Natural Selection and Contrastive Explanation

Joeri Witteveen*†

This article defends the Negative View of natural selection explanation, according to which natural selection cannot explain of any given individual why it has the traits it does. Over the years, this view has been criticized on empirical, metaphysical, and explanatory grounds. I review the debate and offer additional reasons for rejecting the empirical and metaphysical objections. The explanatory objection, which holds that the Negative View is rooted in a flawed account of contrastive explanation, initially seems plausible. However, I argue that a closer consideration of the desiderata of contrastive explanation shows that this objection fails as well.

1. The "Negative View" of Selectionist Explanation and Its Critics. Consider the contrastive explanandum "Why does Rose always have blue-eyed partners rather than brown-eyed ones?" Suppose the answer is simply that Rose has a strong preference for blue-eyed partners over brown-eyed ones. At first glance, it might seem that this answer to the question about all of Rose's partners is also an answer to the singular variant "Why is Rose's current partner, Thom, blue-eyed rather than brown-eyed?" However, a moment's reflection teaches us that this is incorrect. Thom was already blue-eyed before Rose set her eyes on him, and he would have been blue-eyed if Rose's partner preferences had been different. And if Rose's partner selection regime does not make a difference to Thom's eye color, it does not explain why his eyes are one color rather than another. Thus, although Rose's partner selection regime

Received March 2018; revised December 2018.

*To contact the author, please write to: Department of Science Education, Section for History and Philosophy of Science, University of Copenhagen, Denmark; e-mail: jw@ind.ku.dk.

†I would like to thank the audience of the Theoretical Philosophy Colloquium at the Department of Philosophy and Religious Studies of Utrecht University for helpful feedback on an early version of this work. This article benefited greatly from the advice of three reviewers and from the incisive and constructive comments of Jonathan Birch.

Philosophy of Science, 86 (July 2019) pp. 412–430. 0031-8248/2019/8603-0002$10.00
Copyright 2019 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

helps explain a contrastive fact about a population (why all her partners have been blue-eyed rather than brown-eyed) it does not help explain a contrastive fact about a given individual in that population (why Thom is blue-eyed rather than brown-eyed).

Elliott Sober argued that this fact about the explanatory scope of everyday artificial selection explanations holds for selectionist explanations in general—that is to say, it also holds for natural selection explanations (Sober 1984).[1] Although natural selection can explain facts about the spread of traits in a population, it cannot explain of any particular individual in that population why it has the traits it does. This claim about the explanatory scope of natural selection explanations has become known as the Negative View of selection (Pust 2001). We can formulate it in more precise terms as follows:

> **Negative View:** In any population P with individuals $O_1, \ldots, O_n$ in which there has been natural selection for trait T over T', this process of selection cannot explain of any individual O_i that is born in P why it has trait T rather than T'. Only inheritance and development can explain of an individual O_i why it has the traits it does.

Through the years, many critics have contended that the Negative View is wrong. They have argued for the Positive View that natural selection explanations have a more inclusive explanatory scope than artificial selection explanations; the former can (at least under certain conditions) explain why individuals severally have the traits they do. These critics can be classified into three groups. Some have argued that the apparent plausibility of the Negative View depends on *empirical* assumptions about organismal development that do not apply universally. They claim that the exceptions point to a possibility for selection to explain why a given individual has the traits it does. Others have criticized the Negative View for relying on a questionable *metaphysical* assumption, which, once discarded, opens up a space for the Positive View. Finally, there are those who have claimed that the Negative View results from applying flawed *explanatory* criteria. They argue that the adoption of improved criteria for contrastive explanation show why the Positive View is true after all.

I will be brief about the first two lines of objection, which have already been discussed at length in the literature. However, I will give additional reasons for thinking that both objections miss the mark (secs. 2 and 3). My main concern will be with the third line of objection, which, by pointing to issues about the nature of (contrastive) explanation as such, presents the most direct and targeted challenge to the Negative View (sec. 4). Moreover, I will show that an important objection of this kind has not been rebutted by defenders of

1. Nozick (1974) and Cummins (1975) made similar observations, but Sober (1984, 1995) was the first to spell out and defend this view of selectionist explanation in detail.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

the Negative View and therefore deserves close scrutiny (sec. 5). Yet, I will show that ultimately this objection fails too (sec. 6). The charge that defenders of the Negative View have relied on a flawed account of contrastive explanation turns out to be based on a misconstrual of contrastive explanation on the part of critics.

2. The Empirical Objection Reconsidered. Lewens (2001) outlined a scenario in which it appears that natural selection helps explain why a given individual has the traits it does: "Suppose that for a given trait *t* to develop normally the individual's environment must contain a certain proportion of conspecifics with the same (or a different) trait. Suppose also that selection maintains that proportion. Then we can say that changing the selective regime will alter the proportion of conspecifics with these traits, and will disrupt the development of the imagined trait in the individual. . . . Selection, by ensuring the maintenance of the necessary developmental environment for the trait, does partially explain why an individual possesses that trait" (597). Interestingly, after outlining this scenario Lewens downplays its implications for the Negative View. He argues that since defenders of the Negative View have only argued against the relevance of natural selection for explaining why an individual possesses a particular genotypic trait, his scenario of natural selection affecting an individual's phenotypic traits does not challenge the Negative View (597).

McLoone (2013) has responded that we should take Lewens's example more seriously than Lewens himself does. He finds no textual evidence for Lewens's claim that the Negative View has only been defended in relation to the explanation of an individual's genotypic traits. Most defenders of the Negative View write about the explanation of an individual's traits as such, without restricting their view to the genotypic level. Thus, McLoone concludes that Lewens's scenario presents a genuine challenge: "What was wrong with traditional versions of the Negative View is that they ignored that phenotypes can change in response to social environments" (345). This refutes the "traditional" version of the Negative View. Restricting the view to the explanation of genotypes requires a substantial revision of the original thesis that Sober and others defended.

