J Gen Philos Sci (2012) 43:11–27
DOI 10.1007/s10838-012-9185-1

ARTICLE

The Structure of Idealization in Biological Theories: The Case of the Wright-Fisher Model

Xavier de Donato Rodríguez · Alfonso Arroyo Santos

Published online: 19 February 2012

Abstract In this paper we present a new framework of idealization in biology. We characterize idealizations as a network of counterfactual and hypothetical conditionals that can exhibit different "degrees of contingency". We use this idea to say that, in departing more or less from the actual world, idealizations can serve numerous epistemic, methodological or heuristic purposes within scientific research. We defend that, in part, this structure explains why idealizations, despite being deformations of reality, are so successful in scientific practice. For illustrative purposes, we provide an example from population genetics, the Wright-Fisher Model.

Keywords Idealization · Epistemic virtues · Scientific models · Modeling in biology · Population genetics · Wright-Fisher Model

1 Introduction

Although there is unanimous consensus with respect to the fact that idealization is a usual resource in scientific reasoning and an essential aspect in the construction of theories, there is not a similar consensus regarding the use of the term "idealization" and consequently, there is no unique systematic approach to treat idealizations and to understand their role within the structure of scientific theories. The situation seems to be more dramatic in the case of idealizations in biology because there are many ways of understanding the relationship between models, ideal conditions and the real world.

A quick survey of the biological literature shows that idealizations are routinely used in different contexts: for example, they can be used for model construction, in the constitution of concepts or entities highly dependent on a given theory, or in the form of causal

X. de Donato Rodríguez (✉)
Department of Logic and Moral Philosophy, University of Santiago de Compostela, Plaza Mazarelos s/n., 15782 Santiago de Compostela, Spain
e-mail: Xavier_donato@yahoo.com

A. Arroyo Santos (✉)
Facultad de Filosofía y Letras, Posgrado en Filosofía de la Ciencia, UNAM, Mexico City, Mexico
e-mail: agripas@gmail.com

hypotheses of a certain degree of abstraction which refer to the supposed mechanisms necessary to explain biological processes.[1] As a result, there are many interesting questions to be treated specifically regarding idealization in biology. Is there a particular method of idealizing in biology? What aspect (if any) does biological idealization have in common with idealization in physics and other natural sciences? How many kinds of idealization can be distinguished in biology? How are they produced and how are they empirically tested or justified? Our goal in this paper is not to answer all of these questions, but to focus on two major points that will help us to answer these questions in the future. Our major goals in this paper will be to (1) give an account of what an idealization is and (2) to use such account to understand the success of idealization in scientific practice.

Idealization can be considered from at least two points of view: (1) as a mental process or a certain procedure of reasoning (for example, of counterfactual deformation, as we think), and (2) as the product or the result of that process (in this sense, we can say that laws, models and theories are idealizations or are idealized, because they are the product of certain idealizing assumptions).

To simplify our discussion, we will characterize idealizations as statements that, so to speak, would express what is being idealized. As we will see later, these statements are of a very particular kind: they are the consequents of certain counterfactual and subjunctive conditionals whose antecedents express the idealizing assumptions under which the idealization holds.[2] However, this does not mean that we take idealizations to be sets of statements. It is clear that *idealizing* in science is part of a complex cognitive, epistemic and social activity that is not fully captured by abstracting it to a propositional affair. However, we believe that conceptualizing idealizations *as if they were statements* will help us to better introduce our conceptual framework.

In the present account, idealizations will then be understood as being statements built on a series of different conditions. Because such conditions could be easier or harder to meet in the actual world, an idealization will be formed by a network of multilevel conditions that, as we will argue, is in part responsible for both the diversity of idealizations and their success in scientific practice. As will be shown, the idea of a network does not only mean that idealizations are constructed on different conditionals but that these conditionals can interact to form idealizations of different degree.

Speaking of a "network" does not commit us to any particular structure or any particular logic that could represent it; it only has the purpose of highlighting that in this approach, an idealization is formed by a series of conditions that interact with one another to achieve two major goals: (1) to tell what the idealization is about, and (2) to give the idealization a number of epistemic virtues that in turn explain its success in scientific practice.

[1] See for example, Bechtel and Richardson (1993), Machamer et al. (2000), Darden (2002) and Darden and Craver (2002).

[2] The idea of considering idealizations as being statements goes back to Nowak (1980), although the idea already appears in Barr (1971), who speaks about idealized laws as "ideal cases" that are essentially understood as counterfactual statements with ideal conditions as their antecedents. In Barr's approach, theories become classes of ideal cases (statements). As the statement view of scientific theories (for which theories are classes of statements) was abandoned by most philosophers of science in favor of the semantic view, idealizations were understood as the process of model construction or as the product of that process (see, for example, Suppe 1989 and Balzer et al. 1987). In the present article, we don't want to favor any particular view about scientific theories. We just want to defend a certain account of idealization without committing us to any philosophically charged approach to scientific theories. Our proposal that idealizations are certain kind of statements is an idealization itself that is intended to help better understand their functioning in scientific reasoning.

The argument will be presented as follows: first (see Sects. 2 and 3.1), we will explain what we understand by idealization. We will propose the idea of a structured network of idealizations and will defend that thinking in terms of networks helps to explain two of the most salient characteristics of idealizations: on one hand their diversity, and on the other why they are so successfully used in scientific practice despite being deformations of reality. Then we will illustrate our account with the example of the Wright-Fisher model, an example coming from population genetics that has already received considerable attention in the philosophy of biology; we believe that using a familiar example will make it easier to exemplify our analysis.

2 Different Forms of Idealizing

It has been argued that idealization became a mark of scientific method at least since Galileo who is thought to be a champion of the use of idealization in natural sciences (see Nowak 1980, 34; Nowak and Nowakowa 2000, 17 and McMullin 1985). It was in a study about Galileo's method that McMullin (1985) distinguished three different "idealization" techniques: (a) mathematical idealization as the technique that mathematics brings into natural sciences, (b) construct idealization as a kind of counterfactual deformation which consists in the proposal of theoretical models "deliberately fashioned so as to leave aside part of the complexity of the concrete order" (McMullin 1985, 273), and (c) causal idealization, also a counterfactual deformation procedure, but regarding the world itself. This last form of idealization can be achieved by means of experimental control or just in thought (as a particular form of thought experiment). According to McMullin (1985, 248), all these techniques can be grouped under the label "Galilean idealization" not because Galileo invented all of them or even because he had a major responsibility for all of them, but because they played a distinctive part in shaping the "new science" of which Galileo was one of the first and most illustrious advocates.

