Causal Selection and the Pathway Concept

Lauren N. Ross*†

Most philosophical work on causal selection examines the selection of single causes and aims to clarify what grounds, if any, justify this selection. Such analyses overlook the fact that in scientific contexts multiple factors are often selected as the important causes of some outcome. This analysis examines a multicausal case in which factors in causal pathways are selected in biological explanations. This work provides a novel analysis of the pathway concept, its role in causal selection, and the rationale behind this selection. It is argued that this rationale is guided by principled considerations, which have been overlooked in the extant literature.

1. Introduction. Causal selection has to do with a distinction between mere background conditions and the "true" cause or causes of some outcome of interest. Early discussions of causal selection are found in the work of Mill and Lewis, who consider cases of selection in order to understand the rationale behind it (Mill 1874; Lewis 1986). This topic continues to receive attention in current philosophical work, particularly in the metaphysics of causation and philosophy of biology literature (Waters 2007; Franklin-Hall 2015; Schaffer 2016; Weber, forthcoming). These discussions focus on standard examples, which involve the selection of single causal factors (i.e., cases in which X is viewed as *the* cause of Y). For example, when explaining what caused a match to light why do we cite the fact that it was struck and not other factors such as the presence of oxygen? Why do we say that the car crash was caused by the drunk driver and not the ice on the road or the speed of the car? Why do we often select DNA as the cause of biological phenomena without

Received August 2017; revised February 2018.

*To contact the author, please write to: Department of Logic and Philosophy of Science, 3151 Social Science Plaza A, University of California, Irvine, CA 92697; e-mail: rossl@uci.edu.

†I would like to thank Jim Woodward, Ken Schaffner, Ken Waters, Marcel Weber, Marie Kaiser, Oliver Lean, Brian Hanley, and two anonymous reviewers for helpful feedback on this article. Work on this project was supported by the Philosophy Department at the University of Calgary and a grant from the John Templeton Foundation (ID 50191).

Philosophy of Science, 85 (October 2018) pp. 551–572. 0031-8248/2018/8504-0001$10.00
Copyright 2018 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

acknowledging the role of other causal factors? While some debates over causal selection continue, many claim that the philosophical field has reached a consensus—causal selection is "groundless," it lacks any principled rationale, and it is guided by preferences that are arbitrary, subjective, and audience relative (Schaffer 2016).

These analyses have two main shortcomings. First, they overlook the fact that in scientific (and ordinary life) contexts we often select multiple causes in our explanations, as opposed to single causal factors. Consider the biological sciences. It is rare to find biological explanations that cite single factors—most reference multiple causes as is seen in explanations of the cell cycle, physiological processes, the formation of complex phenotypes, and organism development. In this manner, standard philosophical treatments fail to accurately represent and pertain to common cases of causal selection in science. Second, in claiming that selection is "groundless" these analyses are at odds with our conception of science—we typically view causal selection and explanation in science as grounded in principled and justifiable reasons. If these accounts are correct why do we view these practices as guided by more than mere personal preferences? Somehow, scientists manage to reach significant consensus on the causally relevant factors for biological phenomena. How is this possible if selection is guided by arbitrary, personal, and subjective preferences, which are likely to differ across this scientific community?

This article provides an analysis of causal selection that addresses both of these shortcomings. It addresses the first by examining cases in which multiple causal factors are selected as the important causes of some outcome (i.e., cases of "multicausal selection"), and the second by arguing that these cases are guided by a principled rationale. This article focuses on examples of multicausal selection that involve a particular causal concept: the notion of a causal pathway. This concept is frequently found in biology, and it has received relatively little attention in the philosophical literature.[1] The pathway concept refers to a group of causal factors, which are ordered in a sequence that leads to some final outcome of interest. Examples of this concept from biology include gene expression pathways, cell-signaling pathways, metabolic pathways, cell differentiation pathways, and developmental pathways. In this article, I focus on the example of metabolic pathways, which capture processes by which living systems synthesize molecular materials. I use this example to provide a novel account of the features of the pathway concept, and I suggest that these features generalize to other uses of this concept in both biology and ordinary life.

In addition to clarifying features of the pathway concept, I argue that causal selection in these cases is guided by a principled rationale. I provide

1. Exceptions include Thagard (2003) and Schaffner (2008). This article expands on these analyses of the pathway concept.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

a novel account of this rationale that is understood in terms of the *causal control* that selected factors have over some outcome of interest. I characterize causal selection as, roughly, a three-step process. The first step involves identifying an explanatory target, while the second and third steps involve selecting causes that have causal control over this target. In the second step, factors are selected when they have at least some causal control, while in the third step they are selected on the basis of having very particular kinds of causal control.[2] This analysis suggests that causal selection involves both pragmatic and objective considerations, but in a different way than is indicated by mainstream accounts. This selection is not pragmatic in the pejorative sense of being guided by arbitrary, subjective, or audience-relative considerations. It is pragmatic in an alternative and creditable sense of being relative to a practical or useful goal—the goal of control. Furthermore, this selection is also objective in the sense that once this goal is specified there are objective facts and considerations about what means serve this goal. Whether a factor provides causal control depends on concrete features about the factor in the world as opposed to our personal preferences. While I examine this rationale in the metabolic pathway case, I suggest that a similar rationale guides causal selection in other cases in which the pathway concept is used, both in biology and ordinary life. In this manner, this analysis sketches a pattern of reasoning that generalizes to a variety of contexts.