McLoone is right to observe that the Negative View has been predominantly presented as holding for genotypic and phenotypic traits alike; my rendering of the Negative View in section 1 reflects this "traditional" interpretation. However, I submit that McLoone is wrong to conclude that this interpretation is in trouble because of scenarios of the sort Lewens has pointed to. He has overlooked a distinction in the background of the debate over the scope of selectionist explanation, which tells us that the Negative View does not need to be revised to deal with scenarios of the sort that Lewens drew attention to.

Remember that the Negative View identifies a difference between the explanatory scope of (natural) selection *versus* that of development and inheri-

https://doi.org/10.1086/703573 Published online by Cambridge University Press

tance. Whereas selectionist explanations have a population-level explanatory scope, inheritance and developmental explanations have an individual-level scope. What Lewens's scenario does is identify a variety of developmental explanations that involve a selective component—we could call them "selection-on-development" explanations—and show that these have an individual-level explanatory scope too. But given that these selection-on-development explanations are a variety of developmental explanations, this should not be a surprise. It confirms that whereas developmental explanations can explain of a given individual why it acquired the traits it did in the course of ontogeny, natural selection explanations can explain why one trait rather than another spread in a population between generations. The fact that a population-level state can feature in a contrastive explanation of an individual's ontogeny does not challenge the fact that it is development rather than selection that explains why that individual has the traits it does.

3. The Metaphysical Objection Reconsidered. The second critical response to the Negative View argues that it breaks down because of a crucial disanalogy between artificial and natural selection that bears on the metaphysics of identity. Cases of artificial selection, such as Rose's partner selection, act on individuals that existed before the selective episode and that either continue to exist after the selective episode or are removed from the population by selection. Natural selection, however, often involves the reproduction of new individuals that did not exist before the selective episode. The critics argue that this difference between artificial and natural selection affords the latter with the individual-level explanatory scope that the former lacks. More specifically, they argue that natural selection for trait T over T' can help explain why organism O has trait T rather than T' because the following counterfactual is true: if the selective regime had favored trait T' over T, O would have inherited trait T' rather than T by being born to a numerically different, T'-bearing parent. In other words, the idea is that natural selection can help contrastively explain the traits of a given individual because a change in selective regime can make a difference in whom it inherits its traits from. If this is right, then a Positive View of natural selection seems in reach (Neander 1988, 1995a, 1995b; Matthen 1999).[2]

Notice that this response challenges the Negative View on a metaphysical plane, by taking issue with an (implicit) assumption about metaphysics of iden-

2. This counterfactual expresses a general condition that the critics agree on; they disagree on how to fill in the specifics. Neander (1988, 1995a, 1995b) has contended that for natural selection to explain an individual's traits it also needs to have been cumulative and must have operated on compound traits. Matthen (1999, 2003) has suggested that sexual reproduction is required. Subsequent discussion has clarified that none of these further conditions are essential (e.g., Walsh 1998; Lewens 2001; Forber 2005). The pivotal move is the idea that reproduction opens up the counterfactual scenario in which a focal individual does not have its actual parent(s) as parent(s) (Pust 2001, 2004).

https://doi.org/10.1086/703573 Published online by Cambridge University Press

tity: the idea that organisms necessarily have their actual parents as parents. This stance on the metaphysics of identity is known as *origin essentialism* (Kripke 1980; Pust 2001). If origin essentialism is true, a difference in the selective regime cannot cause O to inherit T' rather than T by being born from numerically different parents. But if origin essentialism is false, this possibility cannot be excluded.

Why should we reject origin essentialism? The critics have argued that a scientific account of natural selection explanation should be free from everyday intuitions about the metaphysics of biological identity (Matthen 1999, 2003; Forber 2005). Defenders of the Negative View have justifiably pointed out that this response fails. If the metaphysical critics wish to make a case for the Positive View, they cannot just throw out one view on the metaphysics of identity without offering an alternative. They will need to propose a theory of transworld identity that supports the Positive View. That is, they must specify the criteria by which certain individuals other than an individual's actual parents can and other cannot qualify as its counterfactual parents (Lewens 2001; Pust 2004; Helgeson 2015).

I submit that the challenge the metaphysical critics face is even bigger. Even if they were able to offer a plausible alternative view about which individuals could and could not have been our parents, this would not suffice to vindicate the Positive View. To that end, the critics would also have to show that a counterfactual selective regime in favor of trait T' can be the cause of an actual individual being born from numerically different T'-bearing parents. Only then have the critics shown that a difference in natural selection can make a difference to the trait a given individual has.

So far, the metaphysical critics have not presented a theory of the metaphysics of identity that even begins to address these questions. The burden is on them to show that there is a plausible view that speaks in favor of the Positive View.[3] Until then, the defenders of the Negative View have no reason to worry about the metaphysical objection.

4. The Explanatory Objection. The third line of criticism of the Negative View is prima facie the most promising. The explanatory critics avoid the pitfalls of the other reponses by recognizing that the debate over the scope of natural selection explanation requires us to distinguish it from developmental explanation and by conceding that origin essentialism is a plausible view about the metaphysics of identity. The explanatory critics instead argue that the Neg-

3. Although origin essentialism has also been criticized in the broader metaphysics literature, it is doubtful whether the defenders of the Positive View can draw on these discussions. As Sartorelli (2015) has recently pointed out, debates in metaphysics have problematically conflated issues about identity for biological and nonbiological cases. It is far from obvious that arguments against origin essentialism for chairs and tables apply equally to reproducing biological entities like organisms.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

ative View is rooted in a rudimentary conception of explanation. Better philosophically motivated criteria of explanatory relevance show that artificial and natural selection explanation come apart in a way that favors the Positive View.

