An epistemic argument for evolutionary dispositions
(Un argumento epistémico en favor de las disposiciones evolutivas)

Cristina Villegas*
Universidade de Lisboa

Felipe Morales Carbonell
Universidad de Chile

ABSTRACT: The use of dispositions has been put into question many times in the philosophical literature, especially with regards to how dispositional attributions can be justified. Yet, dispositions are an important part not only of our everyday talk but also of our scientific practices. In this paper, we develop an argument that infers the epistemic justification of dispositional talk from its indispensability for carrying out basic epistemological projects, and we apply it to the use of dispositions in evolutionary biology. For doing this, we first advocate for a function-based strategy for the epistemic justification of dispositional attributions. We next review the functional role of some key dispositional notions in evolutionary biology, such as fitness and evolvability. Then, we show that alternative non-dispositional substitutes of these dispositions fail to fulfill their roles to the same degree. We conclude that the use of dispositions is justified in evolutionary biology.

KEYWORDS: dispositions, indispensability, epistemic justification, fitness, evolvability.

RESUMEN: El uso de las disposiciones ha sido puesto en entredicho en numerosas ocasiones en la literatura filosófica, especialmente respecto a cómo pueden ser justificadas las atribuciones disposicionales. Aun así, las disposiciones son una parte importante no solo de nuestra habla cotidiana sino también de la práctica científica. En este artículo, desarrollamos un argumento que infiere la justificación epistémica del lenguaje disposicional de su indispensabilidad para llevar a cabo proyectos epistemológicos básicos, y lo aplicamos al uso de disposiciones en la biología evolutiva. Para ello, primero defendemos una estrategia de justificación epistémica para las atribuciones disposicionales basada en su funcionalidad. Seguidamente, revisamos el papel funcional de algunas disposiciones claves en la biología evolutiva, como la aptitud o la evolucionabilidad. Luego, mostramos que las alternativas no disposicionales que puedan sustituir estas nociones no cumplen dicho papel con la misma eficacia. Concluimos que el uso de las disposiciones está justificado en la biología evolutiva.

PALABRAS CLAVE: disposiciones, indispensabilidad, justificación epistémica, aptitud, evolucionabilidad.

* Correspondence to: Cristina Villegas. Centro de Filosofía das Ciências, Departamento de História e Filosofia das Ciências, Faculdade de Ciências, Universidade de Lisboa. Campo Grande, Edifício C4, 3º piso, Sala 4.3.20 (1749-016 Lisboa-Portugal) – cvillegas@fc.ul.pt – https://orcid.org/0000-0002-8402-5288


Received: 24 September, 2023; Final version: 15 March, 2024.
ISSN 0495-4548 · eISSN 2171-679X / © UPV/EHU Press

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
1. Introduction

A disposition (e.g., fragility) is a multiply-realizable, functionally defined property of a system. Dispositions are tendencies to show specific behaviors, or to produce certain outcomes, that are attributed to systems carrying different kinds of mechanistic bases responsible for such behaviors or outcomes. Dispositional talk seems to be ubiquitous in the empirical sciences, and in biology particularly (Walsh, 2006; Austin, 2017; Hüttemann & Kaiser, 2018): cells are divisible, immune systems are responsive, populations are adaptable, and ecosystems are robust, to name just a few examples. Early philosophical arguments pointed to the epistemic value of dispositional talk (e.g., Fetzer, 1974), and the idea of dispositions or capacities is at the heart of our modern sense of scientific explanation (Cartwright, 1994). Yet, dispositions have always been surrounded with controversy. Metaphysically, there is a sharp divide between those who advocate for their irreducible existence (Bird, 2007), and those who argue that they must be reduced to sets of conditionals or to other kinds of properties, particularly categorical ones (Choi & Fara, 2021). The raison d’être for such a discrepancy seems to be the difficulty of attributing properties that are not manifested but in latent form: fragile objects are so even if they never break. The epistemic counterpart of this situation is a divide between those who believe we can have knowledge of dispositions, and those who think we cannot access them. Consequently, the support for dispositional concepts is mixed in the philosophy of biology, for there often seems to be reductive positions to counterbalance the dispositional understanding of terms (e.g., Garson, 2017).

Epistemic reduction can come in many ways in biology (Brigandt & Love, 2023). In the case of dispositions, we can identify operationalist views (e.g., Matthen & Ariew, 2002), which resonate with conditional analyses of dispositions (i.e., having a disposition to $X$ is just showing behavior $X$ under conditions $Y$; Choi & Fara, 2021); or views involving the identification of non-dispositional causal components (Wimsatt, 1976), which resonate with the reduction of dispositions to their causal bases. The causal basis of fragility can be a specific molecular structure of the fragile object, while the causal basis of evolutionary fitness (a prominent biological dispositional notion) could be carrying a specific trait in the right environment.

Notice that the problem of empirical accessibility does not disappear in either of these reductive strategies. Indeed, one still needs to make use of counterfactual situations, or perhaps even laws of nature, for the reduction to take the full role of dispositions. Non-direct accessibility is intrinsic to all purely modal reasoning, namely all thoughts and claims not engaged with actual but possible facts, and this includes dispositions, but also counterfactual and nomic reasoning. In fact, the dispute over dispositions parallels Van Inwagen’s (1998) classical concern about the epistemology of judgements regarding unrealized possibilities more generally. That is, how do we access the possible as opposed to the actual?

Given this situation, one can wonder when reductive attempts are epistemically justified, and when they convey a loss of explanatory power. If the main reason for abandoning dispositional talk altogether, namely our impossibility of direct access to the possible, affects in some or other way to reductive strategies too, we must find other grounds for discerning when dispositional talk is epistemically sound. Interestingly, modal epistemology is usually concerned with our general access to the possible, but says little about which kinds of modalities should be preferred in different settings (e.g., dispositional talk vs coun-
terfactual reasoning). The task for a dispositionalist is to find epistemic arguments for the non-reduction of dispositions to other kinds of modality (e.g., counterfactual reasoning).