This article is organized as follows. Section 2 introduces an example of the pathway concept in biochemistry in which this concept figures in explanations of glycolysis. Section 3 outlines the first two steps of causal selection, which involve specifying an explanatory target and identifying factors that have at least some causal control over this target. Before discussing the second step, I describe three characteristic features of the pathway concept. Section 4 clarifies an aspect of the second step, which has to do with the general speed of a factor's causal control. Section 5 discusses the third step of causal selection and the manner in which selected causes are distinguished from background conditions. Section 6 concludes.

2. Pathways in Biology: An Example from Biochemistry. Biochemistry and molecular biology focus on the chemicals, small molecules, and macromolecules that figure in the biological processes of living systems. In these fields there is significant interest in understanding how biological systems manufacture various products, including DNA, amino acids, proteins, sugars, and other substances required for life. When scientists explain these manufacturing processes they often appeal to the pathway concept. How exactly does the pathway concept figure in these explanations?

2. For an application of this framework to situations in which single causes are selected, see Ross (forthcoming).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

Consider glycolysis, a fundamental biological process in which the sugar glucose is broken down to acquire energy. In this process one molecule of glucose is converted into two molecules of pyruvate. The degradation of glucose supplies energy in the form of adenosine triphosphate (ATP), which is the "universal energy currency" of living organisms (Bunney, van Walraven, and de Boer 2001, 4249). When scientists explain and describe glycolysis they rely on the pathway concept—they divide this process up into 10 sequential steps that are represented along a causal chain, called the glycolytic pathway, shown in figure 1. Each step of this pathway involves the conversion of one chemical substance (or metabolite) into another, where these conversions are catalyzed by proteins called enzymes. For example, in the first step of glycolysis glucose is converted into glucose-6-phosphate (G6P) by the enzyme hexokinase. The product of this reaction (G6P) becomes the substrate for the next reaction, and so on down the pathway. The original glucose molecule is transformed into nine distinct intermediates before finally being converted into two molecules of pyruvate. In addition to each step requiring an enzyme some steps require energy (in the form of ATP or NAD^+) or other substances (like cofactors, phosphate groups, etc.).

When biologists discuss metabolic pathways they often rely on concepts and language from manufacturing contexts in ordinary life. Perhaps the most obvious example of this is their frequent comparison of metabolic pathways to factory assembly lines. Biologists claim that metabolic pathways are the "cellular equivalent of an assembly line in a factory," where some starting material goes through a sequence of modifications as it is turned into a final product (Chiras 2015, 45). Consider the assembly line manufacture of a car.

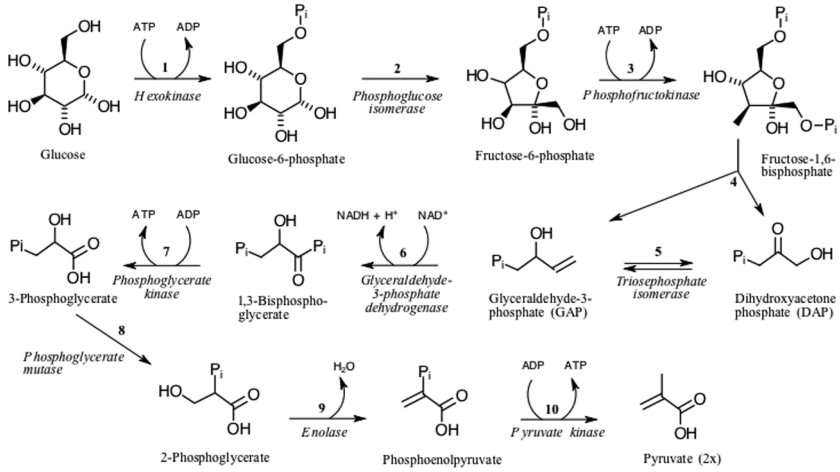


Figure 1. Glycolysis: a "model" metabolic pathway.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

This involves a sequence of steps in which a worker at each step modifies some material, passes it off to the next worker who does the same, and so on, until a complete car comes off the line. Similarly, metabolic pathways involve a sequence of steps that capture sequential modifications to some material substance until a final product is formed. In this analogy, enzymes are the "cellular equivalent" to the workers and tools in the assembly line (45, 66). Similar to workers and tools, enzymes are not consumed by the manufacturing process. Instead, they repeatedly perform their work at a given step as they move product down the line. In both metabolic pathways and assembly lines the product at one step become the substrate for the next, as material flows down the line.

