nomic, i.e. non-accidental regularities. In the case of (DIST) in CM, the relevant

explanation makes use of the nomic correlation between distance and velocity (law of gravitation). In the case of (PUMP) in NS, the relevant explanation makes use of the nomic correlation between pumping and blood circulation and this in turn with survival and differential reproduction, i.e. (approximately) "in context…, heart-pumping facilitates (the survival function) blood circulation, which is beneficial for differential reproduction". And similarly in the other cases, specifying the trait that in the context facilitates the relevant (kind of) function beneficial for differential reproduction: "in context…, colorful tail facilitates (the mating function) sexual attraction, which is beneficial etc."; "…dark wings facilitate (the survival function) escaping predators, which etc." And asking why blood circulation (hence survival and differential reproduction) nomically depends on pumping and not on noise, is tantamount to asking why velocity depends on distance and not on color: the world contains some nomic (causal, if one likes) connections but not others. There might be some concerns with this primitiveness of nomic facts, but whatever our reaction is, it must be the same in NS and in CM. It is worth emphasizing that the facts just mentioned involved in NS explanations are *general* (i.e. they are *not singular* facts, they apply to populations, groups or the like, no matter how small) and *non-accidental* (i.e. with modal force, hence counterfactual supporting). It is true that these facts can be highly environmentally-relational and domain-restricted, but neither context-sensitivity nor domain-restriction cancels out their general and non-accidental character.

- Third, these general non-accidental facts are also of *adaptive* nature. The way in which NS accounts for the relevant counterfactual differences is not "disconnected natural history": the nomic regularities used in different explanations in different branches are all "adaptive" for they are specific instantiations of NSGP. Just as in the case of CM, which (F&PP surely agree) is not disconnected natural history at all: its guiding principle $\sum F = m \cdot a$ unifies the different explanatory models. Thanks to this unifying principle, we can see *different* explanations as "similarly mechanical" and how different mechanical forces have a combined effect. Likewise in NS: its guiding principle NSGP unifies different explanatory models, and because of this we can see *different* explanations as "similarly adaptive" and how different "adaptive forces" and different environmental pressures have a combined effect. Both features are worth stressing. On the one hand, thanks to its net-like unified structure upon CMGP, we have in CM that Earth–Moon and Sun–Mars systems are "mechanically" similar (i.e. gravitational rotation systems); that planets and electrons are "mechanically" more similar to each other (i.e. rotational forces) than to pendula; but pendula and springs are more similar to each other (i.e. oscillators) than to planets. Likewise, because of its net-like unified structure upon NSGP, we have in NS that black wings in moths and green skin in lizards are "adaptively" similar (i.e. escaping predators); that big horns in elks and colorful tails in peacocks are "adaptively" more similar to each other (i.e. mating-adaptive) than to moths' darkness; but moth's darkness and giraffes' necks are more similar to each other (i.e. survival-adaptive) than to peacock's tails. On the other hand, different mechanical forces, e.g. gravitation and friction, have a mechanical combined effect. Likewise, different adaptive forces,

> e.g. mating and survival, also have a combined adaptive effect [some trait may, in a context, be good for partner attraction but bad for escaping predators, e.g. the colorful tail may be good for mating but bad for survival; for a combined effect, cf also the *poeciliid* case mentioned above; see Dobzhansky et al. (1977), for a discussion of combined selective pressures]; we will discuss these features further in Sect. 6.

So this is how guiding principle-driven unified theories support the relevant counterfactuals. Note that, although epistemically subtle, neither CMGP *nor* NSGP is banal or truistic at all. CMGP does not simply say that "there is something" responsible for the trajectory, it says *the kind of something*, i.e. a mechanical force. Likewise, NSGP is what implicitly guides NS explanatory research, saying not merely that "there is something" in virtue of which the trait increases/decreases its presence, but *the kind of something* one has to look for, namely, an adaptive pressure and a trait performing the reproductively relevant function better/worse. Though empirically non-restrict taken in isolation, this unifying factor is neither simply empty nor a banal application of the principle of sufficient reason, nor a dispensable explicit definition.