More recently, Weisberg (2007) has distinguished three kinds of idealization: (1) Galilean, whose goal is to simplify complex phenomena to make them more tractable, (2) minimalist, that aims to single-out the primary causal factors that account for the phenomenon that is to be explained, and (3) the multiple-models idealization, which attempts to understand the causes that give rise to a phenomenon through the construction of different models, which are related but incompatible with one another. As we can see, these kinds of idealization are distinguished according to their different goals in biological model-building and practice, that is, according to different "representational ideals" –as Weisberg calls them (see Weisberg 2007, 639). So what McMullin presents is a conceptual and historical approach to idealization in order to characterize different ways of idealizing in science, whereas Weisberg proposes a pragmatic theory of idealization that does not try to account for the nature and structure of idealizations, but to classify different idealization methods according to their roles in scientific practice.

We agree with McMullin and Weisberg that idealizations can have many different justifications based on their goals in scientific practice or based on the different ways scientists can idealize. It is clear that these kinds of idealizations are not competitors in the sense that only one of them should better capture the essence of idealizing in science. Quite the opposite, they both reflect different legitimate ways of depicting idealization in scientific practice.

In particular, Weisberg's classification seems to be very appropriate from the epistemic point of view, that is, if our interest is to identify the epistemic virtues and goals behind idealizing and to provide a justification for the different kinds of idealization. However, it is important to note that in Weisberg's approach some very important questions are taken for granted, namely, what are idealizations?, why are they so diverse?, and given that they are deformations of reality, how come they are so successful in scientific practice? We try to relate all these questions within the framework of a single account of idealization that attempts to explain their function by means of their nature and structure. This is a major point of departure between ours and previous approaches to idealization. While it is beyond doubt that understanding how and why idealizations are introduced into science is a very important matter, this information does not tell us why they work nor what is it about these numerous strategies that groups them together under the label of *idealization*.

Below we provide an answer to these questions firstly with our analysis of idealization in terms of counterfactual distortion and hypothetical thinking and secondly with the idea of a structured network of idealizations. We will argue that this structure, in part, tells us what counts as an idealization and explains some of their different theoretical, cognitive and heuristic roles in science practice.

3 What are Idealizations?

So far we have seen that we can classify idealizations according to their pragmatic role. We could also classify them by focusing on the level at which they occur (if they occur at the level of model construction or at the level of causal hypotheses, for example). Furthermore, we can ask whether we can classify them by taking into account their "degree of idealization": if we pay attention to the fact that idealization is a process of counterfactual deformation, a departure to some extent from what is considered to be the real world, we may in fact classify idealizations according to their differing degrees of distortion.

This suggestion is not new. Jones (2005) argues that the intuition, which we share, according to which idealizations occur not only at different levels but in different degrees of idealizing could also be reflected in terms of a metric. To measure the degree of idealization a certain idealized system has, we should be able to measure:

(1) *how many idealizations* it contains or involve, and
(2) *how idealized* these idealizations are.

In the case of (1), if we could individuate each idealization contained in a particular idealized system (law or theory), we could eventually provide a satisfactory way to measure (at least partially) its degree of idealization. But, as Jones himself recognizes (see Jones 2005, 191), it seems very difficult to provide an individuation criterion for idealizations.

The second criterion is to be understood in the following sense: the degree to which a model M idealizes the real system S by representing it as having certain property ϕ is the degree to which M distorts the truth about S in representing it as ϕ. That is, we should evaluate *how far from the truth* M represent S as ϕ. This immediately relates the notion of idealization to the notion of truth approximation or truthlikeness, which as we know, is a problematic notion. Even though there have been many attempts to provide a formal characterization of truthlikeness, there is no clear and unanimously accepted definition of what an "approximation to the truth" should be.[3] So, although Jones' way of classifying idealizations is very important, it seems to be very difficult (1) to individuate idealizations,

[3] See Niiniluoto (1987, 1998, 1999). See also Kuipers (1987).

and (2) to provide a formal analysis for comparison between them. To avoid these difficulties, we propose another and perhaps more fruitful way of analyzing idealizations and their interrelations by resorting to the component (2) of Jones' intuition though without trying to give a metric for it and without relating it to the concept of truth approximation or truthlikeness.

3.1 The Network of Multi-Level Idealizations

Before we begin, just one note: this is not a formal account. We will say that idealizations are to be understood as a part (more specifically the consequent) of certain subjunctive and counterfactual conditionals. However, it is important to remember that we talk of idealizations *as if they were statements* in order to simplify the discussion and because, for all that matters in the present context, we think this "idealization" is justified for illustration purposes. More particularly, our intention is to draw on the philosophical machinery that has been developed for counterfactuals, especially by David Lewis, and then to move on. The reason is that, given that idealizations are understood as the intentional introduction of distortions into the actual world, counterfactual analysis seems a very good place to start. However, this is only the beginning. As important as it is, Lewis's account has received numerous and important criticisms that are yet to receive a satisfactory answer,[4] so we are using Lewis's theory as a source of inspiration but we are not committed to his metrics in any sense. Furthermore, we have already said that idealizations are very diverse both in their theoretical and practical uses, so it is not clear that a single formal approach would be powerful enough to account for all of them. Finally, to account for idealizations as they are used in actual scientific practice, the analysis of counterfactuals currently available should be adapted to be applicable to the complexity of the idealizational structure of scientific theories.

We begin the exposition of our framework with an example. Consider a statement such as "What if organisms could live at temperatures above 1,000°C". It could be argued that this statement cannot hold in the actual world because the usual chemical components that constitute organisms (for ex. carbon, hydrogen, or nitrogen) cannot associate in these conditions. However, it could happen that there are life forms made of other elements (for ex. some metal) that could live under such conditions.[5] For the last three decades, we know that there are hyperthermophilic bacteria, like *Thermus thermophilus*, that thrive above 100°C. Not that long ago it was considered that all bacteria were killed by boiling water, so any statement regarding such organisms before the identification of hyperthermophiles would have been as highly hypothetical as our previous example.

Strictly speaking, there are conditions that seem implausible based on our current knowledge but are not known to be false so that statements having these conditions as antecedents are not idealizations, but rather hypotheses. Nevertheless, we include this kind of statements in our framework as limiting-cases in order to cover the whole range of possible subjunctive assumptions (from mere hypothetical suppositions to truly counterfactual assumptions), but above all to account for what we believe is the widespread use of subjunctive conditionals in scientific practice. Consider, for example, the case of hyperthermophiles discussed above, and in general, the introduction of any number of contingent conditions in scientific models.