The glycolytic pathway is considered a "model" metabolic pathway: it was the first metabolic pathway to be discovered, and it is the best understood (Nelson and Cox 2008, 528). An important feature of metabolic pathways is that they outline sequences of causal steps that are shared across most or all life forms. For example, the particular sequence of steps by which glucose is converted into pyruvate are nearly identical in all species. Thus, scientists state that "the pathways of enzyme-catalyzed reactions that act on the main constituents of cells—proteins, fats, sugars, and nucleic acids—are virtually identical in all living organisms" (25). For these reasons, metabolic pathways are said to capture a "unity" of life at the biochemical level (Rogers 2011, 49; Berg, Tymoczko, and Stryer 2012, 3). In other words, "from a metabolic point of view, the cellular processes that take place in a lion are only marginally different from those that take place in a dandelion" (Rogers 2011, 51–52).

3. Causal Selection: The First Two Steps. What causes are selected in explanations of metabolic processes and what rationale, if any, guides their selection? The first step of causal selection involves setting an explanatory target. In the case of glycolysis one phenomenon scientists want to explain is how a final product (pyruvate) is manufactured from some starting material. At the very least, scientists want to explain two outcomes associated with this target: the presence and absence of the product in question. This makes sense because an explanation of product manufacture should account for how the product is successfully made as opposed to not made at all.[3] In this manner, one common explanatory target in cases of metabolic manufacturing processes is this binary contrast of the presence and absence of a final product.

3. This analysis suggests a close connection between explanation and causation. In particular, it focuses on a specific type of explanation (i.e., causal explanation) in which causes are factors that explain their effects. If causes explain their effects, and causal relationships always pick out at least one cause and one effect, information about causal relationships can provide information about causal explanation and vice versa. This basic connection is assumed in standard analyses of causal selection in the literature (Lewis 1986; Waters 2007; Franklin-Hall 2015; Weber, forthcoming).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

3.1. Interventionist Framework. In order for a factor to be selected as a cause of some explanatory target, it needs to pass a second step of causal selection, which ensures that the factor has at least some causal control over this target. What does it mean to say that a factor has causal control? Consider a minimal notion of causal control, which is specified by the following interventionist criterion and motivated by Woodward's (2003) interventionist account of causation:[4]

> **Interventionist Criterion (IC)** X has causal control over Y if and only if there are circumstances S such that if some (single) intervention that changes the value of X (and no other variable) were to occur in S, then the value of Y or the probability distribution of Y would change.

In this framework, variables represent properties that can take on different values and arrows represent relationships of direct causal control. Variables are the relata of causal relationships, and they represent various properties of interest. Suppose that variable X is a light switch on the wall, which can take on one of two values (0,1) representing the switch being down or up. Variable Y represents a light bulb, which can take on the values (0,1) representing whether the light is off or on. The IC relies on the notion of an ideal intervention. An ideal intervention involves an unconfounded experimental manipulation of X with respect to Y in which changes in Y are produced by changes in X and not through any other variable. In other words, as shown in figure 2, this intervention (*a*) is not correlated with another variable W that has causal control over Y, (*b*) it does not directly have causal control over Y, and (*c*) it does not have causal control over any of the intermediate variables between X and Y. Intervening on the light switch in this manner reveals that it has causal control over the light being on or off. In explaining why the light is on or off, we appeal to the state of the switch, because it provides causal control over this outcome.[5]

4. I view the second step of the causal selection procedure—the interventionist criterion—as a mere reformulation of Woodward's account. This is a criterion that some factor should meet in order to be viewed as a cause of some outcome. This is not enough to rationalize causal selection because there should to be some procedure of (not just identifying causes but) distinguishing some causes from others. Part of my original contribution involves specifying what this is, which is captured in the three-step process I outline and, in particular, the third step of causal selection.

5. My analysis focuses on causal explanations of type-level, as opposed to token-level, outcomes. The relationship between these different explanations is nontrivial and has generated significant interest in the literature. While clarifying this relationship is not crucial for my argument in this article, I follow Woodward in viewing both as well understood within a basic interventionist framework, where type-level explanations often abstract from the particularities included in token-level explanations (Woodward 2003, 74–77).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

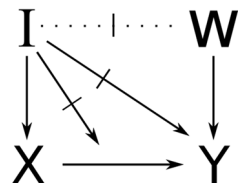


Figure 2. Causal control and ideal intervention (Woodward 2003).

3.2. Selecting Causes: Metabolites. In the glycolysis example there appear to be numerous factors with causal control over the manufacture of a final product.[6] At the very least, these factors appear to include metabolic substrates, enzymes, cofactors, and other regulatory molecules (ATP, NAD^+). The interventionist notion of causal control clarifies which factors should be selected as causes in this case and the rationale underlying this assessment. One set of factors that are cited in explanations of glycolysis is the sequence of metabolites represented along the glycolytic pathway. In this section, I provide an analysis of the reasoning that guides the selection of these factors. In discussing this rationale, I introduce three main features of the pathway concept. In this context, the pathway concept represents a sequence of factors with (i) causal control, (ii) material continuity, and a (iii) fixed order of steps that are shared across some large domain of systems.