This characterization, using the Kuhnian-structuralist framework, of NS as a unified guiding principle-driven theory is in complete agreement with NS practice, as the analysis of Fisher's model and other examples shows. And it is also in agreement with, and helps to understand better, sensible presentations of NS, like the one discussed above attributed by F&PP to the recalcitrant friend of NS, or the following one by Brandon (strikingly similar to, and strikingly independent of, the one by Kuhn quoted above), which we think this Kuhnian-structuralist analysis makes clearer:

> The principle of natural selection is an organizing principle, or to put it another way, a schematic law. As a general schema it is without empirical biological content, but it does serve to structure particular biological explanations of differential reproduction. In such instantiations the dispositional relationship of adaptedness is cashed out in terms of differences in particular traits. […] The instance of the principle of natural selection [for the moth's case] would state:
> (Probably) If moth *a* has darker colored wings than *b* in (this particular) *E*, then *a* will have more offspring than *b* in *E*.
> Clearly an instantiation of the principle concerning a particular population in particular environment has empirical biological content and explanatory value. […] The instantiations of the general schema have empirical biological content but they are not general, they apply only to particular populations under particular environmental conditions. Thus if we are to have a general theory of evolution, as opposed to numerous unconnected low-level theories concerning the evolution of particular populations in particular environmental settings, the principle of natural selection as general schema is necessary.[18]

[18] Brandon (1996, pp. 51–52); see also Brandon (1978, 1982); cf. McShea and Brandon (2010) for a different, and according to us less helpful, comparison to CM. For other references less similar to our Kuhnian-structuralist analysis but pointing also to the puzzling character of a NS general principle, cf. Sober (2008a, p. 47; 1993, p. 129), Kitcher (1982, p. 60), Resnik (1997, pp. 42–47), Rosenberg (1994, p. 122) and Endler (1986, p. 12).

The notion of a guiding principle that specializes in different particular laws for different, particular explananda helps to make coherent the idea contained in this quote, and in other works in philosophy of evolution, according to which the core principle of NS is *neither empirically specific nor empirically trivial*. Without a meta-theoretical analysis along the Kuhnian-structuralist lines we have offered such idea remains, we think, obscure. The proposal elaborated here is thus useful for a correct understanding, and better articulation, of this and other sensible characterizations of NS.

To conclude our appraisal of the use of a priori tools in NS, there are two main senses in which there are theoretical tools that may be qualified as "a priori":

AP1 In the models/explanatory conditionals of the kind "If *I* & *B* & *T*, then *P*", the consequent-explanandum follows, strictly speaking, a priori from the antecedent-explanans

AP2 The explanans contains, for a particular explanandum, a specific version/specialization of a general guiding principle which: (1) is presupposed in an analogous manner by any other explanation of the theory; (2) thus provides the unifying factor of the theory; (3) is empirically non-restrict taken in isolation and can be tested only through some of its specific specializations; and (4) has, in this sense, a specific epistemic status that might be qualified as "a priori", yet not in the AP1 or "bachelors are unmarried" senses

And (this is the crucial conclusion for our response to F&PP) *in neither of these two senses* is NS *specifically different* from other bona fide empirical theories like CM. Classical mechanics, thermodynamics, *…and natural selection* do behave similarly in the respects relevant to the point at stake here. This is what a detailed reconstruction of the structure of highly unified theories, and their explanations, shows. This is just (as Kuhn pointed out and Sneedian structuralism makes precise) what being a unified explanatory theory involves: to be a guiding principle-driven explanatory machinery.

Once it has been shown that NSGP is no more "merely definitional" or "banal" than Newton's second law, F&PP should withdraw, for they themselves have accepted that the friend of NS could block their case with a general adaptive principle (basically identical to our NSGP) "if, but only if" (sic) such a principle is not merely definitional. And it is not: no more than Newton's second law. They should, then, either continue to reject NS, but together with CM and any other guiding principle-driven theory, or acknowledge that NS is a perfectly bona fide explanatory unified theory. There is no other coherent *tertium datur*.

6 Discussion: The Analogy with CM

We will end by clarifying some of the issues related to our key analogy between natural selection and classical mechanics, which is crucial for our strategy against F&PP: the charge of a specific defectiveness of NS, compared with bona fide explanatory theories, for example CM, vanishes if one shows that the alleged defective traits of NS are shared by CM and, actually, are essential to any highly

unified explanatory theory. We will proceed in the form of answers to possible objections.

6.1 Contrary to Newton's Second Law, NSGP is not a law, at least not in the usual sense, but just a heuristic principle

We have tried to be careful in speaking of NSGP and not to use "law" but "principle with counterfactual-supporting force", which is the minimum sense of "nomic component" required for this debate. We claim that its programmatic or heuristic character is compatible with its (though peculiar) counterfactual-supporting nature. The problem is that it is odd to talk of NSGP's nomological character considered "in isolation", without its specializations for different explananda. But if we consider it "together with its specializations" within a unified theory-net, then, terminological differences aside, its counterfactual-supporting nature is undeniable. And, most importantly, it is on a par in this respect with CMGP: if having a heuristic character were enough to disqualify it as nomological-in-the-sense-of-counterfactual-supporting, then Newton's second law $\Sigma F = m \cdot a$ would not be a "law" either, for the same heuristic character applies to it in CM (and to other guiding principle "laws" in other theories).