[4] See, for example, Bennett and Fine (1975), Nute (1976), Schlossberger (1978), Kvart (1992), Krasner and Heller (1994), Fogelin (1998), Bennett (2003), Tooley (2003) and Pruss (2007).

[5] As, for example, has been argued by Cleland and Copley (2005).

Overall, the idea is that by imagining hypothetical and counterfactual conditions we are departing more or less from the actual world, that is, we can remain farther or closer to it (being the hypothetical conditionals the limiting-case). This idea resembles Lewis' notion of distance between possible worlds where the actual world is fixed by the context. Lewis says, for example, that if we speak about physical necessity, we are restricted to worlds where the actual laws of nature hold true: the *accessible* worlds are those constrained by these laws of nature that are said to be met in the actual world (Lewis 1973, 5). We understand this however more in an epistemic than in a metaphysical sense. For us the context that fixes the actual world may be given by the principles of a certain theory or by the background knowledge at a certain time. Therefore, to speak about the "actual world" does not commit us to any form of realism. Here the "actual world" is fixed by the context and is represented by that part of our knowledge that we take for granted in some particular field (physics, biology,… or, more particularly, if we take biology: population genetics, evolutionary biology,…). Suppose we are taking physical knowledge for granted. With respect to this world, we can idealize in different degrees: it is not the same to make the idealization that there is no friction or that there are no external forces acting on a particular system as to make the idealization that bodies are like point masses, in which all the mass is concentrated in an infinitesimal point which lies at the center of the body, because, according to classical physics, no body can have zero mass. Furthermore, if we concentrate all the gravitational mass in a dimensionless point, then we would paradoxically get the result that the point has an infinite gravitational potential, which is impossible according to classical mechanics (see Shapere 1969, 147–149).

In a broader philosophical discussion, Nozick (2001, 148–155) has highlighted the importance of distinguishing between different "degrees of contingency", meaning that statements that are contingently false can be true in possible worlds that differ from our actual world. We will use the same term, degree of contingency, to mean that idealizations can differ in many different ways from the actual world and, therefore, can show degrees of contingency ranging from those closer to our world to those farther apart from it.

There is an immediate problem with the idea of telling what is closer or farther from the actual world. As has been said with respect to Jones (2005), there has not been a successful metrics for this. We are well aware of the problem, but for the purposes of this paper we believe it is possible to consider a qualitative indicator based on whether the conditionals can be tested empirically or not. There is nothing arbitrary here as it is clear that certain conditions can be amenable to experimentation (meaning that they can be reproduced in experimental settings, or directly tested in nature, or translated into a set of quantitative indicators), while other conditions, even though they are not fully amenable to empirical testing can be approximated. Other conditions, though, are simply impossible to reproduce experimentally while others that are very difficult to implement at present could be carried out if certain things happened. For example, consider that before the construction of the Large Hadron Collider some questions in particle physics were simply not amenable to experimentation due to a lack of technology. If this is correct, then we can distinguish different qualitative degrees that can help us to implement the notion of distance to the actual world.

Let us now begin by introducing some terminology. In this paper, idealizations will be considered as statements that are the consequent of certain subjunctive or counterfactual conditionals, in which the antecedent expresses the ideal or hypothetical conditions under which the idealization holds. For the whole conditional we will use the term *ideal-hypothetical case*, whereas for the antecedent we will use the term *ideal-hypothetical* *conditions*. The consequent will be called *idealization*.[6] The structure will then be: $C_1,\ldots, C_n \rightarrow S$, where $C_1,\ldots, C_n$ are the ideal-hypothetical conditions, S is the idealization, and the connective "→" stands for a counterfactual conditional.

The idea is that the hypothetical-ideal conditions should make explicit the possible worlds in which idealizations hold. Also, it is important to note that this formulation can also be applied to other notions of *idealization* that do not involve considering idealizations as statements, as the important matter is that scientific idealizations are formed by a network of concepts, ideas, or epistemic goals to name some examples.

The ideal-hypothetical conditions can have different degrees of contingency as follows (we illustrate each type or degree by providing an example from biology and, in order to better understand each case, also through an example from physics, where idealization is a very common procedure too):

(1) At the highest degree of contingency, we find cases in which $C_1,\ldots, C_n$ are completely idealized, in the sense that they are ideal-types or limiting-conditions that conflict with some theoretical principle. In biology, for example, such idealizations can be found in the construction of models in population genetics where there are infinite populations. In the case of physics, it is common to speak, for example, about bodies or particles as mass points, i.e. dimensionless points having their mass at their very centre as if it were concentrated at one point. We will use the term "abstracting ideal condition" (AI) to refer to this case.

(2) Then we find the case where $C_1,\ldots, C_n$ are contingently false but conflict with a well-established regularity in some or another field of biology. For example, in population genetics we could assume that there is no mutation or that there is no natural selection (these assumptions clearly contravene well-established biological principles and regularities). We will call such conditions "nomologically non-possible ideal conditions" (NONPI). To take an example from physics, take the case when we neglect the influence of the force of attraction of one body on another body which is much greater than the first. In this case, we are justified in doing so, because the effect is very small, but we are in fact violating Newton's third law.

(3) At the next step, we find cases in which $C_1,\ldots, C_n$ are also contingently false but do not explicitly conflict with a well-established regularity. In any case, we have strong empirical reasons to believe that they are false in the actual world so that, in this case, we can speak about "factual falsehoods". Nevertheless, these conditions, though being false, can be *approximately* met under experimental control; for example, in some models of population genetics it is assumed that populations lived in closed systems where there is no migration. This assumption does not hold in the actual world because populations do not live in isolation, but could be achieved experimentally. In the same way, in genetic regulation models two or three genes can be considered to be responsible for a particular trait even if we know that physiological processes are the result of a larger genetic network: although we know that a certain factor or parameter is present, we can isolate it in order to make easier the understanding of a phenomenon; in Levins' (1966) words, we sacrifice realism for precision and generality. We will call "experimental ideal condition" (EI) such case. Again, to take an example from physics, let us think about the case in which we

[6] Compare this with Barr (1971, 261ff). These distinctions are very similar to Barr's own terminology, although our analysis is different from his.

assume that there is no friction or no air resistance. In both cases, you can approximate, by experimental means, a state of practical void.