First, consider glucose, which is the starting point of glycolysis and the starting material for this manufacturing process. Glucose has causal control over the manufacture of pyruvate, but its causal control is indirect in the sense that it operates through a set of metabolic intermediates. Each of these intermediates involves a relationship of direct causal control in which an upstream substrate controls manufacture of a downstream product. Consider the first step of glycolysis, where glucose is converted into G6P. Suppose that the background circumstances S contain the hexokinase enzyme and available ATP. In these circumstances, glucose provides causal control over G6P in the sense captured by the IC—changing the value of glucose from present to absent provides causal control over whether G6P is formed. In this situation, the presence of glucose results in G6P formation, while the absence of glucose prevents it. This is similar to a factory assembly line where the relevant workers and tools for a given step are present but the starting material for the step is either supplied or withheld. The availability of the

6. In this analysis I use the term "factor" to refer to a variable representing some property of interest. While the interventionist framework takes variables to be the relata of causal relationships, other accounts view events or event types as the appropriate relata. Such events will often be translatable into properties or quantities represented by variables, where variables can take on at least one of two values representing the occurrence and nonoccurrence of some event (Woodward 2003, 112).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

substrate controls whether the immediately downstream product is formed. The same relationship of direct causal control holds for the second step of glycolysis. The product of each step provides causal control over the next step in the pathway, and so on, until the final product is reached.

This identifies a chain of metabolites with causal control from glucose to pyruvate. These relationships of direct causal control can be represented with arrows and variables, as shown in figure 3, which represents the first three steps of glycolysis. In this diagram, X_1 represents glucose, X_2 represents G6P, X_3 represents F6P, and Y represents F1,6BP. If the enzymes, cofactors, and other regulatory molecules for this process were present, supplying glucose would result in the sequential production of each metabolite in the pathway until pyruvate was formed. In this environment, supplying or removing glucose provides causal control over whether pyruvate is formed. This is similar to a factory assembly line where all workers and tools are present but the initial starting material is either supplied or withheld. If starting material is supplied the final product can be made, and if it is not, production will not occur. In fact, this reasoning clarifies a strategy that biological systems use to control these manufacturing processes. Biological systems often sequester and compartmentalize starting materials as a way to keep them away from enzymes until the product in question is needed. Intervening on starting materials in this way controls whether the manufacturing process takes place.

In addition to causal control, the sequence of metabolites along the glycolytic pathway has a second important feature. These factors have a type of material continuity in which a significant amount of material from each upstream step carries over to the immediately downstream product. It makes sense that pathway factors have this feature in light of the manufacturing context, because the focus is on explaining how a material substance is changed into a material product. The pathway variables capture changes in a material entity as it undergoes this process. In fact, biologists exploit the material continuity of these processes in order to study and understand them. One way they do this is with tracer experiments, which involve supplying a biological system with tagged starting material and watching it flow through the pathway. By stopping this process intermittently and collecting tagged material, the metabolic intermediates and their location along the pathway can be identified. One of the first uses of this technique revealed the metabolic steps of photosynthesis in plants. In these experiments, plants were supplied with

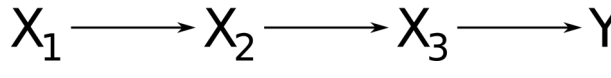


Figure 3. Sequence or chain of factors with causal control.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

tagged carbon isotopes ($^{14}CO_2$), and the tag was identified, successively, in a chain of metabolic intermediates.[7] Tracer experiments provide other types of information about metabolic pathways, such as the flow or flux of material through the pathway, and they are viewed as "perhaps the single most important technique in unravelling the complexities of metabolism" (Rogers 2011, 48).

A third feature of the pathway concept is that it captures the sense in which metabolic processes involve a fixed order of steps. What do I mean by a *fixed* order? In order to clarify this, consider the manufacture of a house. Building a house involves many steps that often need to take place in a particular order: the foundation needs to be poured before the house is framed, and the house needs to be framed before the roof is added. Engineers refer to these fixed relations as "mandatory dependencies" because the particular sequence of dependency relationships is fixed or mandatory. They claim that these relations are constrained by "hard logic" (Kerzner 2009, 501) because they identify a fixed ordering that is "required as a part of the nature of the work" and cannot be changed (Newell 2005, 49). These differ from "discretionary dependencies" that hold between outcomes that are ordered on the basis of personal preferences and that are not fixed by the nature of the project itself. As discretionary dependencies represent a preferential ordering that can be changed, engineers claim that this ordering involves "soft logic" or "preferential logic" (Kerzner 2009, 501).

Metabolic pathways involve sequences of steps that are similar to "mandatory dependencies" in the sense that they capture an ordering that is fixed by the nature of a biological system. This fixed order is sometimes referred to as a "processing chain" that involves "a number of steps which must occur in a certain order" (Bell 2008, 113). In terms of carrying out glycolysis, biological systems have to follow a particular order of steps. As shown in figure 3, glucose must first be converted into G6P, which must be converted into F6P, which must be converted into F1,6BP, and so on. While it is possible to imagine other routes through which glucose could be converted into F1,6BP the only available route in living systems is represented in figure 1. In both biochemistry and ordinary life manufacturing contexts, the pathway concept is used to represent a fixed order of steps in the manufacture of some final product. Not only is this order fixed, but in the case of glycolysis it is nearly universal across all life forms. This sequence of factors will be selected as causes of glycolysis in nearly all organisms because these factors provide causal control over this process in these systems.