Nanay (2005, 2010) and Birch (2012) have developed arguments along these lines. Nanay's argument has already been rebutted successfully by Stegmann (2010a, 2010b) and Helgeson (2013). They have shown that Nanay misinterprets the modal implications of the counterfactual dependence claims, which he thinks speak in favor of the Positive View. I concur, but in what follows I provide a further diagnosis of what I take to be a more fundamental misunderstanding about the nature of contrastive explanation that underlies Nanay's errors. With this misunderstanding in view, we can also see where Birch's account—which has not received any convincing responses—takes a wrong turn. That will be the topic of section 5.

4.1. Nanay on Explanation and Counterfactuals. Nanay (2005) argues that for any environment in which resources are limited and in which there is a selective gradient over two or more different trait types, natural selection can help explain of an individual *O* why it has the traits it does. He reasons as follows: "The fact that *O* inherited trait *A* from her mother is explanatorily relevant to *O*'s having trait *A*. But, *O* could not have inherited *A* from her mother unless her mother had reached reproductive age. Further, *O*'s mother could not have had trait *A* and reached reproductive age unless *O*'s uncle (who did not have trait *A*) had died. Therefore, the fact that *O*'s uncle (who did not have trait *A*) died is explanatorily relevant to the fact that *O* has trait *A*" (Nanay 2005, 1104). Nanay concludes from this that *O*'s possession of trait *A* depends counterfactually on selection favoring *O*'s mother over *O*'s uncle. Therefore, selection helps explain why *O* has *A*.

Stegmann (2010b) has pointed out that this formulation of what natural selection explains fails to capture the contrastive nature of the explanandum at issue. What needs explaining is not why *O* has trait *A* but why *O* has trait *A* rather than *B*. In response, Nanay (2010) refined his argument and presented it in the following revised form:

> (1) *O*'s having *A* (rather than *B*) depends counterfactually on *O*'s mother surviving to reproductive age and having *A*.
>
> (2) *O*'s mother surviving to reproductive age and having trait *A* (rather than *B*) depends counterfactually on selection against organisms with trait *B*. (Nanay 2010, 419; symbols modified)

If 1 and 2 are true, it follows that *O*'s having *A* rather than *B* depends counterfactually on selection against *B*-type organisms. However, as both Stegmann (2010a) and Helgeson (2013) have pointed out, even this revised version of

https://doi.org/10.1086/703573 Published online by Cambridge University Press

Nanay's argument fails because the counterfactual dependence relation in 1 does not obtain.[4]

To see why 1 is false, consider that for any counterfactual dependence relation of the form "*P* depends counterfactually on *Q*" it must hold true that had *Q* not occurred, *P* would have failed to obtain. For counterfactuals like 1, in which *P* is contrastive, we can add that had *Q* not occurred, *P* would have failed to obtain in a particular way: the contrast expressed by *P* must have been reversed. To see why this is a requirement for counterfactuals with a contrastive antecedent, consider a simple example: if Sam being home before rather than after dinner counterfactually depends on his train running on time, it follows that if his train does not run on time, Sam is home after dinner rather than before dinner. The counterfactual in 1 has the additional complication that the consequent *Q* is a conjunction. For a counterfactual of this kind to be true, the contrastive antecedent must be reversed if at least one of the conjuncts is false. Another example helps to illustrate this requirement: if Sam going ice skating rather than roller skating (the contrastive antecedent *P*) counterfactually depends on it freezing outside and on his ice skates being sharp (the conjunctive consequent *Q*) it follows that Sam will go roller skating rather than ice skating if it is not freezing or if his ice skates are not sharp or if both obtain. Now, applied to the counterfactual dependence relation in 1, this means that for it to hold, all of the following statements must be true:

(3) Had *O*'s mother survived to reproductive age and had she not had trait *A*, then *O* would have had trait *B* (rather than *A*).

(4) Had *O*'s mother not survived to reproductive age and had she not had trait *A*, then *O* would have had trait *B* (rather than *A*).

(5) Had *O*'s mother not survived to reproductive age and had she had trait *A*, then *O* would have had trait *B* (rather than *A*).

Statement 3 is true: given the setup of the example, it follows from *O*'s mother not having had trait *A* that she must have had trait *B*, which, given her reproductive success, would have been inherited by *O*. But 4 and 5 are clearly false: if *O*'s mother had failed to produce any offspring, *O* would not have possessed trait *B* rather than *A*. *O* not having been born in those cases, it would have possessed neither *B* nor *A*. Since all of 3, 4, and 5 need to be true for 1 to hold, we can conclude that 1 is false: *O*'s having *A* rather than *B* does not depend counterfactually on a condition that includes its mother's survival as one of

4. Stegmann (2010a) also argues that 2 is problematic, but for present purposes it suffices to show that 1 fails. If 1 fails, Nanay's argument for the Positive View inevitably breaks down.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

its conjuncts. This is sufficient for Nanay's defense of the Positive View to break down, since on his account it is the survival of *O*'s mother that links the explanation of *O*'s having *A* rather than *B* to natural selection (via premise 2).

4.2. Nanay's Error, Contrastive Explanation, and Presupposition. Although I believe that the preceding analysis of Nanay's counterfactuals is correct, it does not strike me as a particularly intuitive and insightful account of where his reasoning takes a wrong turn. Fortunately, there is a more straightforward way of seeing why and how Nanay's argument fails, based on considerations about the nature of contrastive explanation.

In essence, Nanay's argument fails because it violates the requirement that explanantia for contrastive explananda need to satisfy their *explanatory presuppositions*. Consider a request for explanation of the form "Why *P* rather *Q*?" This explanandum presupposes that either fact *P* or foil *Q* occurs (Lipton 2004, 35ff.; Hitchcock 2013). If I ask "Why is Rose watching a movie rather than Thom playing a video game," I am presupposing that either Rose watches a movie or Thom is playing a video game and that there is an explanation for why one rather than another occurred. If you respond that Thom could have played a video game while Rose watched a movie, you are not answering my question, which presupposes that just one of the two activities occurred. You either misunderstood the question, or you are rejecting the presupposition that only one of the two activities could occur. For a response to count as an answer, it must satisfy the presupposition.