(4) Finally, we can distinguish cases in which $C_1,\ldots, C_n$ are purely contingent and, despite seeming implausible, we cannot really exclude the possibility that $C_1,\ldots, C_n$ may be fulfilled, now or in a next future. For example, the previous case of organisms living at very high temperatures. We will call this final case "implausible ideal condition" (IMPI). In order to give an analogous case in physics, you can take the supposition that, by using an ultramodern spacecraft, we can travel at a velocity close to the speed of light.

Of course, with the present classification we don't aim at being exhaustive. We could have distinguished other degrees as well (recall that the difference is only of degree and that in fact we could admit a whole continuum of idealizations). Even so, it is clear that this classification reflects some paradigmatic cases of idealization according to their degree of contingency. Furthermore, it is very important to remember that an idealization is not understood here as an independent entity, but instead as a network formed by different ideal conditions that interact with one another to accomplish three main goals: first, to express what the idealization is about, second to form different degrees of contingency and third, as a consequence of the different degrees of contingency, to endow scientific research with numerous epistemic, heuristic and cognitive virtues. As we have seen, idealizations can be understood in many different ways. We believe that one of the reasons why idealizations are so versatile is their structure: because ideal conditions can exhibit different degrees of contingency, they can serve different methodological, experimental or theoretical needs.

In the following diagram (Fig. 1), we want to represent the different types of idealization we have distinguished according to their degree of contingency, ranging from those closer to our world to those farther from it:

For illustration purposes let us consider a caricature of natural selection:

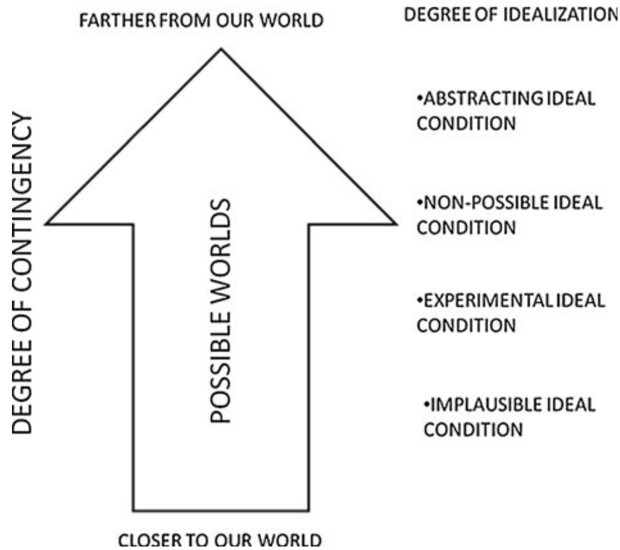

Fig. 1 Diagram representing the degree of idealization for different idealized conditions according to their distance from the actual world

> If, in a given population, every individual is different in some way from all other individuals within that population (C_1), all variable traits are inherited from parent to offspring (C_2), the population size can increase to infinity (C_3) and, if food supply is limited (C_4) → There is an unequal ability of individuals to survive and reproduce that leads to a gradual change in a population, with favorable characteristics accumulating over generations (S).[7]

In this example, the idealization "natural selection" (S) is the consequent of a counterfactual statement that consists of four hypothetical-ideal conditions represented by C_1–C_4. In other words, our natural selection is built on four ideal conditions that not only help us to understand what the idealization is about, but also set the conditions under which the idealization holds. As we will explain shortly, these conditions share some characteristics that in part explain why idealizations are so useful in scientific practice. Here it is important to note that idealization is not synonymous with abstraction: the idea is not to handpick relevant parameters and pretend nothing else matters, quite on the contrary, the idea is knowingly to create a false vision of the world and to use this "lie" as a powerful scientific tool. The idealization can then be used to test a model, to use it as a heuristic device to develop new theories or to test the limits of nature with questions such as *what would happen if…* and a long list of other uses as has been explored in the literature (for example, Levins 1966, Wimsatt 1987, Cartwright 1983 and 1989, Weisberg 2006a, b, 2007, Strevens 2008).

We contend that idealizations can fulfill all of these roles because they are composed of different elements, each being closer or farther from the actual world. In scientific practice, an idealization is not tested as a unit, but different questions target one or some of the elements forming an idealization network. Consider our example in which the idealization that there is a natural selection could be tested by searching for ways offspring could inherit characteristics from their parents (targeting C2), or by seeing if it is indeed true that all members of a population are unique (C1), or perhaps by seeing what happens when large populations live on scarce resources (C4). The moral of the story is that the idealization "natural selection" is actually tested and understood through the network of conditions that form it and not because of a single intuition inspired by it. As each of these conditions could be harder or easier to test, the idealization as a whole (or, more exactly, what we call the "ideal-hypothetical case") can exhibit a great deal of epistemic and heuristic virtues as well as an empirical sense.

For example, consider the case in which none of the conditions forming an idealization hold in the actual world: under this scenario the idealization could not lead to any empirical results as it would not have any connection with the world. Nonetheless, the idealization as a whole could still have heuristic, epistemic or cognitive power as its highly idealized conditions could open ways to construct new models, to improve on existing theories, or to come up with means to make such conditions applicable, to bring them closer to our world. However, this bringing them "closer to our world" should not be considered in a realistic sense or as a kind of truth approximation. As we have previously said, "our world" or the "actual world" has to be understood as something relative to the context and to the epistemic state of the scientists who are reasoning with this kind of counterfactual distortion and suppositions, i.e. those who are idealizing.

[7] An analysis of this example in terms of idealization can be found in Nowak (2000). See Nowakowa and Nowak (2000, chap. 2, 82ff) and ff. We have taken the example from Nowak's article.

