To appear in: Nehm, R. H., Kampourakis, K (expected 2026). *Explanation in Biology Education: Theory and Practice.* Contributions from Biology Education Research. Springer.

Causal Explanation in Biology: A Control Element Account

Lauren N. Ross

Abstract: This chapter discusses the topic of scientific explanation, which has received significant attention in biology and philosophy of biology. In the philosophical literature, explanation is viewed as an important aim of science, which differs from other scientific projects such as mere description, prediction, and classification. Philosophical accounts of scientific explanation aim to clarify how explanations work, which exact factors are (and are not) explanatory, and the principles that guide explanatory practice. This chapter provides an account of causal explanation in biology with a focus on examples from molecular and cellular contexts. I will introduce a control element account of causal explanation, in which biologists explain outcomes by appealing to a set of main causes that are control elements for a target. This analysis suggests that explanatory causes are selected according to three control criteria—whether they provide information about (i) minimal control, (ii) ideal control, and (iii) more control (MIM) over the explanatory target. This account will outline the main steps of explanations, standards that need to be met, and the considerations that guide explanatory relevance, that is to say, how biologists determine what factors explain an outcome of interest.

1 Introduction. The topic of scientific explanation has received significant attention in both biology and the philosophy of biology. In these domains, explanation is often considered "one of the foremost objectives of empirical science" (Hempel 1965) and it is distinguished from other scientific projects such as providing mere descriptions, predictions, and classifications (Woodward 2003; Ross 2025b). In this manner, describing the particular shape of a pea plant's leaves is different from explaining *why* they have this shape. Similarly, predicting the future occurrence of diabetes from physiological markers (such as hemoglobin a1c) is different from explaining *why* diabetes occurs. Additionally, classifying animals into vertebrate and invertebrate groups is different from explaining *why* they have features characteristic of those groups. In all of these cases, explanations are distinguished from these other projects in offering "deep" understanding of the world often through answering *why-questions* about some phenomenon of interest. In answering these why-questions, explanations are expected to identify factors that are "responsible for" and "account for" the phenomenon of interest (Nagel 1961; Hempel 1965). When it comes to explaining traits and phenotypes in biology, explanations appeal to all sorts of factors including: gene variants, environmental details, evolutionary elements, and mathematical properties, among others. Philosophical accounts of scientific explanation aim to clarify how explanations work, which exact factors are (and are not) explanatory, and the principles that guide explanatory practice.

†To contact the author, please write to: Lauren N. Ross, Department of Logic and Philosophy of Science, University of California, Irvine, 3151 Social Science Plaza A, University of California, Irvine 92697-5100; email: rossl@uci.edu.

*Acknowledgements: I would like to thank Jim Woodward for helpful comments on this paper. This work is supported by a John Templeton Foundation Grant (63021).

The topic of biological explanation is not just interesting for mere speculation or "sheer intellectual curiosity" (Hempel 1965, 333). Identifying the principles behind biological explanation matters for practical, concrete, and real-world issues. Scientific explanations have implications for medical recommendations and public policies (such as vaccination, anti-smoking health campaigns, etc.), they capture standards that determine what research gets funded and published, they inform what is viewed as responsible and blameworthy for outcomes, they guide actions for making changes, and they matter for clear science communication in various domains, including in research settings, educational contexts, and to the public. In this manner, getting explanation right helps capture the success of science, it promotes our best scientific work and reasoning, and should support making effective changes in the world. Alternatively, getting explanation wrong can be costly, dangerous, and harmful—we see this in examples such as drugs like thalidomide, incorrect humoral explanations and treatments of disease, vaccine hesitancy, and the impact of species extinctions on ecosystems (such as the grey wolf population in Yellowstone National Park). Getting biological explanation right requires identifying the principles, strategies, and goals that scientists use in their explanatory practice—and how these show us when something counts as a genuine explanation (and when it does not count). If science provides us with our best understanding of the world—better than astrological recommendations, tea leaf readings, and mere guesses—we should be able to say why this is the case and how scientific explanation works.

For these and other reasons, philosophers of science are interested in capturing the "logic," structure, and guidelines of scientific explanation (Nagel 1961; Hempel and Oppenheim 1948; Woodward 2003; Ross 2025b). In particular, there is interest in capturing the (i) standards that genuine biological explanations need to meet, (ii) different types of explanation in biology, (iii) how scientists provide explanations despite various challenges (such as complexity), and (iv) what differentiates explanations that are good, bad, and better. This chapter introduces a philosophical framework for answering these questions, particularly when it comes to causal explanation in biology.

2 Scientific explanation: A framework and two hard problems. A helpful framework for considering types of scientific explanation is one in which explanations have three main elements. These three elements are: (i) an explanatory target, referred to as the *explanandum*, (ii) what does the "explanatory work," referred to as the *explanans*, and (iii) a dependency relation that connects the explanandum (i) and explanans (ii). An illustration that captures these three elements is seen in figure 1.

The first element of an explanation is the (i) explanatory target, which is some natural phenomenon of interest that a scientist is interested in explaining. In the context of the biological sciences, this could be a particular trait or phenotype (such as fruit fly color, plant height, and dorsal fin shape), a disease outcome (Parkinson's disease, measles, etc.), or some complex behavior of a biological system (oscillations in hormones, pulsing behavior of neural firing, etc.). The explanatory target is often something that can be expressed in terms of a why-question.[1] For example, a scientist might ask "Why does this group of fruit flies have brick-red eye color?" or "Why does this patient have Parkinson's disease?" or "Why do these hormones (in the hypothalamic-pituitary-gonadal (HPG) axis) oscillate as opposed to remaining steadily expressed?". Each of these questions specifies an explanatory target or explanandum, which is something that scientists find interesting, that they view as surprising, or that "calls out for explanation" (Ross 2025b).

