

# **Idealization and the Structure of Biological Theories**

Alfonso Arroyo Santos<sup>1</sup> and Xavier de Donato Rodríguez<sup>2</sup>.

---

1 Posgrado en Filosofía de la Ciencia, UNAM.  
Mexico City, MEXICO  
agripas@gmail.com

2 Departamento de Filosofía, UNED.  
Madrid, SPAIN  
xavier\_donato@yahoo.com

## **Idealization and the Structure of Biological Theories**

In this paper we present a new framework of idealization in biology. We characterize idealizations as a network of counterfactual conditionals that can exhibit different degrees of contingency. We use the idea of possible worlds 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, it is this structure what helps explain why idealizations, despite being deformations of reality, are so successful in scientific practice.

Keywords: Idealization, epistemic virtues, biological modeling, Wright-Fisher Model.

Draft: November 2008.

Word count: 5693

## 1. Introduction

Although there is a 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 the same consensus in 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 surely 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 could be used for model construction, in the constitution of concepts or entities highly dependent on a given theory or in the form of causal hypothesis of a certain degree of abstraction which refer to the supposed mechanisms necessary to explain biological processes. As idealizations can be understood in different ways, any attempt to characterize them must not only account for its diversity but must also create a framework that can help us understand its success in scientific practice. This will be our goal in this paper. We will argue that idealizations in biology do not occur in isolation in the context of particular models or theories, but rather form a network of multilevel idealizations and it is this network what in part explains the diversity of idealizations and its widespread use in science. 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 the fact that idealizations are not isolated entities, but quite at the contrary, they are linked entities belonging to a more complex structure. Idealizations are here understood as statements that are usually known to be false in the actual world but nevertheless hold or would hold under

certain conditions. These conditions can depart from the actual world in a major or minor degree, that is, they can differ in their “degree of contingency” in a sense that will be explained later (see section 3).

The argument will be presented as follows: First (section 2) we will argue that there are many different ways of idealizing in biology. We will present some of the products of such idealizations and will defend that such diversity has to do with the internal structure of the idealization. Then (section 3), we will present such structure and will defend that it helps explain two of the most salient characteristics of idealizations: on one hand its diversity and on the other, why, despite being deformations of reality, they are so successfully used in scientific practice. Finally, in section 4, we will illustrate it all with the example of the Wright-Fisher model

## **2. Different forms of idealizing.**

One of the main problems we find in current accounts of idealization is that they analyze idealizations in biology as if they were all of the same kind. This is to say, they have not realized that idealizations in biology are used for numerous and distinct claims within a given framework and have consequently failed to properly analyze the different roles these idealizations play in theoretical accounts. There can be alternative ways of characterizing the different kinds of idealizations, for example, in terms of the processes and goals behind idealizing or in terms of the practice and the products of idealization.

For example, according to Weisberg (2007), there are three kinds of idealization: 1) Galilean, whose goal is to simplify complex phenomena to make it more tractable, 2) minimalist, that aims to discover the primary causal factors that account for the

phenomenon that is to be explained, and 3) the multiple-models idealization, that pretends to understand nature through the construction of different models, which are related but incompatible with each other. 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. In Weisberg (2006, 633), he says that “[e]numerating the representational ideals associated with the idealization approach is very difficult, for there are infinitely many ways to idealize, only some of them valuable.” But in Weisberg (2007), he tries to give a refined theory of idealization based on the different goals and epistemic virtues idealizations can exhibit. We agree with Weisberg that idealization can have many different justifications based on such goals and that these three kinds of idealizations are not necessarily competitors in the sense that only one of the three should better capture the essence of idealizing in science. Quite at the contrary, all three reflect different legitimate ways of idealizing in scientific practice.

Weisberg’s classification seems to be very appropriate at the epistemic level, that is, if we want to identify the epistemic virtues and goals behind idealizing and to provide a justification for the different kinds of idealization. But, in the present article, we want to develop a different approach which has more to do with the role idealizations occupy within the structure of scientific (particularly biological) theories. For this reason, our account will be based on the different products of idealizing (for example, the result of a set of idealizations can be an idealized theory or a particular idealized law). These different products occupy different and characteristic roles within scientific activity. From the wide set of products idealization can produce, we have selected four for expository reasons

(theories, laws, data models and mechanisms); we could have added for example, metaphors or intertheoretic relations<sup>1</sup> as well.