6.2 NSGP seems more void than CMGP in that if, as seems to be our implicit use, the functions we refer to are relevant for reproduction, then the principle reduces, approximately, to something like: "a function or behavior relevant for reproduction that performs better/worse with a heritable trait produces an increase/decrease of the number of individuals with the trait", which sounds more void than CMGP[19]

Yet, even in this simplified version it is essential to include, as we did above, a ceteris paribus clause: "a function…produces, *ceteris paribus* (i.e. if no other—adaptive or non-adaptive—trait transmission pressures intervenes),…". Compare this now with a similar simplified version of CMGP: "a positive/negative force produces, ceteris paribus (i.e. if no other trajectory-modifying factor intervenes) an increase/decrease in the quantity of motion". We don't think that this is more void than the former, in particular if we keep in mind that "force" means in CM (i.e. it is *implicitly* defined as) "mechanically modifying trajectory factor"; as we emphasized in Sect. 4, CMGP is, when considered in isolation, empirically non-restricted. Thus, it is true that if we conceive of adaptive functions as those relevant for reproduction, what we claim in NSGP taken in isolation is very little, but it is also true that if we conceive of mechanical forces as those relevant for trajectory change, what we claim with CMGP taken *in isolation* is very little also. As we said above, without their specializations both principles sound equally "empty". They look more like a promise than like a particular empirical claim: when you see a change in trajectory/trait, look for forces/reproductive-relevant-functions-and-environmental-changes that account for it. In this regard both principles look similar. To repeat,

[19] We want to thank an anonymous referee for this and the next objection.

they just tell you the kind of thing you have to look for. This is very little, but enough to make them stronger than the mere principle of sufficient reason "if something happens it is because of something". Yet, as we also insisted, this becomes apparent only when particular applications of the principles are postulated/discovered. *Without* specific forces/adaptations these principles are explanatorily empty, just empirically void promises. One succeeds in mechanically explaining a trajectory when one comes out with a *specific* force that pre/retrodicts the trajectory. Likewise, one succeeds in adaptively explaining a change in trait distribution if one comes out with a *specific* reproductively relevant function or behavior that, in the environment, performs better/worst with the trait and so pre/retrodicting the trait evolution. And why, as F&PP ask, some explanations include certain factors and no others? Why pumping blood is relevant for reproduction but making noise is not? It is for same reason why a planet's distance to the sun is relevant for its trajectory but planet's color is not: the world contains certain determination (causal, if one likes) relationships but not others. What these relationships are is what special laws postulate/discover. In a nutshell: without specializations, both NSGP and CMGP are equally empirically empty, and with specializations both are equally empirically rich. And even alone, they are not merely a principle of sufficient reason: their specializations must, in each case, have something in common (more on this below). Thus, no difference in theses respects between NS and CM that is relevant to our case.

6.3 There is a relevant dysanalogy between NS and CM even accepting their net-like structure. On the one hand, the special laws in CM are derived from the general CMGP (plus other assumptions), but there is no similar derivational relationship between determinations of NSGP and that principle. On the other hand, and related to this, the construction of NS theory net is more arbitrary than in the case of CM, for in the latter there is a derivational criterion whereas in the former there is none

It is true that, *if* specific mechanical laws were *derived* from Newton's second law (under specific assumptions) but specific adaptive conditionals were not similarly derived from NSGP, then the second law would be essential in a sense in which NSGP were not, and this dysanalogy would invalidate our answer to F&PP. Yet, as we said in Sect. 4, the relationship between special laws and guiding principles is not of inference or derivation, but of specialization: special laws specify certain variables that are not yet specified in the guiding principle. The relationship is then more similar to being an instance of a scheme than to being a conclusion of a derivation. And this applies to both NS and CM: NS special laws specify the kind of function, relevant in the environment, that performs better/worse with than without the trait; CM special laws specify the kind of force that applies to specific trajectories. It is true that, since the second law says that the total incident force equals mass times acceleration, if one adds a specific assumption, e.g. that the total force is $-kx$, then from the assumption and the second law one *derives* the special law, e.g. Hooke's Law. But the same happens with NSGP combined with specific assumptions. Suppose that in a particular context the only selective survival pressure

is due to the presence of specific predators and that specific wing color improves camouflage from this kind of predator: from these assumptions and NSGP we can *derive* the selective moths-law mentioned by Brandon in the above quote. Thus, no difference between NS and CM with regard to derivations.