7. As Zelitich states, "Calvin and Massini (1952) assumed that if a chain of intermediates exists in a sequence, and if during steady-state photosynthesis the uptake of $^{14}CO_2$ occurs at a constant rate, then 'we should find the label appearing successively in the chain of intermediates'" (1971, 90–91).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

4. The Second Step: Clarifications. Explaining glycolysis involves more than just appealing to a sequence of metabolites. These factors alone do not seem to provide the right kind of causal control over product manufacture. At the very least, their causal control seems to depend on the presence of enzymes and other factors associated with steps in the pathway. Relatedly, when scientists explain these processes they often appeal to these additional factors. If additional factors beyond metabolites are explanatorily relevant to glycolysis, on what grounds are they selected as causes of this process?

4.1. Selecting Enzymes: Interacting Causes. A straightforward answer to this question involves viewing enzymes and substrates as *interacting causes* with respect to product manufacture.[8] This can be clarified with an ordinary life example in which we want to explain what causes a flashlight (Y) to turn on and off. Consider two candidate causes, represented in figure 4, which include the flashlight's battery (X) and on/off switch (Z). The switch has causal control over whether the light is on or off, but only when the battery is present. When the battery is present, intervening on the switch turns the light on and off. The battery also has causal control over the flashlight's being on / off, but only when the switch is left in the on position. When the switch is left on, intervening on the battery (by removing it and replacing it) causes the flashlight to turn on and off. Both of these factors have causal control over the outcome, but their control depends on the other factor. This is what it means to say that these factors are interacting causes with respect to some outcome. Explaining whether the flashlight is on or off involves appealing to both factors, because both are required to gain causal control over this contrastive focus. Notice that the causal control of these factors individually is uneven—each factor alone can reliably turn the flashlight off, but both are required to turn it on. As the contrastive focus involves both states of the flashlight, and both factors are required for one of these states (i.e., the on state), gaining complete causal control over this target requires both factors.

It is natural to view metabolic substrates and enzymes as interacting causes with respect to product manufacture. However, unlike the flashlight example, the metabolic case involves a sequence of many interacting causes, as shown in figure 5. Each step in the pathway identifies a substrate-enzyme pair that are interacting causes with respect to the immediately downstream product. The enzyme and substrate each have causal control over formation of the product where the causal control of each factor depends on the other. When the relevant enzyme is present, changing whether the substrate is available or not controls whether the product is formed. Likewise, when the substrate is present, changing whether the enzyme is available or not controls

8. For more on interacting causes, see Spirtes, Glymour, and Scheines (2000, 24).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

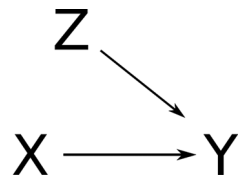


Figure 4. Directed acyclic graph of flashlight battery (X), switch (Z), and bulb (Y).

whether the product is formed. Explaining glycolysis involves citing the pair of interacting causes at each step in the ordered sequence from beginning to the end. This clarifies why scientists cite sequences of enzymes and substrates in these explanations and how there is a principled rationale for doing so—these are factors that provide causal control over formation of the final product.

Scientists exploit the causal control of enzymes to study and understand metabolic processes. One such strategy involves examining biological systems with defective enzymes that lead to metabolic disorders or "inborn errors of metabolism." When a metabolic enzyme is absent or nonfunctional, this blocks a step in the pathway resulting in a buildup of upstream material. The buildup of particular types of material can be used to identify the metabolites along the pathway and the enzymes they are associated with. In fact, the initial discovery of many metabolic disorders has involved identifying a pathological accumulation of metabolic materials. Scientists compare these disorders to disrupted factory assembly lines. They state that nonfunctional enzymes "cause an accumulation of an intermediate, in much the same way as a missing worker on an assembly line halts production at that point" (Windelspecht 2007, 55). Other experiments involve selectively inhibiting enzymes to identify their substrates and location along the pathway. As enzymes are proteins produced by genes, one way to manipulate them is to introduce gene mutations into an organism.[9] This experimental strategy was used in twentieth-century research that helped establish the fields of molecular biology and biochemistry. For example, these experimental methods were used in Beadle and Tatum's (1941) Noble Prize–winning work on metabolic pathways in *Neurospora* and Beadle and Ephrussi's (1937) research on biosynthetic eye color pathways in *Drosophila*. These studies led to the one gene–one enzyme hypothesis, which provided one of the first and most convincing explanations for how genes influence phenotypes, that is, through enzymes, which regulate metabolic pathways (Morange 1998, 21). Hull's (1974) diagram in figure 6 depicts these relationships between genes, enzymes, and metabolites and bears some similarities to the causal graph in fig-

9. For more on this, see Waters's (2004) discussion of the "genetic approach."

https://doi.org/10.1086/699022 Published online by Cambridge University Press

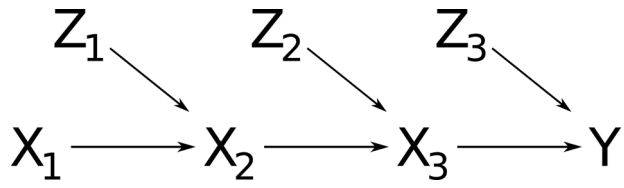


Figure 5. Directed acyclic graph of metabolites (X_1, X_2, X_3, Y) and enzymes (Z_1, Z_2, Z_3).

ure 5. (I discuss the role of genes in these explanations in the next section.) In these experiments, metabolic pathways are studied with strategies that exploit the causal control of enzymes, the location of enzymes and substrates in a causal chain, and the flow of material along the pathway.