Let us now present a list of different kinds of idealizations focusing on the products rather than on the goals. Of course, the list does not pretend to be exhaustive.

Idealizations in biology can occur:

- in the form of idealized laws, such as Mendelian laws, Fisher's Fundamental Theorem or the Hardy-Weinberg equilibrium. In each of these cases, the idealization purports an idealized (generally mathematical) description of the world. For example, the Hardy-Weinberg equilibrium talks about an idealized population of constant size, where there is no natural selection, no mutations and individuals show no tendency to choose partners with particular traits (mating is totally random). These idealizations are necessary for Hardy-Weinberg to then introduce its classical mathematical description of allelic frequency ( $p^2 + 2pq + q^2$ ) and discuss the different factors that might perturb the equilibrium.

in the form of complex idealized systems, i.e. (biological) theories: these theories are said to be idealized because they contain ideal concepts and idealized laws as essential components. For example, Darwin's theory of evolution or immunology's Clonal Selection Theory (see for example, Tauber 1994) are not the cold, objective result of observation and empirical data but rather, complex apparatuses formed in

part (or even guided by) idealizations in the form of metaphors, ontological commitments, heuristics or simplifications of world's phenomena.

- in the form of causal hypothesis of a certain degree of abstraction which refer to the supposed mechanisms necessary to explain biological processes (see for instance (Bechtel and Richardson 1993, Machamer, Darden and Craver 2000, or Wimsatt 1986). In these cases, a mechanism is an idealized explanatory process of a certain physiological phenomenon, for instance the mechanisms of cellular communication where there is an orchestrated sequence of proteic, cellular and genetic signals that explain how cells “communicate” with each other. Likewise, we could mention DNA replication, protein synthesis or axonal growth, biological phenomena that are explained by these kinds of idealized systems.
- in the construction of data models. Here idealizations deal with the constitution of data that are more and more dependent on theoretical interests. For example, in immunology, where the characterization of at least the T- and B-cells was based on the clonal selection theory (Bretscher and Cohn 1970). Based on their original paper, Bretscher and Cohn needed to make such a distinction to satisfy some of the premises of the theory.

These examples show some of the different theoretical and methodological forms idealizations can take in biology. These different forms can be used to construct models, or in theory development, or to produce certain kinds of explanations (as in the case of

mechanisms). However, this does not mean that there are particular idealizations suitable for specific purposes: there is not a “type A” idealization used to construct models or a “type B” idealization used to formulate law-like generalizations. What we have is a general tool whose diversity makes it suitable for different empirical, heuristic or epistemic tasks in biology. If this is correct, then we are entitled to ask *why?* How come idealizations can serve so many different purposes? Why, despite being intentional distortions of reality, they can be used in scientific practice? In what follows, we will discuss that these different forms of idealizing have a common structure, and that they are organized in a particular way. We will show this in the particular case of an idealized model (the Wright-Fisher model), but it is our belief that this account can be applied to all the other types of idealization mentioned: as we will see, biological theories or models are actually composed of different kinds of idealizations forming a network that is in part responsible for their heuristic and explanatory power.

### **3. The network of multi-level idealizations**

All the different kinds of idealization presented in section 2 can be treated as cases of idealization as it is understood in this section.

According to the approach we will develop here, an idealization is to be understood as a part (more specifically the consequent) of a subjunctive, mainly counterfactual, conditional. Idealizations are currently understood as statements that are known to be false in the actual world, holding rather under certain counterfactual conditions, that is, in certain possible worlds. We also propose including as idealizations suppositions that hold in (sometimes highly) hypothetical conditions expressed in subjunctive form.



For example, we are thinking in statements such as “What if organisms could live at temperatures above 1000° 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<sup>2</sup>. For the last three decades, we know that there are hypertermophylic bacteria, like *Thermus thermophilus*, that grow above 100 °C. (Furthermore, we now know that organisms such as *colwellia* need temperatures below -20°C to grow and there are reports of such bacteria having basal activity at temperatures going from -80 to -120°C<sup>3</sup>.) Any statement considering such organisms before the identification of hyperthermophyles would have been as highly hypothetical as the previous example.