On the other hand, it is true that the reconstruction of a guiding principle-driven theory as a theory-net has specific elements of interpretation and in this respect it is not fully determined by text books, articles, etc. Yet, this is always so in any rational reconstruction, and it is indeed the case in this kind of Kuhnian-structuralist reconstructions, of NS but also of CM or genetics or other theories (cf. Balzer et al. 1987, and Balzer et al. 2000). In CM we distinguish a space-dependent node from a velocity-dependent node, inverse versus direct space dependent nodes, etc. In the same regard, in NS we distinguish survival form mating from … nodes; within the survival branch, we distinguish food supply from escaping predators nodes, etc. The net showed in Sect. 4 is just for the sake of exemplification, and a complete reconstruction would show more nodes and should present these criteria in more detail. But it suffices for justifying that the NS net-like structure is not totally arbitrary. We present food supply and escaping predators as two different realizations (among others) of the survival functions, and analogously attracting partners and fighting with sexual rivals as specifications of mating functions. But leaving certain interpretative problems aside (for instance, whether artificial selection is included within NS) the reconstruction has certain objective constraints, for instance one could not put food supply and attracting sexual partners under the same immediate node. Thus, again, there is no difference relevant to our case between NS and CM.

6.4 Even if the analogy applies, this cannot be the whole story, for explanations and laws in physics differ in important respects from theories and laws in biology in general, and in NS in particular

True, but not harmful, for what we said is the whole story *only as far as this debate is concerned*. The similarity argued for here concerns *only* this guiding principle-driven explanatory endeavor: *there is something adaptive/dynamic, non-truistic, counterfactual-supporting and common to different explanations of traits/particles trajectories*. This similarity is compatible with dissimilarities in other respects. For instance, it is actually the case that NS is not similar to CM and other theories in physics in the manner in which this structure is shown in scientific papers and text-books. In physics, the laws/principles, including guiding principles, for example Newton's second law, are often explicitly presented, even articulated, whereas in NS (and other biological theories) this is rarely the case, and one finds instead a seemingly disconnected collection of models, explanations, and problems solved. Related to this, it is also the case that NS literature often explicitly presents explanatory conditionals/models, which we saw are a priori, while the literature in physics does not, or not that often. Or it might also be the case that the counterfactual-supporting (i.e. nomological) elements of NS are, in relevant respects (e.g. their highly relational character, their domain-restricted application), so different from laws in physics that we do better not to label them "laws" [see, though, Dorato (2005) for a criticism of universality and unrestriction as marks of

laws in physics]. Or that NS offers in general more post hoc explanations (retrodictions) than CM.[20] Or maybe other "deeper" differences, e.g. maybe NS is less "metaphysically robust" than CM (i.e. applies to more varied ontological entities—organisms, species, organs, cells,…—while CM systems are ontologically less varied).[21] Or maybe NS is highly dependent on contingent facts, so that its generalizations and explanatory conditionals are implicitly conditionalized in a much stronger way than CM ones [yet see Dorato (2005), for a criticism of the alleged independence of laws of physics on contingent facts]. Or maybe lawhood may come in degrees (Lange 1999) and NS and CM are in different parts of the spectrum. And there might be other dissimilarities. Our point is that, irrespective of dissimilarities in these and other respects, the similarity elaborated here suffices to block F&PP's case against adaptationism as explanatorily void (and to qualify Sober's concerns about an alleged specific aprioricity of NS compared with CM).

6.5 Ok, the analogy could apply, but there is no need to buy it, for one can read different trait-explanations as independent of each other, ignoring the existence, and unifying role, of NSGP, thus obtaining the kind of "disconnected" natural history F&PP contend

True, one could do that. But one could do that with CM (and thermodynamics, and classical genetics, and any other guiding principle-driven unified theory) as well, and see the explanations of the pendulum and of the planets, and of inclined planes, and of … as "disconnected" from each other. If one does this, one loses the essential, and beneficial, features of unified theories discussed in Sect. 5:

(i) Different explanations "resemble each other" in being different instances of a general mechanism, a resemblance on which (at least part of) the theory's heuristic power relies. As emphasized by Kuhn, seeing the pendulum as similar to two inclined planes has an essential heuristic power; likewise, seeing the tongue in chameleons as similar to the spider's web has the heuristic power of explaining the former as a case of distance food supply.

[20] It is worth noting that, even if (though one could, cf. fn 14) one does not include artificial selection (which obviously makes predictions), NS does not make only retrodictions but anticipatory explanations also. For instance, NS makes general predictions such as that the color of prey mammals (with no strong dimorphism) will not differ greatly from the color of their environment. And it also makes other more specific predictions, both in the laboratory [e.g., for Lepidoptera—Kettlewell (1955, 1956)—and fruit flies—Maynard Smith (1993)] and in the field [e.g., the theory predicted an increase in the size of the beak of finches in a long drought season—Winer (1995), Grant (1999)—or the acquisition of resistance to treatments in viruses—Ridley (2004)]; we thank D. Blanco for these references. On the other hand, the greater number of post hoc explanations in NS has to do with the greater difficulty in predicting environmental changes and mutations. Moreover, the fact that an explanation is temporally posterior to the occurrence of the explanandum does not undermine per se the explanatory power of the adaptive mechanism, in the same way as the post hoc mechanical explanation of the elliptic orbit of the planets does not undermine the explanatory unified power of CM explanations.