In this paper, we build an epistemic argument that helps in identifying when reduction of dispositions concurs in an epistemic loss. We particularly focus on evolutionary biology explanations, where many prominent dispositions are already discussed in the philosophical literature, such as fitness, evolvability, or modularity. We argue that these dispositional notions are not merely operational: they are modal notions that refer to the possible, and we further defend that such a modal nature is not epistemically reducible to non-dispositional modalities. While the epistemic value of dispositions has been vindicated in other contexts, so far, no modal epistemology has been applied to this issue. Here we take a step further in advocating for dispositions by arguing for their value to be epistemically justificatory of dispositional talk in specific explanatory contexts.

In Section 2, we develop our justification strategy by proposing an epistemic indispensability argument for the use of dispositions in biology, based on the function that dispositional attributions play in our reasoning, and the role such reasoning plays in rationally required projects (Enoch & Schechter, 2008). In Sections 3 and 4, we apply our argument to fitness and the principle of natural selection, and to variational dispositions such as evolvability, respectively. We conclude that the use of dispositions is epistemically justified in evolutionary biology, and point at potential further implications.

2. The indispensability approach to epistemic justification

We begin by identifying two general strategies for assessing epistemic justification: a mechanism-based strategy, and a function-based one. The mechanism-based strategy consists in scrutinizing the reliability and trustworthiness of the belief-formation mechanisms involved in attributing dispositions. Can we rely on them to form true, justified, knowledgeable, beliefs? For instance, imagine that we formed our beliefs about dispositions by means of inferences. Under what conditions could such inferences be relied on? If we could do so in a range of cases, we could then say that this inferential basis gives epistemic justification to our talk of dispositions.

In contrast with knowledge about actual facts —i.e. whether something is the case—modal claims are those that concern whether something is possible or impossible, contingent or necessary. Dispositional talk is one kind of modal claim, since it points at what an object can do rather than at what it is actually doing. As such, it must be acquired by some form of modal belief-formation mechanism (Strohminger & Yli-Vakkuri, 2017). How modal claims relate to reality is a metaphysical concern relatively decoupled from the undoubted fact that we do make judgements about the possible, and that we tend to do so with different degrees of reliability. The literature on modal epistemology deals with this problem, and it provides a series of accounts for modal belief-formation mechanisms. Some

1 Cf. the first chapter of Hannon’s (2019), where he places his own function based methodology with other methodologies in epistemology. Our distinction is broader: both Hannon’s function-first epistemology and what he calls reverse engineering are function-based approaches in our way to classify things. It is important to note that Hannon focuses on knowledge rather than justification, but much of what he says carries over methodologically.
examples of accounts of modal belief formation are conceivability-, counterfactual-, similarity- and abduction-based accounts (see Strohminger & Yli-Vakkuri, 2017, for a review). Thus, this mechanism-based strategy consists in identifying which of these accounts are right for dispositions and addressing their validity. For example, we could hypothesize that dispositional attributions are made on the basis of an abductive process, and then evaluate whether abduction is reliable enough for there to be justification on those attributions.

While such an exercise would be interesting on its own, we see no a priori reason to assume that dispositional attributions are acquired through a particular belief-formation mechanism that is specific to dispositions only. It is doubtful that dispositional attributions could differ from, for example, attributions of counterfactuals, in terms of the belief-formation mechanisms enabling them. In fact, the literature on modal epistemology does not engage in any fine-grained categorization of types of modalities in dealing with belief-formation mechanisms. Therefore, the farthest away that this strategy could lead us to is a defense or a rejection of modal claims generally, but not dispositions in particular. Since the opponents of dispositions also involve modality (e.g., counterfactuals or laws of nature), we see no payoff in following this strategy.

More promising, we believe, is to address the epistemic justification of a dispositional framework by following what we call the function-based strategy, namely by focusing on how it serves different functions. The functions dispositions serve can indeed be specific of dispositional thought as compared to other means of modality, such as counterfactual reasoning or claims about the laws of nature. The idea behind this strategy is that the epistemic role of dispositions (at least partially) justifies their use as compared to other kinds of modal attributions.

There are reasons to think that a full-fledged defense of the adoption of a dispositional framework will not only benefit from this, but actually require it. Suppose that we knew that we had reliable mechanisms to form true beliefs about dispositions. Even in this case, it could be legitimately asked whether dispositional judgements formed by those mechanisms could not be functionally replaced with equally justified non-dispositional judgements. Taking a biological example: imagine that for any claim about fitness under a dispositional interpretation we could make a claim about fitness in counterfactual, non-dispositional, terms, and that no loss was incurred if we made that move. Why use dispositional talk then, considering the potential shortcomings of alluding to non-manifested properties, rather than the counterfactual framework? In order to discern whether one kind of modality is justified over another, thus, we think that paying attention to the function-base strategy may pay off.

2.1. The scheme of an epistemic indispensability argument for dispositional thought

We will start from the observation that the use of dispositional concepts is embedded in a network of practices, which are connected to one another with different relations of dependence. By mapping out those dependency relations, and examining what they tell us about whether dispositional concepts and talk can be replaced, one can highlight the kind of justification that they may have. This involves taking a primarily synchronic perspective (cf. Gardiner, 2015) of the functionality of dispositions, as opposed to a diachronic one, which involves describing how the concept developed over time and came to have such a role.
To drive this idea further, we now develop an *indispensability* argument for dispositions, similarly to how this has recently been advanced for modality more generally (Morales Carbonell, 2020). The notions of *dispensability* and *indispensability* characterize some important dependency relations of the sort we have just indicated. We will understand indispensability as a relation between some X (e.g. a concept, a practice or a theory), some Y which serves as a context (e.g. a task, project, practice, or discipline) and a set Z of items of the same type as X which could serve as alternatives to X. In these terms, the relation can be characterized as follows:

X is indispensable for Y iff there is no alternative Z’ in Z to fulfill Y that would play this role better than X (by some appropriate standards); that is, X is indispensable for Y iff eliminating X would leave us without a way to fulfill Y or, if it did, none of the alternatives to X left in Z would be better than nor as good as X in fulfilling it.