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. This idea resembles Lewis’ notion of distance between possible worlds, where the actual world is fixed by a particular context (for ex. the principles of a certain theory or family of theories and the background knowledge in a certain time). 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 in a very different manner. We will use the same term, degree of contingency, to mean that idealizations can differ in many different ways from the actual world. In short, we will use Lewis’ notion of distance between possible worlds but that does not mean that we endorse his particular semantics of

counterfactual conditionals (maybe another semantics could serve as well). One important difference is that we accept counterfactuals where the antecedent is true, that is, we do not commit ourselves to the thesis that the use of a counterfactual implies that the antecedent is always false<sup>4</sup>. However, we do agree with Lewis that counterfactuals are related to a kind of conditional based on “comparative similarity of possible worlds” (Lewis 1973: 8).

The idea of different degrees of idealization has been suggested by previous authors, for example by Martin Jones (2005), but, as far as we know, none of them has developed it with enough detail.

Let us now begin by introducing some terminology. As we have said, idealizations are statements that are the consequent of certain subjunctive (or counterfactual) conditionals, where 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, for the antecedent we will use the term ideal-hypothetical conditions. The consequent will be called idealization. The structure will then be:  $C_1, \dots, C_n \rightarrow S_1, \dots, S_k$ , where  $C_1, \dots, C_n$  are the ideal-hypothetical conditions,  $S_1, \dots, S_k$  are the idealizations, and the connective “ $\rightarrow$ ” could be understood according to Lewis’ analysis of subjunctive and counterfactual conditionals. Ideal-hypothetical conditions should make explicit the possible worlds in which idealizations hold. For illustration purposes let us consider a caricature of natural selection:

If, in a given population, every individual is unique and different in some ways 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 ( $C_4$ )  $\rightarrow$  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).

In this example, the idealization “natural selection” (S) consists of four conditional statements represented by  $C_1$  to  $C_4$ . 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 notice that idealization is not synonymous to abstraction: the idea is not to handpick relevant parameters and pretend nothing else matters, quite on the contrary, the idea is to knowingly create a false vision of the world and to use such “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 like *what would happen if...* and a long list of other uses as has been explored in the literature (for example, Weisberg 2006a, 2006b, 2007, Cartwright 1989 and 1983, or Levins 1966).

We contend idealizations can do all that, because they are composed of different elements each being closer or farther apart from the actual world. In scientific practice, the idealization is not tested as a unit but different questions target one or some of the elements forming an idealization network. Consider our example where the idealization that there is natural selection could be tested by searching for ways offspring could inherit characteristics from their parents, or by seeing if it is indeed true that all members of a

population are unique or perhaps by seeing what happens when large populations live on scarce resources. As each of these conditions could be harder or easier to test, the idealization as a whole can exhibit a great deal of cognitive, epistemic and heuristic virtues. For example, consider the case where none of the conditions hold in the actual world and therefore will never lead to any empirical result. In this case, 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). Now consider the opposite situation, that where all of the ideal conditions are closer to the actual world and can be subject to experimentation. In this scenario, the idealization as a whole will eventually lead to concretizations and could become an important tool for empirical research. The general idea is that, due to its internal structure, idealizations can exhibit different degrees of contingency and it is this combination what in part explains why idealizations are so useful for scientific practice.

This idea is in some important aspects related to Mitchell's dimensions of scientific law (Mitchell 1999). 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). Likewise, what we propose for the case of idealizations is that they also show degrees of contingency ranging from those closer to our world to those farther apart from it. In our work, we distinguish four particular cases but as

in her case, it could also be a continuum following a notion of distance between possible worlds similar to that of Lewis.

The ideal-hypothetical conditions can have different degrees of contingency as follows:

(i) At the highest degree of contingency, we find the cases where  $C_1, \dots, C_n$  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. We will use the term “abstracting ideal condition” (AI) to refer this case.