[21] Yet it is not clear at all that NS applies to more varied natural kinds than CM: very different natural kinds have the functional property "adaptedness", but for sure no less different natural kinds also have the functional property "mass". If multiple realizability by different natural kinds were the problem, then every functional theory using multiple realizable properties would be unacceptable, a consequence that we do not believe Fodor would endorse.

(ii) Different effects may combine with each other (Dobzhansky et al. 1977), both in the same direction (e.g. for antelopes, raising the head facilitates both food supply and sexual attraction) or in opposite ones (e.g. a colorful tail may be good for sexual attraction but bad for escaping predators).

Thus, if one expels the guiding principle and retains only particular disconnected explanations, one loses both the heuristic and the combinatorial power of the unified theory and, according to us, misrepresents it both conceptually and historically. On the other side, the crucial point is that, as Sects. 6.1–6.3 emphasize, in regard of its guiding-principle-driven nature NS is in the same boat as CM (and thermodynamics, and classical genetics, and…), therefore applying this disconnected reading to NS and not to CM (and other guiding principle unified theories) simply begs the question. There is no more reason to treat NS trait-explanations as disconnected natural history than to treat CM trajectory-explanations in the same manner. What enables us to see planets and inclined planes as different *mechanic* phenomena is analogous to what enables us to see giraffe's necks and moths' color as different *adaptive* phenomena. This suffices to refute F&PP's case, for their case is that NS is *specifically defective* compared with CM and other theories. In brief: F&PP are correct in saying that the flesh lies in the details, but wrong in then concluding that the unifying factor is trivial and thereby methodologically defective. Quite the contrary, it is essential for seeing different specific applications as relevantly similar and for explaining combined effects, features which are essential to highly unified theories and the source of their distinctive fruitfulness, both in CM *and* in NS.

6.6 But all this makes NS unfalsifiable!

No! All this makes NS unfalsifiable only if we have the wrong, Popperian idea of falsification. Not if we have the more sophisticated, Kuhnian idea according to which theory replacement is a kind of confidence-loss phenomenon related to the persistence of anomalies. Logic obliges us to change something, but not to specifically change the guiding principle; the irrationality involved in abandoning the guiding principle too late, or too soon, is not formal but pragmatic or procedural. And here the analogy with CM also applies: CM is also a guiding principle-driven theory and it was abandoned, in this Kuhnian manner. And all that we have said is compatible with NS suffering from strong anomalies. Actually it has anomalies, and if they persist and intra-NS modifications do not solve them, NS could eventually enter into a Kuhnian crisis and be abandoned and supplanted by another, better theory (which, as in the case of CM, though changing essential features, may preserve some of the NS's central ideas).

7 Conclusion

We have argued that:

- a correct analysis of the structure of evolutionary theory and NS explanations makes it clear that F&PP's argument against NS is flawed;

- there are actual explanations in biological practice which *do* use nomic principles that can rightly be considered both adaptive and explanatory;
- these explanations have in common that their explanans include nomic components that are specific versions for specific explananda of a general NS guiding principle;
- although this general NS guiding principle has a subtle epistemic status, the same happens to analogous principles in other bona fide scientific theories such as CM;
- NS can account for the relevant counterfactual differences that F&PP call for, and it does so in the same manner in which CM accounts for its corresponding counterfactual differences;
- the counterfactuals are grounded in facts that are: *general*, i.e. they apply to populations or groups; *nomological*, i.e. non-accidental, with modal force; and *adaptive*, i.e. they are specializations of the NS general adaptive principle. And all this no matter how environmental-dependent and domain-restricted these facts be.

So there is no more room, based on the alleged specific "a priori" or "empty" components of NS, for F&PP to argue that there is something *specifically* defective in NS. As far as this debate is concerned, there is nothing epistemically specific, much less defective, in NS. The analogy with classical mechanics and its second law, elaborated in a Kuhnian-structuralist manner, enables us to give a complete response to F&PP's case and, for the first time in this debate, explicate at the same time:

the way in which NS accounts for the relevant counterfactuals;
where the unification lies and what its function is; and
the seeming emptiness of its general unifying guiding principle.

In *these* respects, which are those relevant to F&PP's case, NS behaves exactly like CM and other never challenged theories.