Metabolites and enzymes receive the most attention in explanations of metabolic processes. However, other factors are also frequently cited, such as enzyme cofactors and accessory substrates (e.g., ATP, NADPH). The interacting cause framework provides a promising approach to include many of these other factors—in many cases, they are additional interacting causes that allow for causal control over product formation.

4.2. General Speed. I have used the notion of an interacting cause to clarify why enzymes and substrates are selected as causes of metabolic processes. This is because intervening on both of these factors allows for causal control over the manufacture of a final product. There is a puzzle with this analysis. In the case of glycolysis, there is a sense in which metabolic substrates have causal control over product formation without enzymes. Let me clarify. Glycolysis is an exergonic reaction in the sense that it results in a net release of energy. Exergonic reactions are referred to as "spontaneous processes" because their thermodynamic favorability means that they can occur spontaneously without any outside factors or catalysts. This means that glucose can transform into pyruvate without the help of enzymes. If glucose and

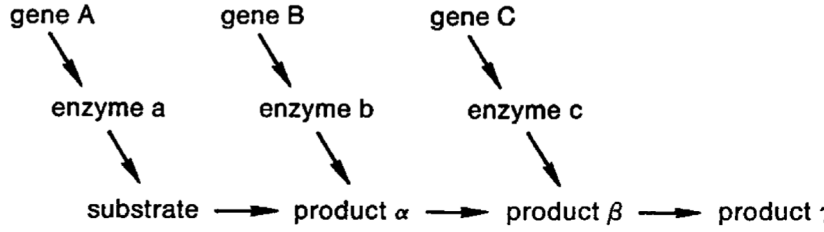


Figure 6. Metabolic pathway (Hull 1974, 29).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

the glycolytic intermediates have causal control over pyruvate manufacture without enzymes, why are enzymes cited in explanations of this process? It turns out that just knowing that a process is spontaneous provides no information about the speed that it takes the process to occur. In the case of glycolysis, the transformation of glucose into pyruvate would take somewhere on the order of billions of years without enzymes. The following quote discusses this point in the context of glycolysis:

> Many of us, for example, consume substantial amounts of sucrose—common table sugar—as a kind of fuel, usually in the form of sweetened foods and drinks. The conversion of sucrose to CO_2 and H_2O in the presence of oxygen is a highly exergonic process, releasing free energy that we can use to think, move, taste, and see. However, a bag of sugar can remain on the shelf for years without any obvious conversion to CO_2 and H_2O. Although this chemical process is thermodynamically favorable, it is very slow! Yet when sucrose is consumed by a human (or almost any other organism), it releases its chemical energy in seconds. The difference is catalysis. Without catalysis, chemical reactions such as sucrose oxidation could not occur on a *useful* time scale, and thus could not sustain life. (Nelson and Cox 2008, 183, emphasis added)

This helps to clarify the type of causal control that matters in the second step of causal selection. Selected factors need to have causal control that operates within a timescale set by the context of inquiry. If a factor fails to meet this standard it will not be selected because its control does not show up at the relevant timescale and, thus, is of little (to no) use in controlling and explaining the outcome within this time frame. In the case of metabolic processes, the relevant timescale is milliseconds to minutes. Metabolic processes need to operate at this scale in order for living organisms to sustain life and functionality. Organisms need to react to their environments, break down materials, and grow at particular speeds in order to survive. Metabolites alone do not have causal control that meets this standard. Without enzymes, they manufacture products far too slowly with respect to the time frame of interest. Enzymes are selected in addition to metabolites, because they increase the speed of causal control that these factors have over product formation, such that it occurs within the relevant timescale of interest. This consideration can be formulated as follows. In order for factors to be selected as the causes of an outcome, they need to meet a *general speed principle*, which ensures that the speed of their causal control over the outcome fits within some general time range specified by the context of inquiry. This clarifies why both enzymes and metabolites are cited in explanations of glycolysis, as opposed to just metabolites alone. It is not just that both are required to gain causal control

https://doi.org/10.1086/699022 Published online by Cambridge University Press

over pyruvate manufacture but that their causal control meets this general speed principle.[10]

Another example helps to support these points and suggests their general role in causal reasoning. Consider the conversion of diamond to graphite, which is another spontaneous process. Although this process is spontaneous, it takes over a billion years to manifest at normal conditions. In ordinary life and scientific contexts, we usually deny that this process occurs. Scientists claim that while "diamonds are thermodynamically unstable with respect to graphite . . . the rate of conversion . . . is so slow, that *for all practical purposes, it does not occur*" (Woodbury 2012, 10, emphasis added). This reaction is viewed as "*negligible* at standard conditions" (Fecht, Brühne, and Gluche 2014, xi, emphasis added) because "it cannot be observed over any *practical* period of time" (OpenStax 2016, 236, emphasis added). We tend to deny that diamond alone decomposes into graphite, because this decomposition does not occur on a timescale relevant to human life, which is our default timescale in these situations. The only way to see this decomposition at this timescale is to subject diamond to extremely high heat or high pressure. When we want to explain this process on a timescale relevant to human life, we appeal to all these factors (diamond and high heat or high pressure) because they provide causal control at this scale. If the context of inquiry assumed a larger timescale (e.g., one at the geological or cosmological level), we would be more likely to just cite diamond as the sole cause of this conversion, because its causal control does manifest at these larger time frames. This reveals that our selection of causes is relative to an assumed time frame of interest. This is likely to be a general consideration that applies to causal reasoning in many cases, both within and outside of biology.