[1]Not all possible why-questions count here, but "explanation-seeking why-questions," as Hempel clarifies (Hempel 1965, 334).

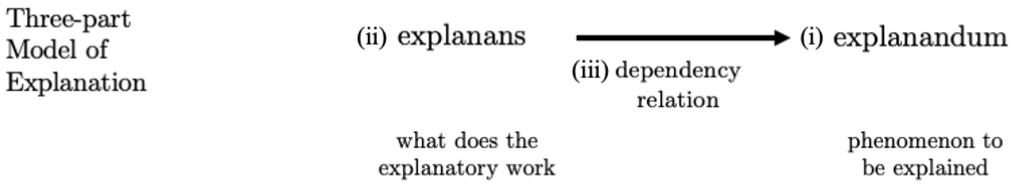

Figure 1: A three-part model for scientific explanation: (i) the *explanandum* is the explanatory target or phenomenon to be explained, (ii) the *explanans* refers to what does the "explanatory work," and (iii) is a dependency relation that connects the explanandum (i) and explanans (ii).

Usually, the first step in providing an explanation involves specifying a well-defined explanatory target of interest.

Once an explanatory target is specified, the next step involves identifying the (ii) explanans, which is what does "the explanatory work" and "is responsible for" the target. In the context of causal explanation, this involves identifying the main causes that "make a difference" to the explanatory target (Woodward 2003; Ross 2025b). So, for example, measles is partly explained by a virus (and ensuing immunologic changes), fruit fly eye color is partly explained by gene variants (and downstream molecular outcomes), and plant height is partly explained by environmental causes (such as fertilizer, sunlight, etc.). After getting partial explanations, scientists often work to gain explanations that are more complete. The idea behind causal explanation is that "causes explain their effects," that causes are "difference-makers" for effects, and that effects "depend" on their causes (Woodward 2003). This second step of scientific explanation requires already having a notion of what type of (iii) dependency relation should be present between the explanandum and explanans to justify that an explanation has been provided. For causal explanations, this dependency relation is causal, but other types of explanations involve other types of dependency relations (such as mathematical dependencies) (Ross 2025b). The dependency relation is viewed as important in justifying that a genuine explanation has been provided.

With this basic model we can introduce two hard problems of scientific explanation—explanatory dependence and explanatory relevance. Explanatory dependence concerns the nature of the connection between the explanandum (explanatory target) and explanans (factors that do the explanatory work). This captures what "powers" or justifies an explanation and also what distinguishes explanation from non-explanatory projects. Providing an explanation often involves identifying what the explanatory target "depends on" and, in this manner, the nature of this dependency is essential to specifying when something counts as explanatory. Some philosophical debates examine whether there are different types of explanation in science that involve different types of dependency relations. For example, while the above causal explanations include a causal dependency, it has been suggested that other types of explanations involve dependencies that are mathematical, deductive, evolutionary or functional, and so on (Baker 2005; Batterman 2001; Lange 2018; Woodward 2019; Ross 2025b). This chapter will focus on causal explanation, while discussion of other types can be found in the literature (Baker 2005; Woodward 2019; Ross 2025b). Questions about explanatory dependence help capture pluralism and monism claims about scientific explanation. One explana-

tory pluralism view claims that there are different "types" of explanation in science, by arguing that there are different types of "explanatory dependence" relations that power explanations (see (Ross 2025b) for more on this.). Alternatively, an explanatory monism view might suggest that there is only one type of explanatory dependence relation and, thus, only one type of legitimate explanation in science. An example of explanatory monism is seen in claims that causal explanation is the only genuine type of explanation in science (Skow 2014).

A second hard problem is explanatory relevance, which involves determining which factors are explanatory for an outcome among all factors that are present or that might be considered. Once a target is specified, we need to know how to determine which factors explain an outcome and which do not. Biologists cannot cite all factors or all details present in a system (or even that seem mildly "relevant"). Also they don't seem compelled to do this. In many domains, scientists are highly selective when identifying factors they view as explanatory for an outcome. When biologists explain fruit fly eye color they focus on genes and ensuing biomolecular factors, but not the presence of oxygen, not the Big bang, and not entities in fundamental physics that might underlie this phenomenon. Additionally, biologists also don't view just any biological factors as explanatory—they identify specific causes as explanatory, which they distinguish from others. Significant debates in philosophy focus on the considerations that guide the selection of explanatorily relevant factors. Some have claimed that this selection process is subjective and "capricious" (Mill 1874; Lewis 1986), some view it as predominantly audience-relative (Potochnik 2017), and others claim it is principled and guided by objective considerations (Woodward 2003; Ross 2023d). The analysis in this chapter follows the third option—it develops a framework that captures principled reasons for biological explanation and explanatory relevance that are not highly subjective or audience relative, but instead guided by more objective, stable, and scientific-goal-focused considerations.[2]

This chapter provides an account of causal explanation in biology with a focus on examples from molecular and cellular contexts. I will introduce a "control element" account of causal explanation, in which biologists explain outcomes by appealing to a set of main causes that are "control elements" for a target. This account will outline the main steps of explanations and the considerations that guide explanatory relevance, that is to say, how biologists determine what factors explain an outcome of interest.