A contrastive question of the form "Why does *O* have *A* rather than *B*?" introduces a further presupposition. Apart from presupposing that either *A* or *B* obtains, it presupposes that *O* exists. If I ask why Thom has blue eyes rather than brown eyes, I am asking something about *Thom* only. If you answer by saying "If Thom had not existed, then . . . ," you have misunderstood what I was asking for. My question presupposed that Thom exists and that he has either blue or brown eyes.

It can be easy to overlook explanatory presuppositions because the grammar of contrastive questions often hides them from our view. In English, the presupposition remains partially hidden in sentences with a *rather-than* locution, because these take a noun phrase rather than a clause as their completion (Hitchcock 2013, 13). We ask, "Why does Thom have blue eyes rather than brown eyes" and not "Why does Thom have blue eyes rather than Thom have brown eyes?" Although the latter sentence is ungrammatical, it does a better job at bringing out the presupposition that Thom exists and has one or the other eye color. We can, however, use a different linguistic device to make the presupposition evident. Instead of asking "Why does Thom have blue rather than brown eyes?" we could ask "Given that Thom exists, why does he have blue rather than brown eyes?" The *given* locution makes salient that a particular el-

https://doi.org/10.1086/703573 Published online by Cambridge University Press

ement needs to be satisfied by a response for it to count as a potential answer to the explanandum (Sober 1986).

With these considerations about explanatory presupposition in mind, it becomes easy to see why Nanay's defense of the Positive View fails. If we rephrase the question "Why does *O* have trait *A* rather than *B*?" as "Given that *O* exists, why does it have trait *A* rather than *B*?" it becomes instantly clear that a response of the sort "If *O* had not existed, then . . ." does not qualify as an answer. It fails to satisfy the presupposition that *O* exists and, therefore, fails to answer the question. This is precisely the problem that Nanay's account faces. We have seen that on his analysis a difference in natural selection can only make a difference to the existence or nonexistence of *O*. Hence, natural selection cannot feature in the explanation of why *O* has the trait it has when it exists. Therefore, natural selection must fail as an answer to the question "Why does *O* have trait *A* rather than *B*?"

To sum up, Nanay's critique of the Negative View is unsuccessful because it fails to recognize an elementary requirement on contrastive explanation. A contrastive explanation of a fact about a given organism requires us to presuppose the existence of that organism. Natural selection explanations do not satisfy this presupposition. Natural selection may help explain why an organism exists but not why, given that it exists, it has one trait rather than another.

5. The Explanatory Objection and the Argument from Overdetermination. Like Nanay, Birch (2012) has aimed to vindicate the Positive View by examining the nature of the counterfactual dependence relation between natural selection and the traits an individual has. However, Birch's argument unfolds in an importantly different and at first glance more promising way. Before offering an account of contrastive explanation that he claims supports the Positive View, Birch provides a close analysis of the criterion for explanatory relevance that defenses of the Negative View rely on. He argues that this criterion is flawed because it cannot handle commonplace cases of *causal overdetermination*.

Before considering Birch's arguments in detail, it is worth noting that the responses it has received so far do not do it justice. First, consider McLoone's (2013) claim that we can safely ignore Birch's criticism of the criterion for explanatory relevance that the Negative View is said to rely on, because it is simply what "traditional versions of the Negative View assume." When Birch offers an alternative criterion, he has "changed the rules of the game by changing the sense of 'explanation' used" (343). This criticism misses the point. If Birch is right about there being a viable alternative account of contrastive explanation on which selection is deemed to be relevant to the explanation of an individual's traits, then surely we should not ignore that account—let alone reject it—simply because it goes against "tradition." Instead of brushing Birch's alternative aside for being an alternative, the defender of the

https://doi.org/10.1086/703573 Published online by Cambridge University Press

Negative View should argue that it fails as an alternative. In fact, McLoone's response should impel us to take Birch's argument even more seriously, because it suggests that Birch has at least correctly identified the explanatory criterion that proponents of the Negative View rely on—a criterion that Birch argues is flawed.

A second response to Birch's account has come from Helgeson (2015). In the conclusion of his critique of the metaphysical objection to the Negative View, Helgeson extends his discussion to Birch's argument by pointing out that although it "concedes origin essentialism for the sake of argument . . . origin essentialism is a purported necessary condition for trans-world identity; it is never regarded as remotely sufficient. What the Positive View requires is a sufficient condition. Birch's novel criterion for picking the relevant counterfactual scenario in no way reduces the need for a standard of judging whether an organism in that scenario is the same individual as the actual organism of interest" (28). This response is puzzling, since Birch makes amply clear that he accepts the theoretical and metaphysical assumptions of the Negative View wholesale (2012, 571). The problem of having to provide alternative necessary and sufficient conditions for establishing modal identity conditions is only faced by those who push the metaphysical line of criticism (i.e., those who deny origin essentialism). Helgeson's remarks suggest that he takes Birch to be presenting a variant of the metaphysical response to the Negative View. This is a mistake; Birch criticizes this view only on explanatory grounds. Metaphysically speaking, the plausibility of his account is on a par with the Negative View. Again, the only proper response is to take a closer look at the substance of Birch's argument. Let us do that now.

5.1. Birch's Counterexamples of Causal Overdetermination. Birch proceeds from the observation that all current defenses of the Negative View rely on the following simple counterfactual criterion (SCC) for contrastive-explanatory relevance:

> **SCC:** Cause *C* rather than *C** helps explain why outcome *E* rather than *E** occurred only if: (if *C** had occurred, then *E** would have occurred).