If we are right, then: who got what wrong? F&PP are wrong to claim that, when the distinction between selection of/for two coextensive properties is made in biological practice as a bona fide explanation of the evolution of traits, there is no non-truistic, counterfactual-supporting, adaptive mechanism common to different explanations that constitutes the unifying component of the theory of natural selection. Sober is right in what we think are related criticisms of F&PP, but according to us he is wrong to argue that NS explanations make use of a priori components (laws/principles/models…) in a sense in which CM explanations do not. Did Darwin get anything wrong? As far as this debate is concerned, the neo-Darwinist Darwin got nothing wrong, except perhaps for a misleading metaphor (though after a good methodological analogy),[22] because "selection" connotes decision, "decision" connotes intentionality, and intentionality presumably implies a mind. But leaving this metaphor aside, neo-Darwinism (Darwin's acceptance of

[22] As Lennox emphasizes, NS "began life as the product of analogical reasoning [from artificial selection]. Sebright [in 1809] sees clearly that the natural processes he is describing will have the same *effects* as the breeder's selection, but he is not about to describe those processes as selection processes. Darwin took that step." (Lennox 2004, 3.2).

the transmission of acquired traits is not part of neo-Darwinism) contains nothing epistemically or methodologically wrong.

In short: F&PP are wrong in that NS is *specifically* defective compared with standard theories like CM. Their central claim that *there is nothing adaptive, non-truistic, counterfactual-supporting and common to different explanations of traits evolution*, is false. Seriously read, it is as false as it is for CM and other never questioned theories. In all these cases, what different explanations have in common is epistemically subtle, but not truistic or "merely definitional" at all. Their claim is true only in an unacceptable superficial reading which, ignoring the essential role of guiding principles, dispenses with NS, but together with classical mechanics, thermodynamics, and all unified, guiding principle-driven theories. In either reading NS survives, or dies, in good company.

Acknowledgments We want to thank D. Blanco, A. Diéguez, M. García-Carpintero, S. Ginnobili, C. Hoefer, P. Humphreys, J. Lennox, M. Martínez, C.U. Moulines, E. Sober, the attendants of the LOGOS Seminar (Barcelona), the CPS Lunch Time Talks (Pittsburgh) and the IHPS Colloquium (Toronto), and an anonymous referee of this journal for comments and criticisms on previous versions of this paper. Research for this work has been supported by research projects FFI2008-01580/CONSOLIDER INGENIO CSD2009-0056 (Spain) and PICT2007-1558-ANPCyT and PIP-112-201101-01135-CONICET (Argentina).

References

Ariew, A., & Lewontin, R. C. (2004). The confusions of fitness. *British Journal for the Philosophy of Science, 55*, 347–363.

Balzer, W., Moulines, C. U., & Sneed, J. D. (1987). *An architectonic for science. The structuralist program*. Dordrecht: Reidel.

Balzer, W., Moulines, C. U., & Sneed, J. (Eds.). (2000). *Structuralist knowledge representation: Paradigmatic examples*. Amsterdam: Rodopi.

Block, N., & Kitcher, P. (2010a). Misunderstanding Darwin. Boston review, March/April. Retrieved from http://bostonreview.net/BR35.2/block_kitcher.php.

Block, N., & Kitcher, P. (2010b). Rejoinder. Boston review, Online March 17. Retrieved from http://bostonreview.net/BR35.2/darwin_exchange.php#reply.

Brandon, R. (1978). Adaptation and evolutionary theory. *Studies in History and Philosophy of Science, 9*, 181–206.

Brandon, R. (1982). A structural description of evolutionary theory. In P. D. Asquith & R. N. Giere (Eds.), *PSA 1980* (Vol. II, pp. 427–439). East Lansing, Michigan: Philosophy of Science Association.

Brandon, R. (1990). *Adaptation and environment*. New Jersey: Princeton University Press.

Brandon, R. (1996). *Concepts and methods in evolutionary biology*. Cambridge: Cambridge University Press.

Brandon, R. (1997). Does biology have laws? The experimental evidence. *Philosophy of Science, 64* (Proceedings), 444–457.

Carrier, M. (1995). Evolutionary change and lawlikeness: Beatty on biological generalizations. In G. Wolters & J. Lennox (Eds.), *Concepts, theories and rationality in the biological sciences* (pp. 82–97). Konstanz: Konstanz University Press and Pittsburgh University Press.

Coyne, J. A. (2010). The improbability pump. *The Nation*, 10 May. Retrieved from http://www.thenation.com/article/improbability-pump.

Darwin, C. (1859). *On the origin of species by means of natural selection* (6th ed.). London: John Murray.

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Dennet, D. (2008). Fun and games in Fantasyland. *Mind and Language, 23*, 25–31.

Díez, J. A. (2002). A program for the individuation of scientific concepts. *Synthese, 130*, 13–48.

Díez, J. A. (2005). The Ramsey sentence and theoretical content. In M. J. Frapolli (Ed.), *Belief, truth and probability* (pp. 70–103). London: Continuum.