(ii) Then we would find the case where  $C_1, \dots, C_n$  are contingently false and conflict with a well-established regularity in some or another field of biology. For example, in population genetics the case would be to assume that there is no mutation or that there is no natural selection. We will call “nomologically non-possible ideal condition” (NONPI) such conditions.

(iii) Then we find cases in which  $C_1, \dots, C_n$  are also contingently false but do not conflict explicitly with a well-established regularity. In any case, we have strong reasons to believe that they are false in the actual world and can only be approximately met under experimental control; for example, the assumption made in population genetics according to which there is no migration 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' words, we sacrifice realism for precision and generality. We will call "experimental ideal condition" (EI) such case.

(iv) Finally, we can distinguish cases in which  $C_1, \dots, C_n$  are purely contingent, that despite not seeming plausible, we do not even know if they are possible in the actual world. For example, the previous case of organisms living at very high temperatures. We will call this final case "implausible ideal condition" (IMPI).

Of course, with the present classification we don't pretend to be exhaustive. There could be other kinds of idealization (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 is presenting 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, it is rather a network formed by different ideal conditions that interact with each other 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 turn idealizations into a polyvalent scientific tool. As we have seen, idealizations can be understood in many different ways. We believe that one of the reasons idealizations are so versatile is its structure: because ideal conditions can exhibit different degrees of contingency, they can serve different methodological, experimental or theoretical needs. Remember the list

presented in section 2 where we mentioned four different products of idealization: if they are now seen through the proposed framework, we will notice that mathematical models such as Hardy-Weinberg equilibrium need highly idealized conditions to play the role of a null hypothesis. That mechanisms (in the sense explored by Machamer, Darden and Craver 2000) need conditions that are closer to our world because they are constructed mostly from empirical data and presumably, refer to the causal entities responsible for a physiological processes, or that data models must be formed by both kinds of conditions given that they rely on the theory but predict the existence of certain entities<sup>5</sup>.

In the following diagram, 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:



Figure 1. Diagram representing the degree of contingency for different idealized conditions according to their distance from our world. See text for details.

#### **4. The idealizational structure of the Wright-Fisher model.**

To better understand what we bear in mind, we will exemplify our approach with an example coming from population genetics. As was briefly summarized in section 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 2002, there is a very nice discussion on 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, so that it is easy to see where the idealizations come from and how they are being used in scientific practice.

##### **4.1 Wright-Fisher Model**

Fisher and Wright developed separately this model in a series of works (see Fisher 1930 and Wright 1931) and was inspired by their disagreement on the importance of genetic drift. Therefore, the basic idea behind the model is to describe how gene frequencies (the



number of a given gene in the 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  $t$ .

$q_t = 1 - p_t$ , 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<sup>6</sup>:

$$(1) \quad P_{tj} = \binom{2N}{j} \left(\frac{t}{2N}\right)^j \left(1 - \frac{t}{2N}\right)^{2N-j}$$

From (1) we have that the conditional expectation<sup>7</sup> of a given allelic frequency is:

$$(2) \quad E(j|t, N) = \sum_{j=0}^{2N} j P_{tj} = t$$

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 10 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. If we take a closer look at each ideal condition, we will notice that: (C1) and (C2) are highly idealized, there cannot be a population 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 wonder 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, it is 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 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 whole condition talks about an ideal population where its 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 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) and Kitikado et al (2006). Finally, the idealization C3-C4-C5-C6 also shows how the whole idealization is not simply formed by a number of conditions but these conditions also interact between one another to form new idealizations. This is then, another very important sense in which idealizations form a multi-level network.

Wright-Fisher model has been used as an example to show how it is the combination of low and high-level ideal conditions what in part is responsible for the model's success. This model has not only been used to study genetic drift but for example, associated to Kingman's coalescent theory (Kingman 1982) is currently used for numerous studies of human evolution (See for example, Helgason 2003, Kingman 2000 or Labate 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, Naglyaki 1979, Kimura 1954)). This shows that it is not a particular idealization, but the entire network, what does the heuristic and epistemic work.