Applied to natural selection as causal factor, this becomes a simple counterfactual criterion of selection (SCCS):

> **SCCS:** Natural selection favoring trait *A* rather than trait *B* helps explain why organism *O* has trait *A* rather than *B* only if: (if selection had favored trait *B* over *A*, then *O* would have possessed trait *B*).

According to Birch, supporters of the Negative View argue that the truth of their view follows from the fact that the SCCS is never satisfied. They hold

https://doi.org/10.1086/703573 Published online by Cambridge University Press

that counterfactual natural selection for *B* over *A* either results in *O* never being born, because selection for *B* is so strong that *O*'s parent or parents die before reproducing, or results in *O* being born without selection affecting which trait it has, because selection favoring *B* does not cause *O* to have *B*: inheritance, mutation, and development do. Either way, natural selection for *B* over *A* fails to cause *O* to have *B*. Therefore, SCCS fails.[5]

Birch agrees that if the SCCS is the right criterion for passing judgment on the scope of natural selection explanation, it follows that natural selection cannot help explain why individuals have the traits they do. However, Birch denies that the SCCS is the right criterion. He argues that the SCC that underpins it constitutes "a proto-theory of contrastive explanation . . . [which] turns out to be highly questionable" (Birch 2012, 571). The reason the SCC fails is that it cannot handle simple "biologically commonplace" cases of causal overdetermination that it should be able to handle (571).

Birch builds his case against the SCC by first considering a nonbiological example of causal overdetermination. Suppose that your train ride from London to Paris gets canceled because of two independent causal factors: the train drivers are on strike and there are leaves on the track. As a result, you are stuck in London (outcome *E*) instead of arriving in Paris (outcome E^*). The strike and the leaves are overdetermining causes of *E*: if one of them had not occurred, the other would have still guaranteed that you remain stuck in London.[6] Birch concludes that both causes fail to satisfy the SCC: for each cause, its contrast does not result in the occurrence E^*. Thus, according to the SCC, neither the leaves nor the strike help explain why you are in stuck in London instead of having arrived in Paris. And yet, we would want to say that both the leaves and the strike help explain why you are in London instead of Paris. The SCC gets it wrong, it seems.

Birch argues that similar cases of overdetermination pervade biology. Consider his example of moth *M*. We would like to know whether the parental genotype of this moth helps explain why it has dark coloration (outcome *E*) rather than pale coloration (outcome E^*). Suppose that the coloration of *M* is fully determined by inheritance of either of two possible alleles: a dark al-

5. The SCCS fails to make allowances for those defenders of the Negative View that have formulated the contrastive criterion in probabilistic terms (Razeto-Barry and Frick 2011; McLoone 2013). To account for these, the SCCS could be rewritten in a probabilistic form, as follows: "Natural selection favoring trait *A* rather than trait *B* helps explain why organism *O* has trait *A* rather than *B* only if: (if selection had favored trait *B* over *A*, then Pr(*O* has trait *B*) > Pr(*O* has trait A))." However, for the remainder of the discussion nothing depends on whether a deterministic or a probabilistic approach is adopted. For sake of simplicity, I therefore focus on Birch's formulation of the SCCS only.

6. It is assumed in Birch's example that the strike and the leaves are time-symmetric causes. If they were time asymmetric (e.g., if the strike had begun before the leaves fell on the track) the example would be one of causal preemption instead of causal overdetermination.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

lele (cause *C*) or a pale allele (cause *C**). Also suppose that there was strong selection for dark over pale coloration in the generation of *M*'s parents, such that had they possessed the pale allele rather than the dark one, they would have died before being able to reproduce. Birch concludes:

> As a result it is *false* that, if *M*'s parents had possessed the pale allele, *M* would have possessed pale coloration. Not-*E** (i.e., the non-instantiation of pale coloration by the offspring moth) is overdetermined by two independent factors (namely, the parental genotype, and the strong selection against the pale allele) in exactly the same way that my non-arrival in Paris was overdetermined by two independent factors. The information that *M*'s parents possessed the dark allele rather than the pale allele consequently fails to satisfy SCC. Hence, either SCC is defective in this context, or parental alleles are genuinely irrelevant to the explanation of offspring traits. The latter possibility seems sufficiently absurd to warrant a preference for the former. (Birch 2012, 572)

In short, Birch argues that we should reject the SCC applied to the inheritance of parental genotypes, since this criterion erroneously judges parental genotypes to have no explanatory relevance to offspring traits. And if the SCC is mistaken about the explanatory relevance of parental genotypes, this gives reason to think that it may be defective in general and, therefore, also in its application to natural selection in the form of the SCCS.

5.2. Overdetermination and Explanatory Relevance. In this section I argue that Birch's analysis fails to show that the Negative View relies on a problematic criterion for explanatory relevance. First, I show that the example of moth *M* does not give us any reason to doubt the soundness of the SCC. Next, I argue that even if we concede that the train cancellation example does present a genuine case of causal overdetermination and thereby throws the SCC into doubt, this does not unsettle the Negative View in any way.

Let us start by taking a closer look at Birch's claim that in the example of moth *M* the nonoccurrence of outcome *E** is overdetermined. What does this outcome not-*E** amount to? Birch describes it as "the non-instantiation of pale coloration by the offspring moth," but this is ambiguous between two different outcomes: one in which *M* exists but does not have pale coloration and one in which pale coloration is not instantiated by *M* because it does not exist. Which of these interpretations supports Birch's claim that not-*E** is overdetermined?

Consider the first interpretation of not-*E**, as "*M* exists but does not have pale coloration." Given that the contrastive setup of the example presupposes that *M* has either dark or pale coloration, *M* existing without pale coloration implies that it has dark coloration. In other words, the first reading of outcome not-*E** equates to the occurrence of outcome *E*. Yet, it is easy to see that outcome *E* is not causally overdetermined. A dark parental genotype and a selec-

https://doi.org/10.1086/703573 Published online by Cambridge University Press

tive regime favoring dark coloration do not independently cause *M* to have dark coloration: a change in selective regime can affect whether *M* exists but not which color it has. In those scenarios in which *M* does exist, the parental genotype is the only factor that is causally relevant to its coloration.