Dobzhansky, T., Ayala, F. J., Stebbins, G. L., & Valentine, J. W. (1977). *Evolution*. San Francisco: W. H. Freeman & Co.
Dorato, M. (2005). *The software of the universe*. Aldershot: Ashgate.
Dorato, M. (2012). Mathematical biology and the existence of biological laws. In D. Dieks, W. J. Gonzalez, S. Hartmann, M. Stöltzner, & M. Weber (Eds.), *Probabilities, laws and structure. The philosophy of science in a European perspective* (Vol. 3, pp. 109–121). New York: Springer.
Düsing, C. (1884). *Die Regulierung des Geschlechtsverhältnisses bei der Vermehrung der Menschen, Tiere und Pflanzen*. Jena: Fischer.
Edwards, A. W. (1998). Natural selection and the sex ratio: Fisher's sources. *American Naturalist, 151*, 564–569.
Edwards, A. W. (2000). Carl Düsing (1884) on *The regulation of the sex-ratio*. *Theoretical Population Biology, 58*, 255–257.
Elgin, M. (2003). Biology and a priori laws. *Philosophy of Science, 70*, 1380–1389.
Endler, J. (1983). Natural and sexual selection on color patterns in poeciliid fishes. *Environmental Biology of Fishes, 9*, 173–190.
Endler, J. (1986). *Natural selection in the wild*. New Jersey: Princeton University Press.
Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Clarendon Press.
Fodor, J. (2008a). Against Darwinism. *Mind and Language, 23*, 1–24.
Fodor, J. (2008b). Replies. *Mind and Language, 23*, 50–57.
Fodor, J., & Piattelli-Palmarini, M. (2010a). *What Darwin got wrong*. (London: Profile Books) Second edition with an Afterword, New York: Picador, 2011.
Fodor, J., & Piattelli-Palmarini, M. (2010b). Misunderstanding Darwin, and exchange. Boston review, Online March 17. Retrieved from http://bostonreview.net/BR35.2/darwin_exchange.php#reply.
Futuyma, D. (2010). Two critics without a clue. *Science, 328*(5979), 692–693.
Ginnobili, S. (2010). La teoría de la selección natural darwiniana. *Theoria, 25*, 37–58.
Godfrey-Smith, P. (2008). Explanation in evolutionary biology: Comments on Fodor. *Mind and Language, 23*, 32–41.
Godfrey-Smith, P. (2010). It got eaten. *London Review of Books, 32*(13), 29–30.
Gould, S. (1977). *Ever since Darwin*. New York: Norton.
Gould, S., & Lewontin, R. (1979). The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London B, 205*, 581–598.
Gould, S., & Vrba, E. (1982). Exaptation—a missing term in the science of form. *Paleobiology, 8*, 4–15.
Grant, P. (1999). *Ecology and evolution of Darwin's Finches*. Princeton: Princeton University Press.
Hamilton, W. (1967). Extraordinary sex ratios. *Science, 156*, 477–488.
Kettlewell, H. B. D. (1955). Selection experiments on industrial melanism in the *Lepidoptera*. *Heredity, 9*, 323–342.
Kettlewell, H. B. D. (1956). Further selection experiments on industrial melanism in the *Lepidoptera*. *Heredity, 10*, 287–301.
Kitcher, P. (1982). *Abusing science: The case against creationism*. Cambridge, MA: MIT Press.
Kitcher, P. (1993). *The advancement of science*. New York: Oxford University Press.
Kuhn, T. S. (1970). Second thoughts on paradigms. In: F. Suppe (Ed.), *The structure of scientific theories* (pp. 459–482). Urbana, Ill.: University of Illinois Press, 1974.
Kuhn, T. S. (1976). Theory change as structure-change: Comments on the Sneed formalism. *Erkenntnis, 10*, 179–199.
Kuhn, T. S. (1990). Dubbing and redubbing: The vulnerability of rigid designation. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 14, pp. 298–318). Minneapolis: University of Minnesota Press.
Lange, M. (1999). Laws, counterfactuals, stability and degrees of lawhood. *Philosophy of Science, 66*, 243–267.
Lennox, J. (2001). History and philosophy of science: A phylogenetic approach. *Historia, Ciências, Saúde-Manguinhos, 8*, 655–669.
Lennox, J. (2004). Darwinism. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Substantive revised version Tue Jan 19, 2010. Retrieved from http://plato.stanford.edu/archives/fall2010/entries/darwinism/.
Lennox, J., & Bradley, E. W. (1994). Natural selection and the struggle for existence. *Studies in History and Philosophy of Science, 25*, 65–80.
Lewontin, R. (1978). Adaptation. *Scientific American, 239*, 213–230.