3 Setting up Causal Explanation: Explanatory Target and Causation. A helpful starting point for understanding causal explanation in biology is with the guiding principle that "causes explain their effects". On this view, an explanation should cite the main causes that are responsible for, account for, and produce the outcome. Fulfilling this requires both a (1) well-defined explanatory target and (2) an account of causation. Identifying a (1) well-defined explanatory target is often viewed as a first main step in providing an explanation. This is because, before we can determine what is explanatory, we need to know exactly what we want to explain. Reviewing various subfields of biology illustrates numerous types of explanatory targets, such as: why a species has gone extinct in some ecosystem, why hippocampal neurons have a (pulsatile) particular firing behavior, why birds in a given region have acquired different beak shapes, why a contaminant in a lake harms some fish and not others, why genetically identical plants reach different heights, and why

[2]Further questions in the subcategory of explanatory relevance include: how causation should be defined, how explanatory causes relate to causal concepts such as "mechanism" ((Bechtel and Richardson 2010; Craver 2007) and what makes explanations "deeper" or "better" and more "powerful" (Hitchcock and Woodward 2003; Ylikoski and Kuorikoski 2010).)

a group of fruit flies have a particular eye color. As mentioned above (sec 2), explanatory targets can often be expressed in terms of "why-questions," where the targets are something that scientists find "surprising," that they are interested in, and that "calls out for explanation" (Hempel 1965; Nagel 1961; Ross 2023d).

3.1 The Explanatory Target (Explanandum): Two Standards. A common first step in providing an explanation involves specifying a well-defined explanatory target and this is often more difficult than is appreciated. This is because well-defined targets need to be precise in ways that are easily (or often) overlooked. In order to see this, consider two candidate explanatory targets. In a first case "we ask for an explanation of the aurora borealis" (Hempel 1965, 334) and in a second we want to know "why do human beings have lungs?" (Nagel 1961, 19-20). While these might seem unobjectionable, they are not yet sufficiently clear in ways required of well-defined explanatory targets. An issue with the first example is that it doesn't specify what exactly it is about the aurora borealis (or which of its many features) we want to explain. This matters, because different features of the aurora borealis are likely to require different explanations and it doesn't provide enough to move explanatory inquiry forward. For the second, the why-question is ambiguous as it can be interpreted in different ways. The second case might be asking why human beings have evolved to have lungs or it could be asking why humans have lungs to serve various physiological goals (such as gas exchange, etc.). Again, this matters because these different explanatory targets will require different explanations and it isn't sufficiently clear what exactly we want an explanation of. As Hempel and Nagel state, these examples do not yet have "clear meaning" and are "ambiguous" as they "may be construed" as different explanatory targets altogether (Nagel 1961, 19)(Hempel 1965, 334). More specifically, "requests for an explanation of the aurora borealis, of the tides, solar eclipses...or of a given influenza epidemic, and the like have a clear meaning only if it is understood what aspects of the phenomenon in question are to be explained"(Hempel 1965, 334).

In order to prevent the above issues, we can specify two standards that well-defined explanatory targets need to meet. These two standards are (i) property characterization and (ii) contrastive focus characterization. First, well-defined explanatory targets should clearly specify the property of interest that is to be explained. This often involves isolating a particular property or feature from a complex, multi-featured phenomenon. Similar to the aurora borealis above, phenomena such as neural firing, protein folding, gene expression, and eye color are all multi-featured. For example, instead of the explanatory target "neural firing" a clearer target is the "pulsatile firing of hippocampal neurons," instead of the target "protein folding" a better alternative is "why some initial amino acid sequence X folds into Y," and instead of "fruit fly eye color" a better explanatory target is "the presence of brick red eye color in fly group B." Sometimes these specifications are intended without being clearly communicated. If not clearly communicated they can lead to misinterpretation and cross-talk. The importance of this requirement is also seen in the choice and characterization of variables in experimental work in science—overly vague descriptions of dependent or independent variables, such as "the action potential" and "protein folding," fail to meet standards for "unambiguous variable definitions" (Miksad and Abernethy 2018).

A second standard for well-defined explanatory targets is (ii) contrastive focus characterization, which involves specifying a contrast for the property in question. In other words, even when a clear property is specified—such as brick red eye color in fruit flies—it needs to be clear what the contrast of interest is for this target which also captures the "as opposed to what" that one wants to explain (van Frassen 1980; Achinstein 1984). Do we want to explain why this group of fruit flies has

"brick red eye color in contrast to white color?," or "brick red in contrast to scarlet?," or "brick red in contrast to all other possible colors?". This is important because each of these contrasts specifies distinct explanatory targets, which can require distinct explanations—in this case distinct genetic and biochemical "difference makers" for each contrast. As another example, consider the explanatory why question: "Why does this patient have pulsatile ringing (tinnitus) in their right ear?". While this question might seem clear, it has different potential contrasts that capture distinct explanatory why-questions. This question might be asking why this patient has (i) pulsatile ringing in their *right* ear as opposed to the *left* ear, or why the patient has (ii) *pulsatile* ringing in their right ear in contrast to *tonal* ringing (or musical ringing), or why the patient has (iii) *ringing* in their right ear as opposed to *no ringing* at all. From this question alone the intended contrast is not sufficiently clear. As these distinct contrasts require different explanations, this reveals again the importance of a well-defined explanatory target.

These two standards encourage precision in formulating explanatory targets. This is important, because clear explanatory targets situate explanatory inquiry and serve as a guide through the process. In some cases, explanatory inquiry can remain stuck at this step due to challenges in characterizing the property of interest, a lack of consensus regarding this characterization, or other issues. One arguable example here are efforts to explain "consciousness"—as their are challenges and debates about how to best characterize this phenomenon, much explanatory progress remains at this first step (which must be partly addressed before explanatory factors can be identified). In other cases, scientists might start with one explanatory target, but change and refine this along the way as progress is made. An example of this includes the shift from the single category "diabetes mellitus" (high blood sugar) to the categories of "type 1" and "type 2" diabetes (which have distinct symptomatic profiles and etiologies). Ideally, we want our framework for explanation to accommodate how explanatory targets are refined over time, but also to capture the need for clear, fixed, precise explanatory targets when an explanation is provided. While these are two important standards for well-defined explanatory targets, they are not exhaustive and the topic of standards for explanatory targets certainly deserves more attention.