Notice that indispensability here is a property something has in the context of a particular project (Y): something is indispensable for achieving a certain goal in a given context. So we do not rule out that an alternative to what we take as indispensable in one context could be indispensable in other contexts (for example, that non-dispositional explanations might be indispensable in some other tasks or projects). Importantly, the relevant context (Y) is also what determines what is an ‘appropriate standard’ for X to play a role in it. Standards are typically fixed at the community level, in terms of what makes sense in view of the goals of the relevant projects —cf. Chang’s (2022) notion of operational coherence, which points towards this same idea. In the case of dispositional notions in evolutionary biology, the standards will be those set by scientific explanations, in particular those with causal content (see Sections 3 and 4 below). Thus, in what follows, we will be evaluating whether dispositions are indispensable in terms of the standards that are at play when looking for causal explanations. Part of this endeavor relies on what has been called theoretical virtues (such as coherence, adequacy, simplicity, and so on), which are supposed to ground our preferences for certain explanatory means over others. This preempts the following worry: if just any appropriate standard is viable, what is stopping someone who wants to endorse intelligent design, just to give an example, from appealing to some standard according to which their theories are better than selectionist accounts, and thus claiming that intelligent design is indispensable for biological explanations? Our response is that we would not expect such appeals to gain any traction at the community level, assuming the availability of selectionist explanations and their acceptance by members of the relevant community.

The strategy of appealing to the notion of indispensability might seem a bit unusual. Arguments concerning the notion have often been used in the context of the debate between metaphysical realism and anti-realism, rather than in epistemological contexts, with realists endorsing them as a way to establish realism. The classical example is the so-called Quine-Putnam argument in favor of the existence of mathematical entities. The argument goes as follows:

(P1) We should be committed to the existence of the entities (all and only) which are indispensable for our best scientific theories,

---

2 We thank a reviewer for raising this worry.
Mathematical entities are indispensable for our best scientific theories, we should be committed to the existence of mathematical entities. The general idea behind the argument is to ‘pump’, as it were, confidence in the existence of the target class of entities (i.e., numbers) from the confidence in our best scientific theories. It is important to notice that the latter is sustained by the fact that those theories already satisfy at least some of the relevant appropriate standards. However, here we are interested in utilizing an argument of this kind for an epistemological, rather than ontological, purpose. For doing this, we need to think of indispensability arguments in a more general way. Panza and Sereni (2016) describe the general function of an indispensability argument in terms of the transmission of certain semantic and epistemic statuses. Following their approach, when we have grounds for establishing the validity of a set of claims (such as a theory), the semantic or epistemic properties justifying such validity are transferred to those other claims that cannot be avoided by the former ones, despite their potential apparent lack of them when considered in isolation. In our case, we want to transfer justification towards the use of dispositional concepts and explanations. To achieve this, we need to do two things: first, indicate where we are transferring this justification from, and then, show how the transfer goes.

An indispensability argument of this sort was introduced by Enoch and Schechter (2008), who argue for the indispensability of basic belief formation mechanisms, such as reasoning by modus ponens and inference to the best explanation (see also Enoch, 2011). They observe that certain epistemic projects, that is, long term recurring tasks which are epistemically assessable, are rationally required: if subjects did not engage in them, they would be criticizable as rational agents. They identify four such projects: the explanatory project of making sense of and understanding the world, the deliberative project of making decisions, the project of planning for the future, and the project of self-evaluation. The argument, then, is that if a particular belief-formation mechanism is indispensable for the realization of those projects, there is prima facie epistemic justification for engaging in it, regardless of there not being other means for justifying its application.

While we remain neutral on Enoch and Schechter’s argument about basic belief formation, we consider that it offers a template for a similar argument concerning the indispensability of dispositional thought on the basis of functionality. The argument scheme is as follows:

(P1) We are epistemically justified in engaging in rationally required epistemic projects; we would be epistemically criticizable if we didn’t.
(P2) We ought to commit to any X that is indispensable for engaging effectively in rationally required projects.
(P3) X₁ is indispensable for engaging in rationally required projects.
(C) (a) X₁ is epistemically justified, and (b) we ought to commit to X₁ as a way to engage effectively in our rationally required projects.

Note that neither (P1) nor (P2) make overt claims about the existence of anything; they are thoroughly epistemic. Indeed, they make the assumption that the practical and epistemic domains are interrelated, enabling the possibility that rationally required projects are epistemic projects. The type of commitment at play is practical: rather than committing to the existence of entities of a particular kind, we ought to commit to the projects themselves
An epistemic argument for evolutionary dispositions

—such as explaining or self-evaluating (cf. Enoch & Schechter, 2008)— (this follows from P1), and more narrowly, to the ways in which we can effectively engage in them (this is the stage at which we are left in at P2). Thus, in the case of dispositions, X₁ ought to refer to dispositional claims or reasoning, rather than to the metaphysical existence of dispositional properties; while the rationally required projects in P1-3 ought to refer to those of the scientific (read: theoretical, experimental, etc.) context where the dispositional claim or reasoning is in use. We are therefore left with a schema that infers that our means of reasoning are epistemically sound (C) if they are indispensable for carrying out rationally required projects (P3), regardless of any metaphysical realism of our concepts and claims.

2.2. How to defend that something is indispensable

The crux of the argument lies on (P3), the premise that states that something is indispensable in a given context, and we will focus on this in this section. An indirect way to defend P3 would consist in claiming that dispositions inherit the indispensability of modality more generally, as a specific part of it. If modal thought is itself indispensable for engaging in rationally required projects (Morales Carbonell, 2020), then one could say that all modality is justified, including dispositions. However, this strategy gives at best only a partial justification of dispositions, since there could still be modal alternatives that fulfill the function of dispositions better (e.g., the laws of nature). On the other hand, one could show that dispositions are directly indispensable for the effective engagement in rationally required projects. Ultimately, we think that this direct strategy is better suited to defend the justification of dispositions in the context of biology, so we will focus on it here.