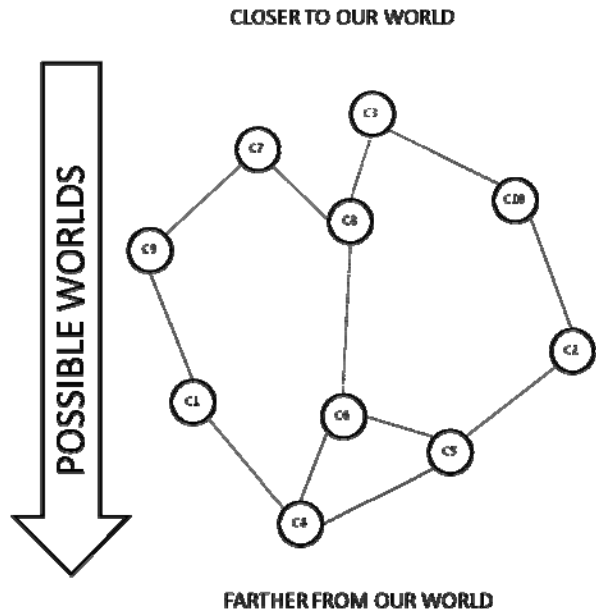


Figure 2. Visual representation for Wright-Fisher's idealization network. The figure is meant only for illustration purposes and should not be taken as representing an actual idealization network. Here, the Wright-Fisher model is represented by the whole structure, nodes represent its 10 idealized conditions, and continuous lines illustrate how all of them work in unison to express the idealization. Finally, dotted lines represent possible relations not discussed on the text but are used to represent that the ideal conditions can interact between them to make new idealizations like that made by conditions C3, C4, C5 and C6 (see text for further details).

## 5. Conclusion.

Certainly, there has been a lot of discussion about idealization in biology, but most of it centered on model and theory construction (for example, Keller 2000, Godfrey-Smith 2006a and 2006b, Griesemer 1990, Levins 1966, Wimsatt 1987, or even Weisberg 2007). We agree with them that idealization is indeed important in these cases, but in the present article we have aimed to show that idealizations cover many other scientific activities that can also include data models, concept formation or intertheoretical relations. 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, it is the interaction of these conditions what gives the idealization 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 those conditions that are closer to the actual world give the idealization empirical power in the sense of concretizations, highly idealized conditions help formulate problems, envision possible solutions, or formulate ways such abstract conditions can be applied in the real world. As Wimsatt has put it, “an oversimplified model may act as a starting point in a series of models of increasing complexity and realism.” (Wimsatt 1987:30). As can be seen, idealizations 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 any of the other kinds of idealizations discussed. 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 and scientific realism.

## **Bibliography**

Bechtel, W. & Richardson, R.C. (1993). *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton University Press, Princeton.

Bretscher, P.A. & Cohn M. (1970). A Theory of Self-nonsel Discrimination. *Science*, 169:1042-49.

Committee on the Limits of Organic Life in Planetary Systems, Committee on the Origins and Evolution of Life, National Research Council. (2007). *The Limits of Organic Life in Planetary Systems*. The National Academies Press, Washington D.C.

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

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

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(Suppl.), 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*. Clarendon Press, Oxford.

Gillespie, J. (2004). *Population Genetics: A Concise Guide*. Johns Hopkins University Press, Baltimore, MD.



Godfrey-Smith, P. (2006a). Abstractions, Idealizations, and Evolutionary Biology. Retrieved March 28, 2008 from Harvard University, Peter Godfrey-Smith online papers: <http://www.people.fas.harvard.edu/~pgs/OnlinePapers/PGSAbstractnIdealizn06.pdf>.

Godfrey-Smith, P. (2006b). The Strategy of model-based science. *Biology & 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 & Philosophy*, 5, 3-36.

Jorde, P.E., & Ryman, N. (2007). Unbiased Estimator for Genetic Drift and Effective Population Size. *Genetics*, 177, 927-935.

Jungea, E., Eicken, H., Swanson, B.D., and Deming, J.W. (2006). Bacterial incorporation of leucine into protein down to  $-20^{\circ}\text{C}$  with evidence for potential activity in sub-eutectic saline ice formations. *Cryobiology*, 52, 417-429.