3.2 An Account of Causation: Interventionism. Once a well-defined explanatory target is specified, the next steps in providing a causal explanation involve identifying the main causes that explain this target. First we will introduce a basic account of causation for this, then in the next section we can build on this to capture a procedure for selecting the main, important causes that explain the target. This analysis will rely on an interventionist account of causation (Woodward 2003). The interventionist account is motivated by scientific methods and reasoning used to assess causation, such as unconfounded experimental manipulations and frameworks in cognitive science, epidemiology, computer science, and other domains. On this account, causes and effects are noted as variables (X, Y, Z, etc.) and these variables represent properties in the world, such as gene variants, dietary factors, phenotypes or traits, and so on. Additionally, these variables can take on distinct values (0, 1, 2, etc.) that represent different states of these properties, such as whether a gene variant (or other property) is present (1) or absent (0), or states of the property that come in degrees (0, 1, 2...). With this variable-value framework in place, causation is understood in the following way:

> Interventionism (I): X causes Y if and only if there are background circumstances B such that if some (single) intervention that changes the value of X (and no other

variable) were to occur in B, then Y or the probability distribution of Y would change (Woodward 2010, 2003)

On this account, in order for X to be a cause of Y it needs to be the case that if an intervention changed X (and only X), this would lead to changes in Y, in some set of background conditions. In this manner, causation is understood in terms of counterfactuals about interventions. Counterfactuals capture "if...would" statements, such as, (i) if vitamin C were removed from her diet, then she would acquire scurvy, (ii) if I had taken aspirin yesterday, I would not have gotten a headache, and (iii) if they had the *huntingtin* gene variant, they would develop Huntington's disease. As seen in each of these statements, the counterfactuals about interventions do not require that the interventions are possible with current technology or permissible ethically. Instead, these causal claims are understood in terms of hypothetical interventions about what would happen to an effect if a candidate cause were hypothetically changed. We often have evidence about what would happened under such hypothetical interventions even if we can't directly perform them. This allows us to capture how we make causal claims (in science and everyday life) about past events and causes we can't manipulate. For the candidate cause, we should be able to associate it with a hypothetical experiment, in which manipulations of the cause lead to changes in the effect (Woodward 2003). In this manner, "an intervention can be thought of as an unconfounded manipulation of the sort that might be realized in an "ideal" experiment" (Woodward 2011, 412).

In sum, the interventionist account captures causation in terms of control. In this sense, causes are like handles in the world that (if manipulated) can provide control over their effects. This account allows us to capture how causes are "difference makers" that are responsible for, that account for, and that explain their effects, and also how effects "depend" on their causes. This interventionist standard (I) provides a minimal criterion for a factor to be a cause or to be causally relevant to an outcome. Note that this criterion requires just that under *some* interventions, in *some* background conditions, changes to X will produce *some* changes to Y. The changes in Y should be those specified in the explanatory target (and its contrastive focus), but this only requires some control over this target. If a candidate cause X fails to meet this criterion (I) it is not causal for the effect or causally relevant to this effect. This causation account helps capture why "correlation does not imply causation"—while a rooster crowing and the sun rising are correlated, they are not causal because hypothetical interventions on crowing roosters do not change or control the sun rising. Similarly, other biomarker-effect relationships are correlated and useful for predictions (Ha1c and diabetes, bulls eye rash and Lyme disease, koplik spots and measles, etc.) without capturing causality, in part, because they don't capture or provide information about control.

This framework uses interventionism as our criterion for causation. While a main role of interventionism is distinguishing causal from non-causal, there are features beyond this criterion that we prioritize in selecting main explanatory causes for an outcome. One way to clarify this is that, while interventionism captures that a cause needs to provide *some* control over an effect, there are *different types of control* that causes can provide, and these types matter for selecting explanatorily relevant causes. I am going to suggest three "control" standards that guide which causes explain an outcome, as discussed more in the next section.

4 Providing a Causal Explanation: A Control Element Account. Recall that one of the hard problems for scientific explanation is explanatory relevance—this involves specifying which of many causes are the ones that explain a target of interest. We need to know how this works

in order to provide an account of causal explanation in biology. Ideally, we want to know what distinguishes causal details that *are* explanatory from those that are *not* explanatory, and which details *improve* an explanation or make explanations better. We can build on interventionism to offer a solution to this hard problem. I am going to suggest a "control element" account of causal explanation, in which biologists explain outcomes by appealing to a set of main causes that are "control elements" for a target. These explanatory causes are often represented in sparse, reduced order models seen throughout biology, often in diagrams, figures and illustrations (and also referenced in verbal, textual accounts). These illustrations often depict a set of control elements, in which upstream and intermediate causes provide control over a downstream effect. I am going to suggest that explanatory causes are selected according to three control criteria—whether they provide information about (i) minimal control, (ii) ideal control, and (iii) more control (MIM) over the explanatory target.

Control standards for explanatory causes:

1. **minimal control**: interventionism
2. **ideal control**: strong, stable, fast
3. **more control**: further control elements

Consider causal explanations for lac operon gene expression, hippocampal neural firing, and the disease measles (where we are interested in the contrast of 'present' and 'absent' for each). Notice that when biologists explain these outcomes they (i) choose some "starting point" or initial, upstream, or root cause(s) and that (ii) their explanations are very selective about which causes they view as explanatory. Explaining lac operon gene expression starts with changes to lactose concentration, explaining hippocampal neural firing starts with excitatory signals to the neuron, and explaining measles starts with exposure to a virus. Furthermore, explanations of each of these outcomes are selective in terms of the particular factors cited. A large number of factors are omitted from these explanations, such as lower-level details from physics and chemistry, details about a person's parents and ancestors, the Big bang, and so on. This process of choosing and omitting explanatory causes is not arbitrary—it is guided by principled considerations, discussed more (in what follows) below.