In order to defend premise P3 for dispositions directly, one has to show two things:

1. That the use of dispositional concepts is effectively involved in engaging in rationally required projects.
2. That there are no alternatives to it for doing so, or that it is the best way to do so.

It is easy to see that dispositional thought is effectively involved in rationally required projects. It is involved in explanatory tasks, especially when it comes to explain general patterns rather than particular effects (Cartwright, 1994; Fetzer, 1974), but also in ordinary cases, such as when we appeal to a mirror’s fragility to partially explain why it broke when it fell.³ It is also involved in decision-making and planning, such as when we try to prevent the manifestation of dispositions by modifying the situations in which we situate objects that we attribute those dispositions. For example, we could protect the mirror from our previous example using a rubber frame in order to prevent its fragility from manifesting. Finally, dispositional thought is involved in self-assessment. In particular, many folk psychological terms have dispositional profiles (Fletcher & Kerr, 2009), such as generosity understood as the capacity to show altruistic behavior.

By setting the rationally required projects where dispositions effectively play a role, we are identifying the relevant context (Y), as well as the “appropriate standards” for

³ This implies that dispositional thought can take the place of ‘X’ to begin with, which could be considered a previous step for defending P3. See (McKitrick, 2005) for an argument against the a priori rejection of dispositions.
evaluating their indispensability. We are left with the task of arguing that, in some contexts, there are no alternatives to dispositions that fulfill their role in these projects better than them. For doing this, we need to identify the potential explanatory candidates in view of those standards, before evaluating which of these alternatives is better given the standards identified in the first step. To proceed, we have to consider, for any of the rationally required projects, two cases: first, the case where there is no alternative to fulfill the project but through the application of dispositional thought, and second, the cases where there are other alternatives, but dispositional thought is part of the best way to pursue the project. Showing that there are no alternatives offers the peculiar complication that any claim of this kind can be countered by the possibility that there might be unconsidered choices (cf. Van Fraassen, 1989; Lipton, 2004; Stanford, 2006). However, on the assumption of success by current standards, it might be sufficient that the supposedly indispensable X is the only known extant alternative. Additionally, rationally required projects are special in the sense that, being non-optional, we need to have at least some way to engage in them; in this sense, potential alternatives are not significant unless they are already accessible.4

As we saw above, if there are alternatives to dispositional thought, we will have to show that, for the standards posed by our epistemic project, dispositional thought is the best option among them. In this case we should appeal to the relevant (to the project at hand) theoretical and pragmatic virtues of the dispositional approaches, such as simplicity, unification, or explanatory depth, turning them into an instance of inference to the best explanation (Biggs & Wilson, 2021). Our task in the remainder of this paper is to show how some of these considerations are applied to the case of dispositions in biology. Our argument will be that dispositional thinking in biology is both a viable explanatory framework and that it provides the best explanations in some cases — hence, that it is indispensable for evolutionary biology. Our opponent will have to show that the explanatory advantages of their proposals are greater than those we argue for dispositional explanations.

3. Indispensability of fitness and the principle of natural selection

 Evolutionary biology studies the origin and history of organisms through diverse means, some of which also allow for the prediction of trends. Changes in a trait within a lineage, the origin of a novel character, the change in gene frequencies in a population, as well as speciation and extinction processes, are all evolutionary phenomena under the scope of evolutionary biology. While this demands a myriad of approaches and methodologies, Darwin’s Principle of Natural Selection (PNS) is considered one of its central explanatory means.

The PNS states that:

If organism A is fitter than organism B in environment E, then (probably) A will have more offspring in E than B. (Brandon, 2006, p. 333. See also Brandon, 1978)

4 This seems to underwrite a form of conservatism: we would not dispense with our current methods unless it was shown that there are better alternatives. Famously, Feyerabend (2010) argues against this.
Together with the inheritance of traits, this principle explains the process of *evolution by natural selection*, where the traits of those individuals that confer them advantages increase in frequency in a given population. Here the biological notion of fitness becomes crucial.\(^5\) In evolutionary biology, ‘fitness’ is a property of biological entities that refers to their survival and reproduction in a given environment. It is attributed to individuals, to traits, or to types of either, and it is supposed to ground the fact that some variants, in surviving and reproducing more than others in their environments, spread through populations, in turn changing their composition in the long run (thus, these evolve by natural selection). Defining fitness is therefore a major focus for philosophers of biology interested in how evolutionary biology explains and predicts phenomena. The debate over the nature of fitness is a vast one, and it encompasses a wide variety of ontological and epistemological concerns. Here we focus on whether typical accounts of fitness rely on its dispositionality for explaining selection causally.

Within biological practice, fitness values are typically established through measurements of actual survival and reproduction rates, which originally led to an operational rather than dispositional understanding of the notion (e.g., Lerner, 1958; Dobzhansky, 1970). However, this operational orthodoxy was questioned from a philosophical viewpoint when an argument for the dispositional nature of fitness was introduced in the literature through the propensity interpretation of fitness (Brandon, 1978; Mills & Beatty, 1979). The propensity interpretation of fitness states that fitness is a probabilistic capacity or disposition of individuals,\(^6\) or their traits, to survive and reproduce in a certain environment. The dispositional nature of fitness establishes a contrast with actual survival and reproductive success of biological systems: fitness as a disposition refers to the potential for survival and reproduction rather than to survival and reproduction itself. This distinction was introduced in order to avoid the so-called ‘circularity problem’ of the theory of evolution by natural selection (Popper, 1974),\(^7\) according to which identifying fitness with actual survival and reproductive success renders the PNS vacuous: if adaptations are explained by the PNS, according to which those organisms that are fitter will leave more offspring, then ‘fitter organisms’ cannot be identical to ‘organisms that survive and reproduce more’. Otherwise, we would have to admit that adaptations are explained by the fact that organisms that leave more offspring simply survive and reproduce more. This is, of course, not what Darwin had in mind when he introduced the PNS, nor what evolutionary biologists do when they claim that natural selection explains a particular adaptive change. Consider this introduction by Darwin himself:

> if variations useful to any organic being do occur, assuredly individuals thus characterised will have the best chance of being preserved in the struggle for life; and from the strong principle of inheritance they will tend to produce offspring similarly characterised. This principle of preservation, I have called, for the sake of brevity, Natural Selection. (Darwin, 1859, p. 127)

---

\(^5\) Darwin introduced the concept of fitness only once in *The Origin of Species*, where he mostly referred to “advantages” and “profitable variation” of individuals in other passages. However, in linking such an idea with higher success in what he called “the struggle for existence”, he set the ground for making fitness a central evolutionary notion, later popularized by the founders of population genetics.

\(^6\) In opposition to ‘surefire’ dispositions, or those that invariably manifest in the same way under the same circumstances.

\(^7\) Sober refers to this as the analyticity of natural selection (Sober, 2000, p. 70).
In other words, the PNS does not merely state that some individuals reproduce more than others, but it relates such a phenomenon to those individuals having “useful” variations. More generally, the PNS relates the causes of fitness differences with its consequences (Sober, 1984). That is, it connects the ecological fact that there are differences in the adaptedness (in the “usefulness” of variations) of individuals and the evolutionary phenomenon of differential survival in populations. Thus, what biologists have in mind when invoking Natural Selection in their explanations is that certain properties make some organisms perform better in a given environment, and these organisms tend to leave more offspring than others in those environments. It is the actual preservation and accumulation of such advantageous (or ‘fit’) properties in populations that generate evolutionary change. The former (i.e., the properties grounding these tendencies) has been called in the literature the ‘vernacular’ sense of fitness, and it is identifiable with a disposition; while the latter (i.e., the actual success) is referred to as the ‘realized’ sense of fitness (Triviño & Nuño de la Rosa, 2016), identifiable with the manifestation of the disposition once triggered.

One objection to this argument is that the propensity interpretation of fitness is particularly vindicated by the causalist side of the larger philosophical debate over evolutionary mathematical models (Rosenberg & Bouchard, 2005; Millstein et al., 2009). In a nutshell, causalists consider that population genetics models represent the causes of evolution, such as selection and drift, and therefore that fitness ascriptions in such models are propensities referring to causal properties of organisms. For statisticians, however, these models represent statistical measures that are independent of the causes of evolution, and thus fitness is just ‘predictive’ in the models rather than ‘vernacular’ (Matthen & Ariew, 2002; Walsh et al., 2017). For developing their argument, nonetheless, statisticians distinguish between two distinct epistemic projects that the notion of fitness fulfills. The ‘predictive’ sense of fitness is involved, for statisticians, in non-causal explanations of evolutionary change, and as such it is not a propensity but a statistical measure. From this point of view, fitness as in the models of population dynamics is certainly not a real disposition but merely an operational notion. But this stance is directly connected to their belief that the models don’t relate to the real causes of evolution, and thus don’t engage in causal explanations. The ‘vernacular’ sense of fitness, on the other hand, is involved in causal, organismal-level explanations of evolution in terms of the Darwinian PNS (Walsh et al., 2017). This refers to a different epistemic project within evolutionary biology whatever, where population dynamics models play no role. This distinction is most evident in the recent summary of their views by prominent statisticians Walsh, Ariew and Matthen (2017), who contend that fitness has two meanings: as the growth rate of an abstract type; or as a causal property of token organisms, which they regard indeed as a propensity (Walsh et al., 2017, p. 3). Thus causalists and statisticians agree that the causal sense of fitness needs to refer to a propensity or disposition of organisms to survive and reproduce if it has to play some role in the PNS. The disagreement between the two is merely on whether population dynamics models provide a causal understanding of such a principle, or just refer to mathematical truths. In turn, the idea that an operationalist view of fitness can fulfill the role that it is supposed to play in natural selection (causal) explanations can be discarded.

Once operationalism has been ruled out (in the project of causally explaining in terms of the PNS), we need to consider other ways of reducing fitness to non-dispositional notions, and particularly to its causal bases. To be sure, philosophers have proposed non-pro-
pensity causal accounts of fitness (e.g., Abrams, 2007; Bourrat, 2017). However, these proposals reject the association between fitness and the propensity interpretation of probability, but they do not entail any support to the idea that fitness can be reduced to its causal bases. In particular, they do not plead for a substitution of “fitness” for the properties realizing it in specific circumstances. Taking the “mechanistic probability” view of fitness (Abrams, 2007), we can posit that measurements of fitness values depend both on causal properties of organisms and on the probability distribution over circumstances they can encounter. However, fitness is still not reduced to any particular causal basis in this interpretation. As a causal property (of an organism, of a setting, of a trait), fitness is still identified through the kind of manifestation it would show when realized (namely survival and reproduction), thus falling into the category of a disposition, even if not interpreted as a propensity in the philosophy of probability sense.

While it is prima facie clear what type of manifestation we are expecting from a disposition such as fitness, it is not always so when it comes to identifying the proper stimulus conditions for it. What do biologists have in mind as triggers when they speak about the fitness of an organism? Is it just being in a certain environment, or does it involve something more specific? This is indeed one of the problems associated with the propensity interpretation of fitness, namely that specifying the stimulus conditions and the relevant environment for it can be problematic (Abrams, 2009). But, as a matter of fact, this ambiguity is the case for most conventional dispositions, where the particular conditions for their manifestation are not typically specified. When we talk about the fragility of glass, we usually leave unspecified the many triggering conditions that can lead to the glass breaking: being dropped, suffering from an earthquake, being hit by a rock, etc. When analyzing dispositional attributions, philosophers have tended to focus on so-called canonical dispositions, where a specific set of stimulus and manifestation conditions are considered (Choi & Fara, 2021), such as fragility as the disposition to break when hit by a rock. However, in everyday usage, conventional dispositions are thought to be “multi-track” versions of canonical ones (Vetter, 2013), or “capacities” that encompass an undefined set of them (Cartwright, 1994).