Keller, E. F. (1998). What are models for? *Philosophy of Science*, 67 part II, 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. (2000). Origins of the Coalescent: 1974–1982. *Genetics*, 156, 1461-1463.

Kingman, J.F.C. (1982). On the genealogy of large populations. (In J. Gani and E. J. Hannan (Eds.) *Essays in Statistical Science*, (pp. 27-43). *Applied Probability Trust*, London).

Helgason, A., Hrafnkelsson, B., Gulcher, J.R., Ward, R., and 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), 1370-88.

Kitakado, T., Kitada, S., Obata, Y., and 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.

Labate, J.A., Biermann, C.H., and Eanes, W.F. (1999). Nucleotide variation at the runt locus in *Drosophila melanogaster* and *Drosophila simulans*. *Molecular Biological Evolution*, 16(6), 724-31.

Lange, M. (1999). Laws, Counterfactuals, Stability, and Degrees of Lawhood. *Philosophy of Science*, 66(2), 243-267.

Levins, R. (1966). The Strategy of Model Building in Population Biology. *American Scientist*, vol. 54, pp. 421–431.

Lewis, D. (1973), *Counterfactuals*. Blackwell, Oxford.

Machamer, P., Darden, L., and Craver C.F. (2000) Thinking about mechanisms. *Philosophy of Science*, 67, 1- 25.

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.

Nozick, R. (2001). *Invariances: the structure of the objective world*. Harvard University Press, Cambridge.

Skalski, G.T. (2007). Joint Estimation of Migration Rate and Effective Population Size Using the Island Model. *Genetics*, 177, 1043-1057.

Schaffner, K. (1992). Philosophy of Medicine. (In M. Salmon, J., et. al. (Eds.), *Introduction to the Philosophy of Science*, (pp. 310-344), Prentice Hall, Englewood Cliffs).

Schaffner, K. (1967). Approaches to Reduction. *Philosophy of Science*. 34, 137- 147.

Tauber, A.I. (1994) *The Immune Self: Theory or Metaphor*. Cambridge University Press, Cambridge.

Tilman, D., and Kareiva, P.M. (1997). *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*. Princeton University Press, Princeton.

Waples, R.S., & Yokota, M. (2007). Temporal Estimates of Effective Population Size in Species with Overlapping Generations. *Genetics*, 175, 219-233.

Weisberg, M. (2007). Three kinds of idealization. *Journal of Philosophy*, CIV(12): 639-659.

Weisberg, M. (2006b). Forty years of “The Strategy”: Levin’s on Model Building and Idealization. *Biology & Philosophy*, 21, 623-645.

Weisberg, M. (Ed.) (2006). Richard Levins’ Philosophy of Science. *Biology & Philosophy*, 21( 5).

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.

Wimsatt, W.C. (1987). False models as means to truer theories. (In M.H. Nitecki and A. Hoffman (Eds.), *Neutral Models in Biology*, (pp. 23-55), Oxford University Press, New York).

Wimsatt, W.C. (1986). Forms of Aggregativity. (In A. Donagan, A. Perovich, and M. Wedin (Eds.), *Human Nature and Natural Knowledge*, (pp. 201-216), Reidel, Dordrecht).

Wright, S. (1931). Evolution in Mendelian populations. *Genetics*. 16, 97-159.

Woods, M. (1997). *Conditionals*. Clarendon Press, Oxford

## NOTES

---

<sup>1</sup> For example, in intertheoretical reduction, Schaffner's General Reduction-Replacement model (Schaffner 1967, 1992) needs "corrected" reducing and reduced theories, this is, idealized conditions reducing and reduced theories have to meet for reduction to take place.

<sup>2</sup> As, for example, has been argued by Cleland & Copley 2005

<sup>3</sup> See for example, Junge et al. (2006)

<sup>4</sup> See for example Woods 1997

<sup>5</sup> For a similar notion to that of "degree of contingency", but in relation to scientific laws and the concept of invariance, see for example Lange (1999) and Mitchell (2000).

<sup>6</sup> To keep things tidy, just remember that in this and following equations,  $i=i_t$  and  $j=i_t+1$ .

<sup>7</sup> Also known as the mean of the binomial distribution.