4.1 Minimal Control: Interventionism. A first criterion that explanatory causes need to meet is that they need to provide (i) minimal control over the specified target.[3] This first "minimal control" criterion requires that, in background circumstances of interest, there are *some* manipulations of the cause that provide *some* control over the explanatory target. This criterion captures basic requirements for causation and the interventionist criteria. Notice how powerful this first criterion is in terms of eliminating and including factors from consideration in counting as explanatory. If our explanatory target is measles in humans, the measles virus meets this standard because their are manipulations of (exposure to) this virus that provide minimal control over measles occurrence (and nonoccurence). Aside from this virus, other immunologic and cellular factors also meet

[3]For these criteria, when "provide control" is mentioned it nearly always means "information about control". As mentioned, many genuine causes cannot be directly manipulated to gain control, but they still offer information about control under hypothetical interventions. An example is the location of the moon causing changes in tides on each—the moon's location provides information about control, even if we can't manipulate the moon's location to harness this control.

this minimal control standard, as changes to these factors also offer some control over the disease outcome (such as immune status when the virus is present). However, notice that many factors do not meet this standard and, thus, are not explanatory. Examples here include the Big bang, factors in fundamental physics, and other considerations such as the individual's diet, childhood, exercise regimen, and so on. We can consider hypothetical changes to the Big bang (whether it occurred or not), but this offers no control over the presence and absence of measles in human patients—it might control whether humans exist or not, but, of course, this isn't the target of interest. Similarly, there are no known factors in fundamental physics that if manipulated, would control the specific outcome of measles presence and absence in humans. This does not deny that lower-level physics constitutes (or makes up) biological phenomena, it denies that lower-level physics *explains* these biological phenomena. This is consistent with biomedical explanations of measles, which cite viruses and not the Big bang or fundamental physics. Similarly, if hypothetical changes to dietary factors, exercise routine, and childhood environment fail to provide minimal control over measles, these will not count as explanatory causes of this disease. This minimal control criterion is a threshold required to get factors "in the door" to count as causes, causally relevant, and explanatory for the target of interest.

This first criterion helps with puzzles regarding "necessary factors," such as situations when they are mistakenly viewed as explanatory. For example, when considering human diseases such as measles, it seems like various factors are "necessary" or "required" for this disease outcome to occur—these include the presence of oxygen, the presence of food, other requirements for life, the occurrence of the Big bang in past history, and so on (Lill Anjum and Mumford 2018). While seemingly necessary, these factors do not show up in our scientific explanations of these outcomes—some (mistakenly) suggest that technically they should show up in these explanations or at least there is confusion about how to handle such cases. We can show why such factors, while necessary (or required) for these outcomes to occur, are not explanatory for these outcomes—this is because changes to these factors do not control the precise explanatory target of interest. Hypothetical changes to the presence of oxygen (or the Big Bang) *do not control the precise target of presence and absence of measles in living human patients*, although they can control whether humans exist or not (which is not the explanatory target of interest). This helps correct a mistaken view that if something is necessary for an outcome it automatically counts as explanatory for the outcome. This can identify standards that are too low for explanation and causation, as they let in factors that are not actually explanatory. One advantage of precise and well-defined explanatory targets in biology, is that this reduces the number of candidate factors likely to make the cut and that have this type of control.[4]

4.2 Ideal Control: Speed, Strength, and Stability. While minimal control is an important criterion for selecting explanatory causes, we expect explanatory causes to provide more than *just* minimal control. We expect explanatory causes to provide *ideal* types of control over the target. These ideal types of control are captured by various "secondary features" of causation that can be understood within an interventionist framework. These secondary features of causation are features that causes can have or lack, or have to different degrees (Ross 2025a). While "primary features" capture the "objective core" of causation (the focus of definitions of causation), these secondary

[4]It is worth examining whether the specification of precise explanatory targets may be used as a strategy or heuristic to manage complexity. Specifying narrow explanatory targets helps limit the number of candidates to consider in identifying explanatorily relevant factors.

features capture "distinctions within causation" (Ross 2025a; Woodward 2010, 2003). For example, in biological contexts, it is often expected that explanatory causes meet the ideal control standards of speed, strength, and stability. In this manner, it is expected that these causes provide control that is sufficiently fast, strong, and stable over the explanatory target. In general, causes that meet these ideal control standards will be (and should be) viewed as "more explanatory" compared to other causes. In fact, most upstream, root causes in biological explanations often have these three types of control over the explanatory target (as discussed in more detail below).

Starting with speed, it is a common expectation in biology that explanatory causes should provide control that operates within a relevant timeframe. The relevant timeframe is set by the context of inquiry and explanatory target. For example, explanations of metabolic outcomes in human physiology include enzymes as these offer control within the relevant timescale of seconds, minutes, and hours. If enzymes were omitted, some substrates would still convert into products, however, this would take place far outside the physiological timescale of interest (years, decades, etc.)(Ross 2018). This timescale of control expectation is seen in scientists' claims that if a cause is too slow, it simply fails to count as a cause for "practical" purposes. For instance, in chemical conversions that lack catalysts (or enzymes) scientists claim that "the rate of conversion...is so slow, that for all practical purposes, it does not occur" (Woodbury 2012). Speed and timescale expectations differ greatly depending on the scientific domain, from the timescales of nanoseconds in biochemistry, to minutes and hours in human physiology, and years, centuries and longer for evolutionary processes. Whether a cause has influence that is faster or slower is always relative to some default timescale of interest.[5] ideal control Each of these secondary features—speed, strength, and stability—are ideal because they offer control that is useful, functional, and practical. Explanations are expected to cite causes with these features because they serve pragmatic purposes beyond "mere control".