This is the case for fitness. When biologists point at a trait, or a type of organism, being fitter than other, they are not specifying, nor willing to, any set of conditional situations that lead to their greater survival and reproduction. At most, they are pointing out that there are more situations that lead to this particular outcome, or that the situations that do so tend to take place more often. Crucially, this is an aspect in direct connection to the (indispensable) explanatory role of fitness. When we say that fitness explains differences in survival and reproduction, grounding the PNS, we are pointing at a connection between the properties of biological systems and survival and reproduction. If white fur is fitter than brown fur in mammals living in snowy environments, it is not because of a higher abundance of counterfactual conditional situations leading to the higher survival of white animals. Rather, it is because of a nomic relationship between the whiteness of the fur and sur-

---

8 Still other causal views are possible. For example, Millstein (2006) explicitly avoids using the term ‘fitness’ when discussing natural selection in order to “prevent entanglement in disputes over [its] definition” (p. 643). To this, Pence and Ramsey (2013) correctly reply that, in order for her view to work, she needs to acknowledge a relevant “causal connection of a physical difference with survival and reproduction”, and therefore that it must fall “in an implicit reference to a concept of fitness” (p. 855).
vival, which is expressed in the dispositional notion of ‘fitness’. It is important to remark here that it is this nomic-like relation that makes the PNS explanatory: the “survival of the fittest” is the survival of those whose properties are causally responsible for higher survival and reproductive success. The use of fitness rather than a hypothetical set of counterfactual conditionals is a way to point at the source of relevant explanatory factors for survival and reproduction: they are in the bearers of fitness rather than in external aspects, such as particular triggers or antecedents of counterfactuals.⁹

This brings us to our last reductive possibility: to lower-level causal bases or mechanisms. White fur mammals are fitter in snowy environments, and we may say that it is the whiteness of fur that realizes fitness in this particular example. Can we still make natural selection explanatory just pointing at differences in fur whiteness instead of differences in fitness? We could do so in a narrow sense if our goal is to explain particular short-term changes in populations on the basis of fur color, but only by adding an explicit connection between whiteness and survival capacity. In this case, pointing at fitness differences based on whiteness is simpler and more explanatory than pointing to whiteness plus its explicit connection to survival. But, more importantly, it allows for the application of the same explanatory means to different traits or organisms, since the causal basis for fitness differs widely across them.

Like most dispositions, fitness is multiply realizable, meaning that it can be instantiated in a variety of causal bases (Prior et al., 1982; McKitrick, 2003). Fragility, for example, can be instantiated in the physical structure of glass, but also in that of porcelain and of eggshells. Similarly, the fitness of a squirrel in a particular environment is realized by different categorical properties than the fitness of an ameba in its own environment. This is not as trivial a point as it might seem. An unquestionable virtue of natural selection explanations is that they can be generalized to any biological system: they all have a certain capacity to survive and reproduce that is causally relevant for the evolutionary fate of their kind. The unificatory potential for fitness resides precisely in not being reduced to the particular categorical bases that ground it: it explains the widespread biological phenomenon of adaptation. Whether it is useful to be more fine-grained in particular situations does not undermine the fact that considering fitness as a general disposition plays a key explanatory role in evolutionary biology. For example, a reduction of fitness to its causal bases would render claims over some species being fitter than others completely meaningless. Prima facie, the whiteness of one particular mammal species cannot be compared with the swimming speed of fish without establishing how much these properties contribute to the fitness of their bearers.

In sum, neither mechanistic explanations of ecological interactions, nor operationalists predictions of evolutionary trends fulfill the explanatory requirements of the PNS. The PNS signifies the broad theoretical idea that differences in how well organisms perform in their environments change the composition of populations. This idea plays a pivotal role

---

⁹ Notice that even accounts that emphasize the importance of circumstances external to organisms for defining fitness measures, such as Abrams’ “organism circumstance” probability view (2007), still fall under this consideration. The reason is that, while the expected number of offspring of a particular organism is highly dependent on the potential circumstances it may encounter, differences in such expected numbers among individuals in a population will still depend on organismal differences. It is these differences that matter for the PNS to be explanatory.
in preserving evolutionary biology as a robust theoretical domain capable of generalizable causal explanations, independently of whether one is a causalist or a statisticalist about populational models, and thus it is epistemically justified on functional grounds (P1). We believe that our analysis shows that explaining through the PNS requires that fitness is a dispositional property (P3). Thus, applying our argument above to the present case, we have that:

(P1) We are epistemically justified in engaging in causal explanations that make use of the PNS.

(P2) We ought to commit to any notion that is indispensable for engaging effectively in PNS causal explanations.

(P3) Causally explaining through the PNS requires that fitness is a dispositional notion.

From these premises, it follows that the dispositional nature of fitness is epistemically justified and that we should commit to it as a way of engaging with the PNS (C).

Importantly, it follows from our discussion here that the use of the disposition fitness is more or less epistemically justified depending on the explanatory context. Thus, by no means our analysis implies that one should reject other causal views of fitness or an operationalist view of it in all circumstances. What it does imply is that considering it a disposition is justified in many, if not most, explanatory contexts where it is made use of, and particularly when it is used for causal explanations in terms of the PNS.

4. Generalizing the argument to other evolutionary dispositions

Given the idiosyncrasy of this argument, one may wonder whether the justification of fitness as a disposition is an exception rather than the rule in scientific explanations, and in evolutionary biology in particular. In this section, we show that this is not the case by referring to the explanatory role of other dispositional notions in evolutionary biology.