In the present context, this process of "approaching the actual world" is based on the idea that we can better approach empirical data by relaxing our idealizations, that is, by *concretizing* them.[8] Concretization involves tentative hypotheses stating whether a certain (new) factor has or not an appreciable influence on our computations, models or laws, to name some examples. Furthermore, these concretizations can be made years or centuries after a theory has been formulated, as in the case of Newton deriving Kepler's laws from his gravitation theory. But they can also be made immediately as a way of testing our hypotheses. The idealization-concretization process, understood as the process which leads from idealizations by constructing models to the concretizations which make our theories more accurate, seems to be essential in scientific practice. As this method of making more accurate our theories is inseparable from empirical testing, it then involves observation and experimentation as necessary components. Idealizations can then be justified by showing how they allow for practical computability: "If it can be shown that more realistic initial conditions will lead via theory to correspondingly more accurate predictions, then the original highly idealized initial conditions are justified in the sense that they provided the starting point for a successful confirmational process".[9] The well-known puzzle of how to confirm idealizations can be solved in turn by invoking the concept of approximation: it is argued that it is "impossible" to test idealizations, because their antecedents (ideal conditions) are never realizable. To this objection we reply: well, maybe idealizations themselves cannot be directly confirmed, but can be tested if the idealized factors *do not assume the limit value*, but *approach it* .[10]

4 The Idealizational Structure of the Wright-Fisher Model

To understand better what we have in mind, we will exemplify our approach with an example coming from population genetics. As it was briefly summarized in Sect. 2, in biology, idealizations can be found in different forms and, accordingly, can accomplish different methodological, epistemic, heuristic or cognitive functions. Evidently, biology is a rich field where many different examples can be taken from, but we chose the Wright-Fisher model for three important reasons: population genetics is one of the most popular fields in philosophy of biology, so many readers will be already acquainted with the example and will make it easier for us to explain what we have in mind (for instance, in Walsh et al. 2002, there is a very nice discussion of this model). Furthermore, being mathematized, it is an easy example for people not familiar with biology (had we chosen some other field, we would have had to explain very detailed physiological mechanisms). Finally, because the beauty of population genetics is that everything is kept simple and straightforward, it is easy to see where the idealizations come from and how they are being used in scientific practice.

[8] For a detailed account of this process of idealization-concretization see Nowak (1980, Chap. 2), Laymon (1982) and (1985), Cartwright (1989), and Nowakowa and Nowak (2000).

[9] Laymon (1982, 115). Compare Niiniluoto (1999, 141–143).

[10] Authors such as Laymon (1982), Kuipers (1992a) and Kuipers (1992b) and Niiniluoto (1999) invoke the process of idealization-concretization in their different attempts to argue in favor of a convergent realism and method of approaching the truth. We do not want to commit ourselves to any form of scientific realism. The process of idealization-concretization may well serve to justify the use of idealizations, but not a substantive philosophical answer to the problem of scientific realism. See also Strevens's (2008) take on this issue.

4.1 Wright-Fisher Model

Fisher and Wright developed separately this model in a series of works (see Fisher 1930 and Wright 1931) that were inspired by their disagreement on the importance of genetic drift. Therefore, the basic idea behind the model is to describe how gene frequencies (the proportion of a given gene in a population) change within a population by chance alone. To achieve this, the model studies an ideal population (sometimes called Wright-Fisher population, see for example Gillespie 2004) under the following ideal conditions:

(C1) Mutation is not occurring.
(C2) Natural selection is not occurring.
(C3) The population is diploid, finite and constant (every generation will have N individuals).
(C4) Adults make an infinite number of gametes having the same allele frequency.
(C5) From the pool in (C4), 2N gametes are drawn at random to constitute the N diploid individuals for the next generation.
(C6) Every parent contributes equally to the gamete pool.
(C7) All members of the population breed.
(C8) All mating is totally random.
(C9) There is no migration in or out of the population.
(C10) There are no overlapping generations.

Now, for the case of a single locus with two alleles A_1 and A_2 we have:

i_t = number of A_1 alleles in time t. $p_t = i_t/2N$, the frequency of allele A_1 in time t. $q_t = p_t - 1$, the frequency of allele A_2 in time t.

Because there is an infinite number of gametes, the transition probability of going from i copies of A_1 to j copies of A_1 in the next generation is given by the binomial probability distribution:[11]

$$P_{ij} = \binom{2N}{j}\left(\frac{i}{2N}\right)^{i}\left(1 - \frac{i}{2N}\right)^{2N-j} \tag{1}$$

From (1) we have that the conditional expectation[12] of a given allelic frequency is:

$$E(j/i, N) = \sum_{i=0}^{2N} iP_{ij} = i \tag{2}$$

where it can be seen the important result that there is no net change in allele frequency between generations.

There are other details about this model but for our purposes this will suffice. Notice that the Wright-Fisher model is heavily idealized: it is composed of ten ideal conditions some of which could hold in the actual world and some others that definitely cannot. As said before, we contend that it is this combination of conditions exhibiting different degrees of contingency what explains why this model has been and still is so useful in population genetics.

Remember that in Sect. 3.1 we developed a series of qualitative indicators based on whether the different conditions can be met in nature, or approximated under experimental

[11] To keep things tidy, just remember that $i = i_t$ and $j = i+1$.

[12] Also known as the mean of the binomial distribution.

conditions, or if they are impossible to approximate even in experimental settings. Based on these indicators, we will notice that: (C1) and (C2) are highly idealized, there cannot be populations not undergoing mutation and natural selection. In our analysis these are nomologically non-possible ideal conditions (type ii) given that they conflict with known regularities. However, the idea is that these conditions will serve epistemic or heuristic purposes, for example, to ask what the effect of mutations could be on a given allelic frequency by modeling what happens when the genes of an entire population are invariant. (C4) is clearly an abstracting ideal condition (type i) as no adult can make an infinite number of gametes. As it happens with (C1) and (C2), this condition gives the whole idealization epistemic, heuristic and cognitive virtues. In this particular example, (C4) helps to envision the idea that it does not matter how many alleles there are, but if they are really chosen at random they will eventually have a predictable frequency.

Up to here, Wright-Fisher seems to be talking about impossible things. As an idealization that could serve certain epistemic virtues but having no contact with nature's phenomena, it could be of little use in actual scientific practice. This is why the idealization (or, which is better, the ideal-hypothetical case) must also include conditions that are closer to our world. Wright-Fisher achieves this with conditions (C3) and (C5) to (C10): (C7) is a condition that could be met in the actual world, but in most cases (for example packs of wolves or troops of primates where males mate according to their hierarchy within the flock), it may be considered an implausible ideal condition (type iv), given that we know that not all the members of a given population will breed and produce offspring. (C8) looks possible because in principle mating does seem to be random, only that we know sexual selection in general, and culture when dealing specifically with human populations turn the outcome not to be so random at all; therefore, it is an implausible ideal condition (type iv). Likewise, (C9) may be considered as an experimental ideal condition (type iii), given that it could be approximated under certain experimental conditions; for example, we could approximate it by placing a particular population in an isolated area. On the other hand, (C10) is something that happens in the real world, think for instance in the cases of cicadas or salmon where parents die shortly after laying the eggs. However, (C10) is aimed at any population, so for other cases it could only be approximated under experimental conditions.