Our next ideal control standard is strength, which refers to how probability boosting a cause is for its effect (Ross 2018). In this sense, a cause is stronger if the presence of the cause (after intervention) leads to a higher probability of the effect occurring and the cause is weaker if it leads to a lower probability of the effect occurring. In biological explanation, causes that rank higher in terms of strength are viewed as more explanatory and offering more explanatory power. Consider explanations of scurvy in humans that cite dietary deficiency of vitamin C.[6] When a human's diet is deficient in vitamin C this leads to a very high probability that they will acquire scurvy. In this manner, this disease has a single main upstream cause with strong control over the disease outcome. Sometimes we find causes that are only low in strength for an explanatory target, such as "unhealthy diet" in explanations of cardiovascular disease (CVD). While eating an unhealthy diet does have some control over CVD, this control is weaker as it doesn't always produce this disease (and a healthy diet doesn't always prevent this disease). In these situations, we remedy this by including additional "interacting" causes in the explanation that all together

[5] While considerations of ideal control and timescale deserve more attention, this analysis provides some initial guidelines. This work suggests that in explanatory contexts there is (i) some general timeframe within which a cause should exhibit control and that (ii) within this timeframe, causes that provide faster control are often privileged in explanations. This depends on other features of a causal relationship (such as how strong and stable it is), as there many be tradeoffs among different types of control.

[6] In this case, it is actually the absence of the cause (vitamin C) that results in disease, which is compatible with this causal framework (viz., absences can be causal). What matters here is that there is a causal variable X representing some property, where changes in the variable (present, absent) result in control over the effect (absent, present), and provide strong control, in particular.

increase the strength they offer over the target (Ross 2023d). So for example, we include diet, genetics, and exercise in causal explanations of CVD, as these causes all work together to provide a higher probability of producing (and preventing) this disease outcome.[7] For some biological targets there are single upstream causes that provide strong control (measles, scurvy, and Huntington's disease), while other targets require multiple causes to meet this high strength standard (such as in cardiovascular disease, diabetes, phenylketonuria). In both cases, it is appreciated that explanatory causes should offer strong control over the target of interest. Again, it makes sense that biologists expect explanatory causes to have this type of control, as it is very useful for the purposes of changing, controlling, and predicting outcomes.

A final ideal control standard is stability, which refers to causes that provide control over their effects across many situations (of interest) in which background factors are changed or different (Woodward 2010). Stability is important to keep distinct from strength, as they can be mistakenly conflated, partly because many paradigmatic causal relationships have both features. Consider the vitamin C–scurvy example again, with respect to the ideal control standards of speed and strength. If a human consumes a vitamin-C free diet they would get scurvy within a relevant time frame (speed) and this diet has an extremely high probability of resulting in the disease (strength). However, in addition to these features, deficiency of vitamin C also causes scurvy in a manner that is "stable" as this causal relationship holds (and doesn't break down) across many contexts where various background factors differ, such as a person's exercise regimen, childhood environment, the climate they live in, and genetic variations. This stable control is incredibly useful because it means we can focus on only this cause to predict and potentially control, change, and treat scurvy occurrence, no matter what state these other factors are in. Indeed, if we just target dietary vitamin C, we can cure all cases of scurvy worldwide across highly varied contexts. Alternatively, if dietary deficiency of vitamin C only produced scurvy in warm climates, or in individuals with a particular gene variant, or with a particular exercise program, then it would be far less stable (than we know it to be). When a cause depends on or is contingent on factors or background conditions we view it as less stable, and importantly, stability assessments always involve a default relative notion of "relevant" background conditions. We care about whether the vitamin C-scurvy causal relationship holds on planet Earth, in temperatures we experience, and so on, but not in wildly divergent background circumstances from these. The choice of relevant background circumstances is important here, and part of assessing how stable (along a continuum) a causal relationship is. Again, similar to our other ideal control standards, causes with stable control over effects are often privileged in biological explanations

For many explanations in biology, a main goal is to identify causes that meet these ideal control standards. While the minimal control standard helps capture whether a factor is explanatorily relevant or not, ideal control standards capture when causes are more explanatory, when they provide better explanations, and when they have more explanatory power. These ideal control standards specify secondary features of causation that come in degrees, such that causes have them

[7]Our earlier examples of measles, hippocampal neural firing, and lac operon gene expression, also fit this picture as their selected upstream causes meet this standard. Exposure to the measles virus leads to a high probability of producing measles, presence of excitatory neurotransmitter leads to a high probability of hippocampal neural firing, and presence of lactose (when glucose is absent) leads to a high probability of gene expression. These causal claims are contingent on particular background conditions (an unvaccinated patient susceptible to measles, hippocampal neurons that are not in a refractory state, etc.), which is the case for nearly all causal claims.

to "more-or-less" and in different amounts (Ross 2025a).[8] While causes are more explanatory when they exhibit a higher degree of these features, they rarely "max-out" on them (although we prioritize causes that get closer to this). The types of ideal control outlined here (speed, strength, and stability) are likely shared across other contexts as well, although this list is not exhaustive and different contexts may prioritize different ideals. While these three ideal control standards are important there is one more important criterion that helps capture explanatory power and what it takes to provide more complete explanations.