Philosophers of biology have been recently paying attention to evolutionary disciplines that study the way variation is generated, and, as a matter of fact, these studies make large use of dispositional notions. The most salient example is evolvability (Wagner & Altenberg, 1996), which is defined in different ways in the literature, but broadly refers to the capacity to generate viable —or potentially adaptive— variation and is explicitly understood in dispositional terms (Brigandt et al., 2023). Similarly to the case of fitness as a disposition, the notion of evolvability was introduced in order to solve a lacuna in the explanatory capacity of classical evolutionary theory: it fell short in explaining the evolution of organization and complex phenotypes (Wagner & Altenberg, 1996; Dawkins, 2003; Wagner & Draghi, 2010). For such evolution to occur, variation has to fulfill some requirements, and biological systems differ in how, and how much, the variation they generate does so. Heritable changes in phenotypes come in many different ways, such as a change in the chemical composition of a particular substance, a change in the shape or size of an organ, or a change in the timing of a life-history trait like reproductive age. Mutations and genetic recombinations are ultimately responsible for these phenomena, but the fact that these produce viable changes in distinct ways in different biological systems remained poorly explained before the rise of the notion of evolvability (Wagner & Draghi, 2010). Impor-
tantly, evolvability claims are not claims about how variation actually behaves, but about how systems can produce such variation under particular circumstances. Indeed, in introducing the idea, Wagner and Altenberg (1996) stressed this dispositional nature by explicitly associating evolvability to a certain kind of variability rather than variation:

> it is essential to clearly distinguish between “variation” and “variability,” even though these words are often used synonymously in the literature. The term variation refers to the actually present differences among the individuals in a population or a sample, or between the species in a clade. Variation can be directly observed as a property of a collection of items. In contrast, variability is a term that describes the potential or the propensity to vary. Variability thus belongs to the group of “dispositional” concepts [...] [the] variability of a phenotypic trait describes the way it changes in response to environmental and genetic influences. (Wagner & Altenberg, 1996, p. 969)

Thus the general understanding is that both variability and evolvability (as a special kind of the former) are dispositional properties that need to meet some triggering conditions to be manifested, such as mutational perturbations inputs or genetic recombinations caused by selective pressures. This shows that evolvability is not a categorical property of biological systems, since it could be the case that the specific triggering conditions are not met. For example, if an animal species that is very prone to evolve feathers and wings due to its developmental resources —and given the right selective pressures— goes extinct for being hit by a meteorite before producing them, we may say that the species was evolvable but that it never met the appropriate triggering conditions to show such evolvability.

Operational understandings of evolvability are of course possible, as is the case for the quantitative genetics sense of the term, that associates it with a specific way to measure standing genetic variation (Houle, 1992) —that is, the manifestation of the disposition. However, there are two important things to remark about this statistical sense of evolvability. One is that the operationalism here refers to the way evolvability is measured, rather than what explanatory role it is supposed to play: even those who commit to the idea of measuring evolvability through its manifestation define it as a disposition; for instance, as the “ability of a population to respond to selection” (Houle, 1992, p. 195). On the other hand, the situation of the operationalist notion of evolvability is analogous to the “predictive” sense of fitness of population genetics models discussed in the previous section. While it may be the case that the operational notion is detached from a causal understanding of evolvability in some predictive models, there are theoretical contexts —even within quantitative genetics— where the notion needs to be understood in dispositional terms (Houle & Pélabon, 2023). This is the case, for example, when considering the very evolutionary explanation of differences in evolvability (Hansen, 2006).

The extent to which the idea of evolvability can be reduced to some set of causal bases is also highly problematic. Evolvability can be manifested in a variety of ways and is realized in a wide diversity of causal bases, such as robust developmental systems, mutational mechanisms, a specific developmental architecture, or quantitative traits with high additive genetic variance (Brigandt et al., 2023). Importantly, the central explanatory role that evolvability plays is completely independent of the particular mechanisms realizing it in different systems. While studying these mechanisms surely enhances our understanding of evolvability in particular systems —like the study of ecological mechanisms enhances our understanding of fitness in particular organisms— considering them realizers of evolvability
fosters the study of the connection between causal properties of biological systems and variation at a theoretical level. What makes all these causal bases important in this theoretical context is that they contribute to the **manifestation** of evolvability, namely to the production of potentially adaptive variation.

Other related notions referring to variation play key explanatory roles in dispositional terms: versatility (Roth, 2005), developmental variability (Hallgrimsson & Hall, 2005), modularity (Schlosser & Wagner, 2004), robustness (Wagner, 2008) and plasticity (Pigliucci, 2001) are chief among them in both scientific and philosophical discussions. All of these properties refer to how biological systems *can* vary given certain conditions (Austin & Nuño de la Rosa, 2018). **Developmental variability** is a general term that refers to the capacity of developmental systems to vary: it includes the possible outcomes, realized or not, of variation (Nuño de la Rosa & Villegas, 2022). The variability of a system will depend on how reachable are different phenotypes for a given developmental system —i.e., on how easily genetic and environmental changes will lead a particular developmental system to a phenotypic change. **Modularity** is the capacity to generate modular variation, namely variation that does not correlate in different traits, allowing them to change relatively independently. For example, changes in the genes regulating limbs in tetrapods will result in a change in all limbs for most tetrapod species, where limbs act as one module. However, in some groups this correlation has been lost —e.g., in humans— and fore and hindlimbs change modularly, or independently. **Developmental or phenotypic plasticity** is the capacity of genotypes to express their phenotype differently under different environmental conditions (Pigliucci, 2001), while **robustness** is the capacity of phenotypes to maintain themselves under mutational inputs (de Visser et al., 2003). Indeed, it has recently been argued that these developmental properties are dispositions in an evolutionary sense insofar as they specify as their manifestation “a demarcated morphospace” in evolution (Austin, 2017, p. 379. See also Brigandt, 2007; Austin & Nuño de la Rosa, 2018). Finally, in paleontological studies, versatility refers to the phenotypic diversification of traits, which is interpreted as supporting the idea that some traits have greater capacity than others to vary under different environmental circumstances (Roth, 2005).