(C3), (C5) and (C6) are so close to our world that they do not seem to be idealized at all. In the case of (C3), we know lots of organisms are diploid, populations are certainly finite but under real conditions, there is no way the population will remain constant across generations, that is, it is impossible that the population will have the same number of individuals generation after generation. However, such a condition could be achieved under experimental conditions. Therefore, this is what we call an experimental ideal condition (type iii).

Finally, (C5) and (C6) seem trivial. We all know that parents normally contribute equally to their daughters' genetic heritage and that the gametes involved are completely random. However, (C5) and (C6) are not individual conditions but form part of a larger one that includes (C3) and (C4). The condition formed by (C3)–(C4)–(C5)–(C6) illustrates the way in which the conditionals forming an idealization form a network that in turn can lead to new conditionals. In this case, the whole (C3)–(C6) condition refers to an ideal population whose members produce an infinite number of gametes that are all thrown into a basket, shuffled and drawn at random to produce the next generation. Notice that, under this scenario, individuals are the least of our concerns because all that matters is that there are 2N genes and that some of them are alleles A_1 and A_2. Following this logic, our population could very well be monoecious (producing both gametes) but that does not mean Fisher and Wright only cared about hermaphrodites. What happens is that (C5) and (C6) are terribly idealized conditions that form part of an abstracting ideal condition formed by (C3)–(C6). Now, just for the sake of argument, if what we have just said is correct, what should we do with conditions (C7) and (C8) that specifically talk about parents? Let us remember that Wright-Fisher needs them for empirical reasons. The idealization formed by (C3), (C4), (C5) and (C6) is so abstract that the model would again lose contact with natural phenomena. Idealizing that, even under such conditions, parents should behave in a certain way for their offspring to become a Wright-Fisher population, opens the gate for possible experimental testing in the sense of approximations if real populations are considered and some of the conditions are relaxed (what actually happens in scientific practice, see for example Jorde and Ryman 2007, Skalski 2007, Waples and Yokota 2007, Kitakado et al. 2006). Again, the network results in a new conditional formed by the (C3)–(C6) plus (C7) and (C8) showing how the whole idealization is not simply formed by a number of conditions, but these conditions also interact between one another to form new ideal conditions.

Finally, from what we have said, it could happen that different ideal conditions gave way to new idealizations. For example, it could happen that some of the ideal conditions discussed in the Wright-Fisher model opened the gate to a theoretical model that in turn rested on a new idealization. If this is correct, then this new idealization and the Wright-Fisher model would be part of an idealization-network sharing some ideal conditions and probably new concretizations.

In the following table (Table 1), we illustrate some of the ideal conditions discussed. The aim is just to exemplify the different kinds of hypothetical-ideal conditions we have distinguished above and some of their epistemic virtues as well as their different roles in scientific practice.

The Wright-Fisher model has been used as an example to show how it is the combination of low and high-level ideal conditions the element responsible for the model's success. This model has not only been used to study genetic drift but, for example, in

Table 1 Ideal conditions in the Wright-Fisher model and some of its epistemic, cognitive or theoretic virtues

Level of contingency	Ideal condition	Epistemic virtues	What for?
Abstracting (AI)	(C4): Adults make an infinite number of gametes having the same allele frequency	Heuristic power, simplification of parameters, cognitive virtues	Considering infinite populations can serve to simplify calculations in mathematical models, or theoretically, to consider "null" hypotheses
Nomologically non-possible (NONPI)	(C1): Mutation is not occurring	Simplification of calculations, cognitive viruses	As an heuristic in the sense of a reductio ad absurdum strategy
Experimental (EI)	(C10): There are no overlapping generations	Empirical power, prediction, simplicity, precision	Trade offs: models are less "realistic" but gain in precision and generality
Implausible (IMPI)	(C7): All members of a population breed	Prediction, empirical power	Can serve as confirmatory evidence, gives place to concretizations, empirical power

The table is meant only for illustration purposes

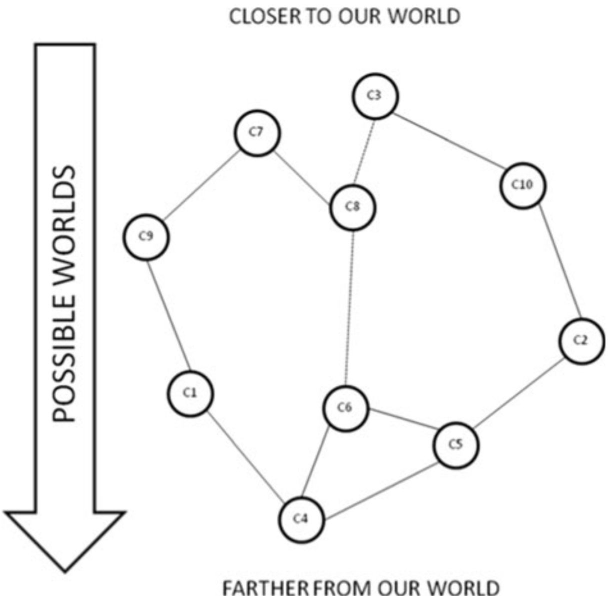

Fig. 2 Illustration of the Wright-Fisher Model as a network of ideal conditions of different degrees of idealization

combination with Kingman's coalescent theory (Kingman 1982), it is currently used for numerous studies of human evolution (see, for example, Helgason et al. 2003, Kingman 2000, Labate et al. 1999). The idea is that, while the model itself may be highly idealized, its conditions can be relaxed, meaning, as we said before, that different conditions can be targeted for concretizations; for example, introducing a mutation factor that will affect all genes in the gene pool (C4), or by exploring what happens if selection or even random genetic drift is introduced (see for example, Nagylaki 1979, Kimura 1954). This shows that it is not a particular idealization, but the entire network, what does the heuristic and epistemic work. The following figure (Fig. 2) illustrates how this idealizational network could look like.