4.3 More Control: Further Control Elements. Notice that causal explanations can meet our ideal control standards even if they contain very few upstream causes without much more detail. So far, these look like sparse "input-output" model causal explanations. This is consistent with common claims that measles is *explained* by the measles virus, lac operon gene expression is *explained* by lactose presence (when glucose is absent), and that hippocampal neural firing is *explained* by presence of excitatory neurotransmitters. However, we also sometimes admit that these explanations are "partial" in contrast to being more "complete". Providing a more complete causal explanation often involves specifying further "how" details about "intermediates" that span the upstream cause(s) to downstream effect (also sometimes called "mechanistic" information). A final standard of "more control" helps capture the importance of and rationale for including more of these details about intermediate causes. Even when we have causes that offer minimal and ideal control, including these additional "intermediate" causes improves explanatory power by providing *more control* over the explanatory target of interest.

Once the explanatory target is specified and upstream causes are fixed, this helps anchor and bookend the explanation, respectively, so that relevant intermediate details can be filled in. Even if we know that the measles virus is causally and explanatorily relevant to measles in humans, an explanation is still improved by identifying additional intermediate "control elements" along the way. These intermediate causes are "control elements" in the sense that (i) if fixed or changed, they can prevent the upstream cause from producing the final effect and, in some cases, (ii) direct interventions on these intermediates can also control the final outcome of interest. In our lac operon gene expression example, our upstream cause is lactose (X) presence, which operates through the causal intermediates of a repressor (R) and then RNA polymerase (P), to produce gene expression (E). The presence of lactose binds to a repressor, removing the repressor from the promoter of the lac operon gene, which allows RNA polymerase to bind to the promoter and express the gene. If we intervene on R and P in this system, they can both stop influence of the earlier cause (X) and if directly manipulated they can also exert control over the effect. These additional intermediate causes offer further places of control that we could target to stop causal influence, to initiate the the causal process, and that also help "explain" how X produces E, in terms of capturing the steps through which X produces E. In some cases, we can use these causal intermediates to control and change an outcome when the upstream cause was already triggered or when it is easier to intervene on these intermediates (than the upstream cause itself).[9]

A main challenge that this returns us to is explanatory relevance, as we need to know how

[8]This is in contrast to a cause having these features in a binary fashion—either being stable or unstable.

[9](In some cases, even if upstream causes with ideal control are identified, it is easier to target downstream intermediates to control, prevent, or treat a disease outcome. For example, measles vaccines don't limit exposure to the measles virus, but instead target intermediate immune cells to prevent the virus from invading the body.)

much and what types of this intermediate cause detail should be included in an explanation. We can answer this question by using the "grain" of our explanatory target and upstream causes, with our notion of minimal control (or interventionism). Both the explanatory target and the upstream causes fix the "grain" of the explanation—the causal intermediates that are explanatory are those that offer control at the same "grain" or "level" as both of these.[10] If we focus on the upstream causes X (lactose presence) we can "chain forward" to specify, when this cause X is manipulated, what subsequent intermediates would be altered (in this case, repressor levels) and continue through intermediates to the final effect. Additionally, we can start with the explanatory target E and "chain backward" to identify what this target depends on (expression of lac operon genes depends on mRNA polymerase, tRNA and so on). These guidelines reveal what details are explanatory based on whether they provide information about control over specified variables. If our explanatory target is changes in "expression of lac operon genes" Y, this depends on (immediately upstream) RNA polymerase levels, however, it does not depend on the color or density of this polymerase enzyme, as changes in these features do not provide (minimal) control over expression levels of these genes. It is common to find these extra details provided in descriptions and stories about biological processes, such as lac operon gene expression, but they are not all explanatorily relevant or features that explain, account for, or are responsible for the outcome.[11] We often have a tendency to over-include details, often reductive or lower-level details, despite the fact that we don't think all are needed or necessary for scientific explanations. They can be useful to have for various reasons, but when it comes to explanation and control it is important to specify limits on the detail and information needed to explain. The principles outlined in the last two sections help capture such limits (on how far back in causal history and how far down in scientific levels) on detail we need to consider in providing explanations in biology, and also what makes such details more explanatory for an outcome of interest.

5 Conclusion. This analysis has introduced a control element account of causal explanation in biology. The focus on control and explanatory standards distinguishes this analysis from other accounts that view explanation as more audience-relative, subjective, and even "capricious" (Potochnik 2017; Mill 1874; Lewis 1986). This analysis specifies standards that both the explanatory target (explanandum) and explanatory factors (explanans) need to meet. As discussed, well-defined explanatory targets need to meet the standards of property characterization and contrastive focus characterization. Once these are met, a genuine causal explanation for the target should include causes that meet three control standards: minimal control, ideal control, and more control. These control standards help capture when explanations are good, bad, and better. They reveal when candidates do not count as explanatory, but also when explanations are partial in contrast to more complete.

Having standards for scientific explanation is important for providing guidance when there is debate about what counts as explanatory, when audiences disagree, and when we need to convey the justification of scientific work. Accounts of scientific explanation that are overly subjective and audience relative can struggle to provide these forms of guidance. For example, some claim that

[10] The explanatory target is the main element that fixes the "grain" of an explanation—this it because it is specified first and it fixes the target that the upstream causes need to "fit with" and match the "grain" of. Once both of specified, the can each be used to fill in the intermediates, keeping them at a similar grain.

[11] For more on the grain and level of cause-effect relationships, and notions of proportionality, see (Woodward 2010).

"explanation is at root an act of communication" and that "what counts as a good explanation depends on both the explainer and the audience" (Potochnik 2017, 127). If this is correct, it's unclear how to settle debates when audiences disagree and it is unclear how to capture forms of consensus across diverse scientific audiences. Admittedly, it is important to appreciate the significance and challenge of communicating explanations to various audiences. However, instead of viewing the standards of scientific explanation as audience-relative, it is more helpful to view "science communication" or "communicating" these explanations as audience-relative. Whether measles is caused by a virus is not determined by the audience we communicate to—it is determined by principled scientific standards, such as the control criteria outlined above. Once these standards are met, we can communicate this explanation well or poorly—and this matters immensely—but the standards for explanation are distinct from the standards for clear communication. Effective science communication has become increasingly important, both in the form of communication to various general public audiences, but also among experts in various scientific domains. What is also clear, of course, is that specifying standards for scientific explanation is often more challenging than expected (Ross and Bassett 2024).