One important aspect of this family of related dispositions is that it allows for the development of theoretical connections among them precisely because of this dispositional nature. Take, for example, the well-studied relationship between evolvability and modularity (Wagner & Altenberg, 1996; Pavličev & Hansen, 2011). While different operationalizations of evolvability and modularity may somewhat be disconnected from each other, theoretical work in different disciplinary contexts does imply a connection between modular systems and evolvability. The basic idea is that a (moderate) decoupling of how different phenotypic traits vary within the same organism enhances the probability that random variation will be beneficial. Importantly, this association between the capacity to decouple variation and the capacity to produce beneficial variation is independent of the particular mechanisms realizing these capacities in specific biological systems (Villegas et al., 2023).

---

10 Interestingly, the opposite idea of modularity, **integration**, has also been vindicated as a disposition, even if the term does not take the form of a dispositional notion, in order to avoid the conflation in evolutionary studies between correlation patterns (manifestation) and developmental causes of such patterns (disposition itself) (Hallgrimsson et al., 2009). Perhaps one could think of modularity and integration as modalities of variability.
One might argue against these examples that the theoretical framework sustaining these dispositional notions is not as robust as the PNS discussed in the previous section. The fact that variation is structured by internal properties of biological systems is not as historically a central explanatory component of evolutionary theory as the PNS. As a matter of fact, referring to these phenomena in non-dispositional terms seems to be detached from such theoretical ideas, like when phenotypic plasticity is reduced to the idea of statistical norms of reaction, or when evolvability is equated to its manifestation in operationalized notions of it, as mentioned before. However, notice that these positions are associated with epistemic projects where the study subject is not variation itself. In the case of statistical evolvability, this is used for predicting quantitative responses in phenotypes to natural selection in the short term. However, when it comes to explaining the differences in evolvability, both synchronically and evolutionarily speaking, the explanatory endeavor demands recognizing evolvability as a dispositional property on its own (Hansen, 2006). Here, like in the case of fitness and the PNS, the dispositionality of variational properties such as evolvability is epistemically justified by virtue of the specific explanatory context considered within evolutionary biology.

5. Concluding Remarks

In this paper, we have argued that the use of dispositions is epistemically justified in evolutionary biology by showing their indispensability for carrying out basic explanatory evolutionary projects. In applying our indispensability argument, we have come to see that dispositional notions such as fitness, evolvability, or modularity are after all justified in an epistemic sense. This justification is inferred because dispositions allow evolutionary biology explanations to be of a certain kind (particularly, causal). Even if we have constrained ourselves to the evolutionary sciences, it is worth recalling here that the use of dispositions is widespread in the life sciences more generally (Hüttemann & Love, 2011; Hüttemann & Kaiser, 2018; Suárez, 2023). Expanding our argument to dispositions in, e.g., molecular, cell biology or ecology, would be quite an interesting pathway to follow in subsequent works.

Although our argument does not refer to the epistemic mechanisms through which dispositional attributions are made—that is, it does not concern the first strategy outlined in Section 2—it points at some consequences regarding them. In particular, since they play a better explanatory role in some epistemic contexts, which justifies their use, they are plausible instantiations of inferences to the best explanation in those particular contexts (Lipton, 2004; Biggs & Wilson, 2021). Note that variational dispositions (section 4) actually have no opposing modal views to consider, perhaps with the exception of operationalist views that nevertheless fulfill a different explanatory role than the disposition. Fitness, however, faces different interpretations in the literature, and we have shown that those still demand understanding it as a disposition if it is to fulfill a causal interpretation of the PNS. This may point to some epistemic virtues of dispositions, such as their simplicity or their unificatory potential (see Rozeboom, 1984). Considering that skepticism about dispositions is typically motivated by empiricist views, we regard our vindication in evolutionary biology—an empirical science—as a potential starting point for thinking about the indispensability of dispositions more generally.
Acknowledgements

We thank Vanessa Triviño and an anonymous reviewer for very helpful comments to previous versions of this paper. CV’s work was funded by national funds through FCT — Fundação para a Ciência e a Tecnologia, I.P., in the R+D Center for Philosophy of Sciences of the University of Lisbon (CFCUL), strategic project FCT I.P. UIDB/00678/2020 and UIDP/00678/2020, and through the contract 2021.03186. CEECIND/CP1654/CT0008. FM’s work was partially funded by ANID (Chile), through a doctoral scholarship and the project FONDECT/Postdoctorado 3220017.

REFERENCES


https://doi.org/10.1387/theoria.25137
Cristina Villegas, Felipe Morales Carbonell


An epistemic argument for evolutionary dispositions


https://doi.org/10.1387/theoria.25137


**Cristina Villegas** is a postdoctoral researcher at the University of Lisbon. She works on causation, probabilities, and chance in evolutionary biology explanations.

**Address:** Centro de Filosofia das Ciências, Departamento de História e Filosofia das Ciências, Faculdade de Ciências, Universidade de Lisboa. Campo Grande, Edifício C4, 3º Piso, Sala 4.3.20. 1749-016 Lisboa, Portugal. E-mail: cvillegas@fc.ul.pt - ORCID: https://orcid.org/0000-0002-6402-5288

**Felipe Morales Carbonell** is a postdoctoral researcher at Universidad de Chile. He works on know-how, abilities, imagination, and modal epistemology.

**Address:** Universidad de Chile, Departamento de Filosofía, Facultad de Filosofía y Humanidades. Avda. Capitán Ignacio Carrera Pinto 1025, 7800284, Ñuñoa, Santiago. E-mail: ef.em.carbonell@gmail.com - ORCID: https://orcid.org/0000-0001-5492-0759