This leaves us with the second interpretation of not-E^*, as "*M* does not exist (and therefore fails to instantiate pale coloration)." For this outcome to be causally overdetermined, selection against paleness and a pale parental genotype must independently guarantee *M*'s nonexistence. Yet, the pale parental genotype clearly does not satisfy this condition. If there had been no selective pressure against the pale parental genotype, such that *M*'s parents could have reproduced, their pale genotype would not have prevented *M* from being born. Instead, *M* would have been born with pale coloration.

We can conclude, then, that Birch was wrong about the example of moth *M*: on neither interpretation of the outcome not-E^* does it present us with a case of causal overdetermination. But there is a further problem with this example. On closer inspection, it appears that it fails to present a genuine test of the SCC in the first place. Remember that the SCC offers a criterion for determining whether the occurrence of *C* rather than C^* answers the contrastive explanandum "Why *E* rather than E^*?" Applied to the case of moth *M*, this comes down to the question whether the occurrence of a dark parental genotype (*C*) rather than a pale parental genotype (C^*) helps explain why *M* has dark (*E*) rather than pale (E^*) coloration. Now, as we have seen in section 4.2, the explanandum "Why does *M* have dark rather than pale coloration?" requires us to presuppose that *M* exists in order to answer why it has one color rather than another.[7] Considerations about counterfactual scenarios in which *M* fails to exist are irrelevant to answering why *M* has a certain color in those scenarios in which it does exist. Yet, the only counterfactual scenario that Birch evaluates is one in which *M* does not exist, since selection has prevented its parents from reproducing. Hence, Birch has not presented us with an actual test of the SCC. In order to test the SCC, we need to consider the counterfactual situation in which *M* exists with pale coloration (E^*) and ask whether the pale parental genotype (C^*) is the cause of *M* being pale (rather than dark). Clearly, the answer is positive: *M*'s inheritance of a pale genotype rather than a dark one caused it to have pale rather than dark coloration. But this means that the SCC gets it right: the occurrence of C^* accounts for the occurrence of E^*.

To summarize, Birch's example of moth *M* fails to show that the SCC is flawed because (*a*) it fails to make a case for causal overdetermination that can be held against the SCC and (*b*) the example as presented by Birch does

7. We could make this explanatory presupposition explicit by writing SCC in the following, logically equivalent form: "Given *x*: (*C* rather than *C** helps explain why *Fx* rather than *Gx*, only if: (if *C** had occurred, then *Gx*))." In the case at hand *x* refers to *M*'s existence, and *F* and *G* are alternative colorations.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

not constitute a test of SCC at all. This second point illustrates that Birch has failed to comply with the basic requirement on contrastive explanation that admissible answers to contrastive explananda need to respect their explanatory presuppositions. Birch has overlooked that a response to the question "Why does *O* have trait *T* rather than *T′*?" needs to presuppose that *O* exists for it to count as an answer to this question.

This leaves us with the example of the train cancellation. Even though this is not an example that involves natural selection, one might argue that it still poses trouble for the Negative View by showing that the SCC (and by implication, the SCCS) cannot deal with cases of causal overdetermination. The worry would be that the defenders of the Negative View have unwittingly drawn on cases in which natural selection is an overdetermining cause. The inability of the SCC to detect the explanatory relevance of natural selection in those instances would then explain why they have arrived at the erroneous conclusion that the Negative View is true.

Whether defenders of the Negative View have made this error should be easy to ascertain. All we need to do is check whether natural selection can affect the traits an organism has in scenarios that control for causal overdetermination. If natural selection is not a cause of an individual's traits in those scenarios, we know that it cannot be an overdetermining cause either. What we should do, then, is consider the following criterion for contrastive explanatory relevance:

> **SCCS′:** Natural selection favoring trait *A* rather than trait *B* helps explain why organism *O* has trait *A* rather than *B* only if: (if selection had favored trait *B* over *A* and no independent factor *D* caused *O* to exist and possess trait *A*, then *O* would have existed and possessed trait *B*).

It is easy to see that, using this criterion, natural selection still comes out as having no explanatory relevance. Just consider the case of the dark-colored moth *M* with its dark-colored parents once more. If selection had favored pale genotypes in the parental generation and there had not been any causal factor that independently guaranteed *M* to exist and have dark coloration, would *M* then have existed and been pale? The answer is that it would not. There are two possible outcomes of a scenario in which *M*'s dark-colored parents are subject to a selective regime that favors paleness. First, if selection for paleness is sufficiently strong, *M*'s parents fail to reproduce, and *M* is not born. Second, if selection for paleness is weak enough for *M* to be born, the selective favoring of one trait rather than another fails to contrastively explain *M*'s existence, since *M* exists regardless of which trait is favored. Thus, natural selection can at most explain whether *O* exists or not. If *O* exists, natural selection cannot make a difference to how it exists. Inheritance, mutation, and development will explain why, when *O* exists, it has the traits it does.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

This suffices to show that if we exclude the possibility of overdetermination by applying the SCCS′, a difference in the selective regime cannot make a difference to which traits an organism possesses. Therefore, selection is causally irrelevant to explaining why a given organism has the traits it does. And if natural selection is not a relevant cause of an organism's traits, we can exclude the possibility that it features as an overdetermining cause in tests of the SCCS. Thus, although overdetermination may present a problem for the SCC, it cannot present an objection to the SCCS. The problem of overdetermining causes never arises for the SCCS.

6. Room for a Positive View? Since Birch's argument against the SCCS fails, his account does not give any immediate reason to think that the Negative View is in trouble. Nevertheless, it is worth considering the alternative counterfactual criterion he introduces. If Birch is right that this alternative account of contrastive explanation supports the Positive View of natural selection explanation, further work will be needed to determine which account of contrastive explanation is to be preferred.