While this chapter focuses on examples from molecular and cellular biology, it is also valuable to consider explanation in other biological domains (ecology, evolutionary, etc.) and from other systems-level perspectives (systems, computational biology). This framework relies on a "causation as control" heuristic that is likely to be fruitful in understanding explanatory practice in other contexts. Other important considerations, include the varied causal concepts, terms, and language that biologists use in their explanations. Examples of these causal concepts include references to mechanisms (Craver 2007; Ross and Bassett 2024), pathways (Ross 2021), cascades (Ross 2023a), and circuits. In other cases, causal language distinguishes types of causes, such as structural causes (Ross 2023c), causes that are deterministic versus probabilistic, causes that are reversible or irreversible (Ross and Woodward 2023), and causal constraints (Raja and Anderson 2021; Silberstein 2021; Ross 2023b), among many other examples (Ross 2023b). It is important to provide rigorous accounts of how these causal systems and types of causes differ and how these differences matter for methods used to study these systems, scientific reasoning, and for the explanations scientists provide. The control element account of causal explanation introduced in this chapter provides one framework that can be used to make progress in these areas and to capture the standards that biological explanations need to meet.

References

Achinstein, P. (1984). The pragmatic character of explanation. pages 1–19. PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.

Baker, A. (2005). Are there Genuine Mathematical Explanations of Physical Phenomena? *Mind*, 114(454):223–238.

Batterman, R. W. (2001). *The Devil in The Details*. Oxford University Press, Oxford.

Bechtel, W. and Richardson, R. C. (2010). *Discovering complexity*. The MIT Press, Cambridge, MA.

Craver, C. F. (2007). *Explaining the brain: mechanisms and the mosaic unity of neuroscience*. Clarendon press, Oxford.

Hempel, C. (1965). *Aspects of scientific explanation*. The Free Press, New York, NY.

Hempel, C. G. and Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, 15(2):135–175.

Hitchcock, C. and Woodward, J. (2003). Explanatory Generalizations, Part II: Plumbing Explanatory Depth. *Noûs*, 37(2):181–199.

Lange, M. (2018). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford University Press.

Lewis, D. A. (1986). Causal explanation. In *Philosophical papers*, volume II. Oxford University Press.

Lill Anjum, R. and Mumford, S. (2018). *Causation in Science and the Methods of Scientific Discovery*. Oxford University Press.

Miksad, R. A. and Abernethy, A. P. (2018). Harnessing the Power of Real-World Evidence (RWE): A Checklist to Ensure Regulatory-Grade Data Quality. *Clinical Pharmacology & Therapeutics*, 103(2):202–205.

Mill, J. S. (1874). *A system of logic*. Harper & Brothers Publishers, eighth edition.

Nagel, E. (1961). *The Structure of Science*. Harcout, Brace and World, Inc, New York.

Potochnik, A. (2017). *Idealization and the Aims of Science*. The University of Chicago Press, Chicago.

Raja, V. and Anderson, M. (2021). Behavior considered as an enabling constraint.

Ross, L. (2023a). Cascade verus mechanism: The diversity of causal structure in science. *The British Journal for the Philosophy of Science*.

Ross, L. (2023b). Causal Constraints in the Life and Social Sciences. *Philosophy of Science*.

Ross, L. (2023c). What is social structural explanation? A causal account. *Nous*, (https://doi.org/10.1111/nous.12446).

Ross, L. (2025a). Causal Pluralism: Primary and Secondary Features of Causation. *PhilSci Archive*.

Ross, L. (2025b). *Explanation in biology*. Cambridge University Press: Elements Series.

Ross, L. and Woodward, J. (2023). Irreversible (One-hit) and Reversible (Sustaining) Causation (Forthcoming). *Philosophy of Science*.

Ross, L. N. (2018). Causal selection and the pathway concept. *Philosophy of Science*, 85:551–572.

Ross, L. N. (2021). Causal concepts in biology: How pathways differ from mechanisms and why it matters. *The British Journal for the Philosophy of Science*, 72:131–158; https://doi.org/10.1093/bjps/axy078.

Ross, L. N. (2023d). Causal Control: A Rationale for Causal Selection. volume XXI. Minnesota Studies in Philosophy of Science.

Ross, L. N. and Bassett, D. S. (2024). Causation in neuroscience: Keeping mechanism meaningful. *Nature Reviews Neuroscience*.

Silberstein, M. (2021). Constraints on localization and decomposition as explanatory strategies in the biological sciences 2.0. In *Neural Mechanisms*.

Skow, B. (2014). Are There Non-Causal Explanations (of Particular Events)? *The British Journal for the Philosophy of Science*, 65(3):445–467.

van Frassen, B. (1980). *The Scientific Image*. Oxford University Press.

Woodbury, C. P. (2012). *Biochemistry for the Pharmaceutical Sciences*. Jones & Bartlett Learning.

Woodward, J. (2003). *Making things happen*. Oxford University Press, Oxford.

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, 25(3):287–318.

Woodward, J. (2011). Mechanisms Revisited. *Synthese*, 183(3):409–427.

Woodward, J. (2019). Some varieties of non-causal explanation. In Reutlinger, A. and Saatsi, J., editors, *Explanation beyond causation: Philosophical perspectives on non-causal explanation*. Oxford University Press.

Ylikoski, P. and Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies*, 148(2):201–219.