The general speed principle clarifies pragmatic and objective considerations that figure in causal selection. If we want to control the manufacture of a product in a living system or in some ordinary life context, identifying factors that take billions of years to influence these systems is not useful to us. We can target and intervene on these factors, but we will never see and enjoy the results of their causal control. While there is a sense in which metabolites alone have causal control over pyruvate, their control is not "useful" for biological systems because it is too slow to operate within the required time frame of interest. The same goes for the diamond-graphite example. While diamond alone does technically transform into graphite, this causal relationship is not useful to us because it does not occur within a "practical period of time" given various constraints (like the length of our life spans). This is not a "merely

10. In other words, selected causes should not have causal control that is too slow relative to some time frame of interest. They also should not have causal control that is too fast. Selected causes should not provide control that is faster than the speed of light, as this would violate a law of nature (Woodward 2016).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

pragmatic" consideration in the sense of being arbitrary, subjective, or audience relative. It is a pragmatic consideration in the sense of utility and usefulness—causes that take billions of years to produce their effects are (typically) not going to be of use to us, especially not if our goal is to control or change outcomes during our lifetimes.[11] This suggests that our selection of causes is relative to a practical goal in which causes need to meet some general time frame specified by the context of interest. Once this time frame is specified, there are objective facts about what factors meet it and what factors do not. Whether a factor meets this specification is not dictated by our personal or subjective preferences. Our personal preferences do not determine how fast high pressure converts diamond into graphite or how fast enzymes break glucose down into pyruvate. The speed of the causal control these factors exhibit is determined by objective facts about these entities and their nature in the world.

5. The Third Step. So far I have discussed the first two steps of causal selection. The first step involves setting an explanatory target, and the second step involves identifying factors that have at least some causal control over this target. In order for a factor to have some causal control, it needs to meet the IC and the general speed principle. These first two steps make up a significant part of the rationale that guides causal selection, and they are likely to figure in causal selection more generally. In the case of glycolysis, these steps help clarify why metabolites and enzymes are selected as causes. However, there are further reasons for why these factors are selected, which have to do with the very particular types of causal control that they have. The third step of causal selection captures the sense in which selected factors are expected to have very particular types of causal control. These types of control are often associated with important features of the explanatory target. In section 3, the explanatory target was characterized as the presence and absence of a final product. There are further features of this target that scientists want to explain. For example, they often want to explain (1) what specific type of product was formed and (2) the rate of product manufacture. These questions are related to the manufacturing character of this example, and they are an-

11. Consider that my use of "pragmatic" compares to the notion found in van Frassen's work. For van Frassen (1980), pragmatic features of explanation concern the fact that they are relative to a context and contrast class that are set by our interests. My account admits that such factors figure in identifying explanatorily relevant information—setting the contrast class is the first step of causal selection, and the general speed of a factor's causal control can be thought of as set by our specification of the relevant context. However, while these factors are present in my account I mean something different by "pragmatic" than simply the role of personal interests in setting the context. In this analysis, I characterize "pragmatic" in terms of utility and usefulness, which relate to general goals found in many contexts.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

swered by assessing two types of causal control: the specificity of causal control and the particular speed of causal control.

5.1. Specificity. In the glycolysis example, scientists want to explain how a particular product is produced. Biological organisms need to manufacture particular products in order to carry out various functions and sustain life. Minor alterations in the makeup of a product can completely change its functional capacity or render it nonfunctional in ways that lead to deleterious outcomes and even death. For example, minor changes in the stereochemical structure of amino acids and the amino acid sequence of proteins are known to cause various diseases. For these reasons it matters that a very particular product is produced, and scientists want to explain what causes the manufacture of one product and not another.

Each step in the glycolytic pathway involves the conversion of an upstream substrate into a downstream product. For every upstream substrate there are many downstream products that it could be converted into. For example, the first step of glycolysis captures the conversion of glucose into G6P, but glucose can also be converted into many other types of products. It could be converted into G5P, or it could be rearranged into the five-carbon sugar fructose or split into two three-carbon molecules (among many other possibilities). Factors that explain this first step of glycolysis should explain why glucose is converted into G6P and not into these other products. This explanation is provided, in part, by the fact that the enzyme at this step has specific causal control. There are two important ways in which enzymes have specific causal control: first, (*a*) they only interact with specific substrates as opposed to others, and second, (*b*) they only produce specific products as opposed to others.[12] Both of these notions are suggested by scientists's claims that "enzymes are highly specific both in the reactions they catalyze and in their choice of reactants which are called substrates" (Berg et al. 2012, 220). I will briefly discuss these two types of specificity and their relevance to causal selection.