5 Conclusion

Certainly, there has been a great deal of discussion about idealization in biology, but most of it focused on model and theory construction (for example, Keller 2000, Godfrey-Smith 2006a, b, Griesemer 1990, Levins 1966, Wimsatt 1987, Weisberg 2007). We agree with them that idealization is indeed important in these cases, but, as we have pointed out at the beginning of the article, idealizations cover many other scientific activities that can also include data models, concept formation or intertheoretical relations. Furthermore, idealizations in biology may serve very different purposes so that they may become successful or not depending on their ability of satisfying them. We have argued that this can happen because idealizations are not individual entities used in isolation in the context of model construction, but are a network of multilevel idealizations. This network is constituted by idealized conditions of different "degrees of contingency" and, as has been shown in the case of the Wright-Fisher model, the interaction of these conditions is in part responsible for the idealizations as a provider of numerous virtues. The idea is that the more highly idealized conditions enrich the model (or law or theory) with many different heuristic and epistemic virtues, whereas the conditions that are closer to the actual world give the idealization empirical power by opening the door to concretizations. Highly idealized conditions help to formulate problems, envision possible solutions, or formulate ways such abstract conditions can be applied in the real world. As Wimsatt has put it, "an over-simplified model may act as a starting point in a series of models of increasing complexity and realism." (Wimsatt 1987, 30). See also Wimsatt (2002).

Going back to our account of idealizational structure, compare what we have said in relation to Weisberg (2007) or McMullin (1985). Under McMullin's approach, the most important idea would be to understand how idealizations are introduced in scientific practice. For example, the ideal conditions of the Wright-Fisher model make it an example of construct idealization, because the distortion behind the conditions could be regarded as a means to abstract from a more complex phenomenon. In contrast, because Weisberg's approach is focused on representational ideals, the idea would be to understand the epistemic role of the idealization, for example in the building of models in population genetics. Contrary to these authors, what we wish to reflect is that idealizations are a network of conditionals (Weisberg and McMullin take the definition of idealization for granted), that these conditionals can have different "degrees of contingency" and that the combination of these conditionals explains, at least in part, why idealizations are so successful in scientific practice. Similar to Weisberg's account, we believe idealizations work because they introduce a number of virtues into scientific practice. However, we also believe our account is more general because Weisberg uses such "representational ideals" to link idealization with different model-building strategies, whereas we get inside the idealization to understand how it can provide Weisberg's representational ideals via the different degrees of contingency.

Our idea of different degrees of contingency is in some important aspects related to Mitchell's dimensions of scientific law (Mitchell 2000). Mitchell presents a novel account of scientific laws in biology based on the idea that there is only a difference of degree between accidental generalizations and scientific laws (she uses the expression "continuum of contingency" ranging from accidental truths to highly idealized laws). This degree could be measured in terms of strength, abstraction and stability (invariance). In her paper, Mitchell is not particularly interested in idealizations, but in arguing for a pluralistic notion of scientific laws as part of her wider project of defending a pluralist view of biological sciences (see Mitchell 2000). As part of this project, Mitchell believes idealizations are useful for explanatory reasons but does not provide a characterization. Rather, she is particularly concerned about the misuses of idealized models but without telling us how to tell when idealizations have gone too far and have lead to reification. To solve the problem of reification, perhaps it could help to have a characterization of idealizations such as ours, which could serve as a starting point to distinguish which aspects perform the epistemic and the empirical work. Perhaps in this way it would be easier to tell which aspects of the idealization have been inappropriately concretized. Finally, that our account shares with Mitchell the idea of a continuum speaks of the different rhetoric, cognitive, empiric and methodological strategies used in biological sciences that makes it impossible to talk about particular notions of laws, idealizations, theories or models, to name some examples.

As can be seen, idealizing conditions work together and this important feature is what we have tried to capture through the metaphor of a network. Despite the fact that we have only exemplified this approach in the case of a model, it is our belief that the same framework could be used in other forms of idealizing in biology. As we are here introducing this account, it would be interesting to study other cases and, above all, to use this discussion in relation to other topics related to idealization such as explanation or scientific realism.

Acknowledgments This paper has received financial support from the Spanish Ministry of Science and Innovation (Research Project Ref.: FFI2009-08828/FISO). A postdoctoral fellowship was awarded to Alfonso Arroyo Santos by the CONACyT ("Programa de Estancias posdoctorales vinculadas al fortalecimiento de la calidad del posgrado nacional" 2010). Previous versions of this article were presented at different conferences, including the PSA Biennial Meeting held on November 6–8, 2008 in Pittsburgh. The authors would like to thank the audiences on those occasions, and specially Mark E. Olson and two anonymous referees of this journal for their helpful comments.

References

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

Barr, E. W. (1971). A syntactical and semantic analysis of idealization in science. *Philosophy of Science, 38*, 258–272.

Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as strategies in scientific research*. Princeton: Princeton University Press.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.

Bennett, J., & Fine, K. (1975). Critical notice: Counterfactuals. *Mind, 84*, 451–458.

Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Clarendon Press.

Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.

Cleland, C. E., & Copley, S. (2005). The possibility of alternative microbial life on earth. *International Journal of Astrobiology, 4*, 165–173.

Darden, L. (2002). Strategies for discovering mechanisms: Schema instantiation, modular subassembly, forward/backward chaining. *Philosophy of Science, 69*, S354–S365.

Darden, L., & Craver, C. (2002). Strategies in the interfield discovery of the mechanism of protein synthesis. *Studies in History and Philosophy of Biological and Biomedical Sciences, 33C*, 1–28.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Clarendon Press.

Fogelin, R. J. (1998). David Lewis on indicative and counterfactual conditionals. *Analysis, 58*(4), 286–289.

Gillespie, J. (2004). *Population genetics: A concise guide*. Baltimore, MD: Johns Hopkins University Press.

Godfrey-Smith, P. (2006a). *Abstractions, idealizations, and evolutionary biology*. Retrieved March 28, 2008 from Harvard University, Peter Godfey-Smith online papers: http://www.people.fas.harvard.edu/~pgs/OnlinePapers/PGSAbstractnIdealizn06.pdf.

Godfrey-Smith, P. (2006b). The strategy of model-based science. *Biology and Philosophy, 21*, 725–740.

Griesemer, J. R. (1990). Modelling in the museum. On the role of remnant models in the work of Joseph Grimell. *Biology and Philosophy, 5*, 3–36.

Helgason, A., Hrafnkelsson, B., Gulcher, J. R., Ward, R., & Stefánsson, K. (2003). A populationwide coalescent analysis of Icelandic matrilineal and patrilineal genealogies: evidence for a faster evolutionary rate of mtDNA lineages than Y chromosomes. *American Journal of Human Genetics, 72*(6), 1370–1388.