The alternative Birch offers is adapted from Strevens (2008), who presents his account of contrastive explanation in the form of a recipe. Strevens proposes that in order to explain why outcome E rather than E^* occurred, we need to construct two separate explanatory models that satisfy the following conditions:

(1) The model for E is veridical; it is a (noncontrastive) standalone explanation for why E occurred.
(2) The model for E^* is veridical except for one or more states of affairs that you might call the *switching events*. The model falsely represents the switching events as not having occurred.
(3) Each of the switching events appears in the model for E. They are therefore difference-makers for E.
(4) The model for E^* is in some sense the most plausible story as to how E^* might have occurred. (Perhaps it is the model instantiated in the closest possible world in which E^* occurs.) (Strevens 2008, 175; symbols modified)

To illustrate the application of this recipe, Strevens gives the example of explaining why Rasputin was murdered by being drowned rather than by being thrown into a volcanic crater. The two causal models for this contrast agree that Rasputin was murdered but disagree on the method. They differ with regard to two switching events: the actual model contains a river whereas the counterfactual one does not (switch 1), and the actual model contains no nearby volcano whereas the counterfactual model does (switch 2). On the basis of these two models, we can say that the explanation for why Rasputin was influviated

https://doi.org/10.1086/703573 Published online by Cambridge University Press

rather than incinerated is that a river rather than a crater was available (Strevens 2008, 176).

Birch contends that by using Strevens's account it can be shown that natural selection helps explain why an individual has the traits it does. The selective regime can feature as a switching event in a Strevens-style pair of models. Birch illustrates this using an example reproduced here as figure 1. It depicts an actual and a counterfactual model of inheritance and selection in a population of asexual individuals with two loci and four possible genotypes: *AB*, *Ab*, *aB*, and *ab*. Birch argues that, according to Strevens's procedure, natural selection helps explain why outcome *E* (O_6 having genotype *AB*) rather than outcome E^* (O_6 having genotype *ab*) occurred. Here is why:

> To apply Strevens's procedure, we consider the most plausible model of how O_6 might have ended up with *ab*. This is, of course, a model in which the relevant mutations do appear in O_6's lineage. But it is *also* a model in which the selection pressures are different; for, had they been the same, the mutations would not have been passed down to O_6 (indeed, O_6 would never have existed). There are therefore multiple switching events: mutations which in fact failed to occur would have needed to occur, and selection pressures which actually obtained would have needed to be different. (Birch 2012, 572–73)

I agree that if the selective regime could function a switching event in a set of models for which trait a given organism inherits, natural selection would help explain why that organism has the traits it does. However, I believe that something is amiss with Birch's argument for why the selective regime can indeed count as a switching event. Remember the question we are aiming to answer: "Why does O_6 have *AB* rather than *ab*?" This explanandum presupposes that O_6 exists and has either *AB* or *ab*. The switching events that appear in the relevant models for answering this contrastive question must therefore be factors that make a difference to whether O_6, given that it exists,

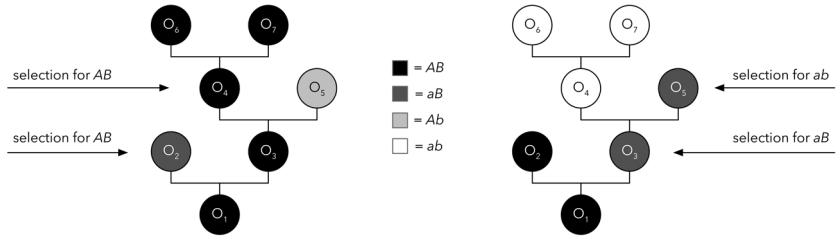


Figure 1. *Left*, model in which individual O_6 inherits genotype *AB* after selection for *AB*. *Right*, counterfactual model in which individual O_6 inherits genotype *ab* after selection for *ab*. (Drawn after Birch [2012] figs. 2 and 3.)

https://doi.org/10.1086/703573 Published online by Cambridge University Press

has *AB* or *ab*. The inherited mutations from O_6's lineal ancestors obviously combine to present a difference that makes a difference to which traits O_6 has; they are switching events. But the selective regime is not a difference maker of this kind: a change in the selective regime does not constitute a difference that makes a difference to which traits O_6 inherits; it only determines whether O_6 comes to exist. In fact, Birch admits this much when he states that selection only affects whether a given genotype can be "passed down" to O_6 by causing it to be born. In other words, Birch's error has been to think that the change in selective pressures that is required to make the counterfactual model *about* O_6 can also feature *in* the counterfactual model of O_6 as a switching event.

One might object that even if natural selection as such cannot make a difference to O_6's traits, it nevertheless contributes something to the contrastive explanation of O_6's traits by showing that a difference in selection is needed to satisfy the presupposition that O_6 exists in the counterfactual scenario. In the actual scenario, the birth of O_6 requires selection for *AB* in the two generations downstream from O_1, whereas in the counterfactual scenario it requires selection for *Ab* and selection for *ab* in these generations. One could argue that natural selection is therefore part of an "inclusive contrastive explanation" of O_6's traits.[8]

The problem with this response is that it either collapses into a defense of the Negative View or presents a confused notion of contrastive explanation. The first outcome results if an "inclusive contrastive explanation" is understood to be a conjunction of two contrastive explanations: a combination of the answer to the question "Why does O_6 exist rather than not?" that refers to natural selection and the answer to the question "Why, given that O_6 exists, does it have *AB* rather than *ab*?" that refers to inheritance and mutation in its lineal ancestors. This pair of explanations fails to challenge the Negative View. Defenders of the Negative View do not disagree with the answer to the first question, and their view is vindicated by the answer to the second question: for any given organism natural selection does not help explain why that organism has the traits it does.