Lewontin, R. (2010). Not so natural selection. The New York review of books, May 27. Retrieved from http://www.nybooks.com/articles/archives/2010/may/27/not-so-natural-selection/?pagination=false.
Lorenzano, P. (2000). Classical genetics and the theory-net of genetics. In W. Balzer, C. U. Moulines, & J. Sneed (Eds.), *Structuralist knowledge representation: Paradigmatic examples* (pp. 251–284). Amsterdam: Rodopi.
Lorenzano, P. (2006). Fundamental laws and laws of biology. In G. Ernst & K.-G. Niebergall (Eds.), *Philosophie der Wissenschaft, Wissenschaft der Philosophie* (pp. 129–155). Paderborn: Mentis-Verlag.
Martínez, M., & Moya, A. (2011). Natural selection and multi-level causation. *Philosophy & Theory in Biology, 3*, e202.
Matthen, M., & Ariew, A. (2002). Two ways of thinking about fitness and natural selection. *Journal of Philosophy, 99*, 55–83.
Maynard Smith, J. (1993). *The theory of evolution*. Cambridge: Cambridge University Press.
McShea, D. W., & Brandon, R. (2010). *Biology's first law*. Chicago: University of Chicago Press.
Midgley, M. (2010, February 6). What Darwin got wrong, by Jerry Fodor and Massimo Piattelli Palmarini. The Guardian.
Mitchell, S. D. (1997). Pragmatic laws. *Philosophy of Science, 64* (Proceedings), S468–S479.
Mivart, G. J. (1898). *The groundwork of science*. London: John Murray.
Moulines, C. U. (1984). Existential quantifiers and guiding principles in physical theories. In J. J. E. Gracia, E. Rabossi, E. Villanueva, & M. Dascal (Eds.), *Philosophical analysis in Latin America* (pp. 173–198). Dordrecht: Reidel.
Moulines, C. U. (2002). Structuralism as a program for modeling theoretical science. *Synthese, 130*, 1–12.
Okasha, S. (2010, March 26). Review of Fodor and Piattelli-Palmarini's *What Darwin got wrong*. Times Literary Supplement.
Papineau, D. (2010). Review of Fodor and Piattelli-Palmarini's What Darwin got wrong. *Prospect, 168*, 83–84.
Pigliucci, M., & Kaplan, J. M. (2006). *Making sense of evolution*. Chicago: The University of Chicago Press.
Popper, K. (1963). *Conjectures and refutations*. London: Hutchinson.
Popper, K. (1972). *Objective knowledge*. Oxford: Oxford University Press.
Popper, K. (1976). Darwinism as a metaphysical research program. In K. Popper (Ed.), *Unended questions* (pp. 167–179). La Salle: Open Court.
Resnik, D. (1997). Adaptationism: Hypothesis or heuristic? *Biology and Philosophy, 12*, 39–50.
Ridley, M. (2004). *Evolution*. Oxford: Oxford University Press.
Rosenberg, A. (1978). The supervenience of biological concepts. *Philosophy of Science, 45*, 368–386.
Rosenberg, A. (1983). Fitness. *The Journal of Philosophy, 80*, 457–473.
Rosenberg, A. (1985). *The structure of biological science*. New York: Cambridge University Press.
Rosenberg, A. (1994). *Instrumental biology or the disunity of science*. Chicago: The University of Chicago Press.
Rosenberg, A., & Bouchard, F. (2002). Fitness. In E. N. Zalta (Ed.), *Stanford enciclopedia of philosophy*. Substantive revised version Thu Apr 17, 2008. Retrieved from http://plato.stanford.edu/archives/summer2008/entries/fitness/.
Rosenberg, A., & McShea, D. W. (2008). *Philosophy of biology*. New York and London: Routledge.
Shapin, S. (2010). The Darwin show. *London Review of Books, 32*, 3–9.
Sneed, J. D. (1971). *The logical structure of mathematical physics*. Dordrecht: Reidel. Second revised edition, 1979.
Sober, E. (1984). *The nature of selection*. Cambridge, MA: MIT Press.
Sober, E. (1993). *Philosophy of biology*. London: Oxford University Press.
Sober, E. (1997). Two outbreaks of lawlessness in recent philosophy of biology. *Philosophy of Science, 64* (Proceedings), 458–467.
Sober, E. (2008a). Fodor's bubbe meise against Darwinism. *Mind and Language, 23*, 42–49.
Sober, E. (2008b). *Evidence and evolution*. New York: Cambridge University Press.
Sober, E. (2010). Natural selection, causality and laws: Selection-for: What Fodor and Piattelli-Palmarini got wrong. *Philosophy of Science, 77*, 594–607.
Sober, E. (2011). A priori causal models of natural selection. *Australasian Journal of Philosophy, 89*, 571–589.
Sober, E., & Fodor, J. (2010). Discussion: Who got what wrong? Retrieved from http://bloggingheads.tv/videos/2492.
Winer, J. (1995). *The beak of a finch: A story of evolution in our time*. New York: Vintage.