Jones, M. (2005). Idealization and Abstraction: A Framework. In M. Jones & N. Cartwright (Eds.), *Idealization XII: Correcting the model* (pp. 173–217). Amsterdam: Rodopi.

Jorde, P. E., & Ryman, N. (2007). Unbiased estimator for genetic drift and effective population size. *Genetics, 177*, 927–935.

Keller, E. F. (2000). Models of and models for: Theory and practice in contemporary biology. *Philosophy of Science, 67*(Proceedings), S72–S86.

Kimura, M. (1954). Process leading to quasi-fixation of genes in natural populations due to random fluctuation of selection intensities. *Genetics, 39*(3), 280–295.

Kingman, J. F. C. (1982). On the genealogy of large populations. In J. Gani & E. J. Hannan (Eds.), *Essays in the statistical science* (pp. 27–43). London: Applied Probability Trust.

Kingman, J. F. C. (2000). Origins of the coalescent: 1974–1982. *Genetics, 156*, 1461–1463.

Kitakado, T., Kitada, S., Obata, Y., & Kishino, H. (2006). Simultaneous estimation of mixing rates and genetic drift under successive sampling of genetic markers with application to the mud crab (Scylla paramamosain) in Japan. *Genetics, 173*, 2063–2072.

Krasner, D., & Heller, M. (1994). The miracle of counterfactuals: Counterexamples to Lewis's world ordering. *Philosophical Studies, 76*(1), 27–43.
Kuipers, Th. (Ed.). (1987). *What is closer-to-the-truth? A parade of approaches to truthlikeness. (Poznań studies in the philosophy of the sciences and the humanities, vol. 10)*. Amsterdam: Rodopi.
Kuipers, Th. (1992a). Naive and refined truth approximation. *Synthese, 93*, 299–341.
Kuipers, Th. (1992b). Truth approximation by concretization. In J. Brzezinski and L. Nowak (Eds.), *Idealization III: approximation and truth. Poznań Studies in the Philosophy of the Sciences and the Humanities* (Vol. 25, pp. 159–179). Amsterdam: Rodopi.
Kvart, I. (1992). Counterfactuals. *Erkenntnis, 36*(2), 139–179.
Labate, J. A., Biermann, C. H., & Eanes, W. F. (1999). Nucleotide variation at the runt locus in Drosophila melanogaster and Drosophila simulans. *Molecular Biological Evolution, 16*(6), 724–731.
Laymon, R. (1982). Scientific realism and the hierarchical counterfactual path from data to theory. In P. Asquith and T. Nickles (Eds.), *PSA 1982* (Vol. 1, pp. 107–121). Michigan: East Lansing.
Laymon, R. (1985). Idealizations and the Testing of Theories by Experimentation. In P. Achinstein & O. Hannaway (Eds.), *Experiment and observation in modern science (147–173)*. Boston: MIT Press and Bradford Books.
Levins, R. (1966). The strategy of model building in population biology. *American Scientist, 54*, 421–431.
Lewis, D. (1973). *Counterfactuals*. Oxford: Blackwell.
Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.
McMullin, E. (1985). Galilean idealization. *Studies in the History and Philosophy of Science, 16*, 247–273.
Mitchell, S. (2000). Dimensions of scientific law. *Philosophy of Science, 67*, 242–265.
Nagylaki, T. (1979). The island model with stochastic migration. *Genetics, 91*(1), 163–176.
Niiniluoto, I. (1987). *Truthlikeness*. Dordrecht: Reidel.
Niiniluoto, I. (1998). Verisimilitude: The third period. *British Journal for the Philosophy of Science, 49*, 1–29.
Niiniluoto, I. (1999). *Critical scientific realism*. Oxford: Oxford University Press.
Nowak, L. (1980). *The structure of idealization*. Dordrecht: Reidel.
Nowak, L. (2000). Darwin's theory of the natural selection. In L. Nowak and I. Nowakowa: *Idealization X: The richness of idealization* (pp. 63–94). Amsterdam: Rodopi.
Nowak, L., & Nowakowa, I. (2000). *Idealization X: The richness of idealization*. Amsterdam: Rodopi.
Nozick, R. (2001). *Invariances :The structure of the objective world*. Cambridge: Harvard University Press.
Nute, D. (1976). David Lewis and the analysis of counterfactuals. *Noûs, 10*(3), 355–361.
Pruss, A. R. (2007). Conjunctions, Disjunctions and Lewisian Semantics for Counterfactuals. *Synthese, 156*(1), 33–52.
Schlossberger, E. (1978). Similarity and counterfactuals. *Analysis, 38*(2), 80–82.
Shapere, D. (1969). Notes towards a post-positivistic interpretation of science. In P. Achinstein and S. Barker (Eds.), *The legacy of logical positivism* (pp. 115–160), Baltimore: John Hopkins University Press.
Skalski, G. T. (2007). Joint estimation of migration rate and effective population size using the island model. *Genetics, 177*, 1043–1057.
Strevens, M. (2008). *Depth: an account of scientific explanation*. Cambridge: Harvard University Press.
Suppe, F. (1989). *The semantic conception of theories and scientific realism*. Urbana and Chicago: University of Illinois Press.
Tooley, M. (2003). The Stalnaker-Lewis approach to counterfactuals. *The Journal of Philosophy, 100*(7), 371–377.
Walsh, D. M., Lewens, T., & Ariew, A. (2002). The trials of life: natural selection and random drift. *Philosophy of Science, 69*, 452–473.
Waples, R. S., & Yokota, M. (2007). Temporal estimates of effective population size in species with overlapping generations. *Genetics, 175*, 219–233.
Weisberg, M. (2006a). Forty years of 'The Strategy': Levin's on Model Building and Idealization. *Biology and Philosophy, 21*, 623–645.
Weisberg, M. (ed.) (2006b). Richard Levins' Philosophy of Science, *Biology & Philosophy*, 21 (5).
Weisberg, M. (2007). Three kinds of idealization. *Journal of Philosophy, CIV*(12), 639–659.
Wimsatt, W. C. (1987). False models as means to truer theories. In M. H. Nitecki & A. Hoffman (Eds.), *Neutral models in biology* (pp. 23–55). New York: Oxford University Press.
Wimsatt, W. C. (2002). Using false models to elaborate constraints on processes: blending inheritance in organic and cultural evolution. *Philosophy of Science, 69*, S12–S24.
Wright, S. (1931). Evolution in Mendelian populations. *Genetics, 16*, 97–159.