Alternatively, one might take an "inclusive contrastive explanation" to be a singular explanans for a unified explanandum of the form "Why does O_6 exist rather than not, having the traits it does?" However, this explanandum is ill formed. It calls for an answer that explains the existence of O_6 while simultaneously presupposing its existence. No single explanans can satisfy these two demands. An answer that presupposes O_6's existence will fail to explain why O_6 exists, while an answer that explains O_6's existence will fail to presuppose that it exists in order to explain why it has the traits it does.

8. I thank a reviewer for raising this objection.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

We can conclude that Birch's attempt to save the Positive View fails. Contrary to what he argued, Strevens's account of contrastive explanation supports the Negative View rather than the Positive View. Moreover, we have seen that in his attempt to argue for the Positive View, Birch commits the same error as we already encountered in Nanay's account. Both of them fail to recognize the essential role of explanatory presuppositions in requests for contrastive explanation. Neither has fully come to grips with the fact that presuppositions are not the target of contrastive explanantia but need to be assumed by responses to count as relevant explanantia. An explanation of why an organism has certain traits rather than others needs to presuppose that this organism exists, rather than account for it.

7. Conclusion. I have argued that the empirical response to the Negative View misses the mark and provided additional reasons to conclude that the metaphysical response has no teeth. The explanatory response, being well targeted and metaphysically unproblematic, seemed like a promising candidate to topple the Negative View. Moreover, an important defense of the Positive View along these lines had not received any convincing rebuttal. However, I have argued that the explanatory response fails nonetheless. Not the defenders of the Negative View but their critics make an error about the nature of contrastive explanation. The critics overlook that, since natural selection already features as presupposition in the explanation of why a selected individual has the traits it does, it cannot also feature in the explanation itself.

REFERENCES

Birch, Jonathan. 2012. "The Negative View of Natural Selection." *Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (2): 569–73.

Cummins, Robert. 1975. "Functional Analysis." *Journal of Philosophy* 73 (20): 741–65.

Forber, Patrick. 2005. "On the Explanatory Roles of Natural Selection." *Biology and Philosophy* 20:329–42.

Helgeson, Casey. 2013. "What Selection Can and Cannot Explain: A Reply to Nanay's Critique of Sober." *Philosophy of Science* 80 (1): 155–59.

———. 2015. "There Is No Asymmetry of Identity Assumptions in the Debate over Selection and Individuals." *Philosophy of Science* 82 (1): 21–31.

Hitchcock, Christopher R. 2013. "Contrastive Explanation." In *Contrastivism in Philosophy*, ed. Martijn Blaauw, 11–34. London: Routledge.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Lewens, Tim. 2001. "Sex and Selection: A Reply to Matthen." *British Journal for the Philosophy of Science* 52 (3): 589–98.

Lipton, Peter. 2004. *Inference to the Best Explanation*. London: Routledge.

Matthen, Mohan. 1999. "Discussion: Evolution, Wisconsin Style; Selection and the Explanation of Individual Traits." *British Journal for the Philosophy of Science* 50 (1): 143–50.

———. 2003. "Is Sex Really Necessary? And Other Questions for Lewens." *British Journal for the Philosophy of Science* 54 (2): 297–308.

McLoone, Brian. 2013. "Selection Explanations of Token Traits." *Studies in History and Philosophy of Biological and Biomedical Sciences* 44 (3): 342–46.

https://doi.org/10.1086/703573 Published online by Cambridge University Press

Nanay, Bence. 2005. "Can Cumulative Selection Explain Adaptation?" *Philosophy of Science* 72 (5): 1099–112.
———. 2010. "Natural Selection and the Limitations of Environmental Resources." *Studies in History and Philosophy of Biological and Biomedical Sciences* 41 (4): 418–19.
Neander, Karen. 1988. "What Does Natural Selection Explain? Correction to Sober." *Philosophy of Science* 55 (3): 422–26.
———. 1995a. "Explaining Complex Adaptations: A Reply to Sober's 'Reply to Neander.'" *British Journal for the Philosophy of Science* 46 (4): 583–87.
———. 1995b. "Pruning the Tree of Life." *British Journal for the Philosophy of Science* 46 (1): 59–80.
Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic.
Pust, Joel. 2001. "Natural Selection Explanation and Origin Essentialism." *Canadian Journal of Philosophy* 31 (2): 201–20.
———. 2004. "Natural Selection and the Traits of Individual Organisms." *Biology and Philosophy* 19 (5): 765–79.
Razeto-Barry, Pablo, and Ramiro Frick. 2011. "Probabilistic Causation and the Explanatory Role of Natural Selection." *Studies in History and Philosophy of Science* 42 (3): 344–55.
Sartorelli, Joseph. 2015. "Biological Process, Essential Origin, and Identity." *Philosophical Studies* 173 (6): 1603–19.
Sober, Elliott. 1984. *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Chicago: University of Chicago Press.
———. 1986. "Explanatory Presupposition." *Australasian Journal of Philosophy* 64 (2): 143–49.
———. 1995. "Natural Selection and Distributive Explanation: A Reply to Neander." *British Journal for the Philosophy of Science* 46 (3): 384–97.
Stegmann, Ulrich. 2010a. "Reply to Bence Nanay's 'Natural Selection and the Limited Nature of Environmental Resources.'" *Studies in History and Philosophy of Biological and Biomedical Sciences* 41:420–21.
———. 2010b. "What Can Natural Selection Explain?" *Studies in History and Philosophy of Biological and Biomedical Sciences* 41:61–66.
Strevens, Michael. 2008. *Depth: An Account of Scientific Explanation*. Cambridge, MA: Harvard University Press.
Walsh, Denis M. 1998. "The Scope of Selection: Sober and Neander on What Natural Selection Explains." *Australasian Journal of Philosophy* 76 (2): 250–64.

https://doi.org/10.1086/703573 Published online by Cambridge University Press