First, *a* represents specificity at the level of interacting causes in the sense of having to do with the particular factors that interact together to provide causal control over an outcome. The enzyme at the first step of glycolysis (hexokinase) exhibits this specificity in the sense that it only interacts with a particular type of substrate (i.e., six-carbon sugars). This enzyme interacts with six-carbon sugars and not with other molecules like five-carbon sugars, four-carbon sugars, or six-carbon nonsugars. But, *b* has to do with specificity in the sense of producing a particular outcome. Enzymes exhibit specificity

12. These types of specificity differ from the notion of fine-grained specificity discussed by Waters (2007) and Weber (forthcoming). The types of specificity I rely on here relate to a form of one-to-one specificity discussed by Woodward (2010).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

in this sense because they produce particular products from some range of potential products. The enzyme at the first step of glycolysis produces G6P, as opposed to G5P, fructose, or two three-carbon molecules. This enzyme adds a specific component (a phosphate group) to a specific location (carbon 6) of a specific substrate (glucose), which results in a specific product (G6P). Both types of specificity, *a* and *b*, are relative to some class of alternatives. In *a* the enzyme interacts with a particular substrate among a set of alternatives, and in *b* the enzyme makes a particular product among a set of alternatives.

The glycolytic pathway, like other metabolic pathways, represents a causal sequence of specific metabolic substances. These substances are particular chemicals and biomolecules, which are represented in fine-grained molecular detail. Impressively, this sequence of metabolites is common to glycolysis in nearly all biological systems. What this means is that explaining glycolysis in any living organism will involve appealing to these factors because they provide causal control over this process in all organisms. The universal relevance of these causal factors is partly why they receive so much attention and why they are so useful in biological explanation. It also explains why these pathways are said to be "conserved" across living systems and representative of a "unity of life" at the biochemical level. But, enzymes in these pathways are not universal in this sense and are not represented in fine-grained molecular detail. This is because there is no single molecularly distinct enzyme that catalyzes each step in all organisms. Instead, there are families of molecularly diverse enzymes in which each enzyme in the family catalyzes the same chemical reaction. The enzymes in these families are called isozymes or isoforms, as they are "different proteins that catalyze the same reaction" (Nelson and Cox 2008, 583). The criterion that groups together a family of enzymes is their specificity in the senses of *a* and *b* and not their shared molecular detail. If the goal is to explain glycolysis in all living organisms, the enzyme variable represents an isoenzyme family. Alternatively, if the goal is to explain glycolysis in a particular system (e.g., a tissue, cell), the enzyme variable represents a particular enzyme from the family.[13] In both cases, the enzyme variable represents a factor that has specific causal control in the senses of *a* and *b*; it just differs in whether it represents a molecularly distinct enzyme or an enzyme family.

Enzymes are selected in explanations of metabolic manufacturing processes, not just because they have causal control over product formation but because they have causal control that is specific, in the senses of *a* and *b*. These types of causal control explain why specific causal factors interact and which specific products they produce. In particular, the specificity

13. For more on explanations of how particular enzymes catalyze chemical reactions and their role in "reaction mechanisms," see Goodwin (2012).

https://doi.org/10.1086/699022 Published online by Cambridge University Press

of enzymes in the sense of *b* explains how an enzyme-substrate pair produces a specific product from some range of alternative products. As product specificity matters for biological functioning and makes up an important feature of the explanandum in this context, enzymes are selected as causal factors, in part, because they explain this feature of the outcome of interest.

5.2. Particular Speed. In the glycolysis case, another feature of the explanatory target that scientists want to explain is the rate of product manufacture. Biological organisms have to make (and degrade) products at very particular speeds, in order to sustain life and proper functioning. For example, immediate and sustained physical activity requires the rapid conversion of energy-rich products into available energy. Additionally, responding to environmental stimuli requires extremely fast nerve signaling, which is facilitated by the speed of neurotransmitter degradation. In other cases, the functionality of a process requires that it not operate too fast relative to some time frame. For example, enzymes that cleave DNA (e.g., restriction endonucleases) need to be slower than those processes that protect it (e.g., methylation), otherwise DNA would be destroyed before it could figure in transcription or other activities (Berg et al. 2012, a9). For these reasons the particular rate of production matters, and scientists often want to explain what causes these processes to operate at one rate and not another.

What explains the rate at which a metabolic process manufactures some final product? In metabolic pathways the rate of product manufacture depends on the particular speed of the enzymes in the pathway. This is similar to a factory assembly line in the sense that the speed of making a final product depends on the speed of each step in the line. As these steps occur in sequential fashion—where a downstream step only occurs once an upstream step is completed—production is no faster than the slowest step. In fact, if each step in the pathway is saturated with material the rate of production is approximately determined by the slowest step, which is called the "rate-determining" or "rate-limiting" step. The enzyme at this step is called the "pacemaker enzyme" because it sets the pace of production and the flow (or flux) of material through the pathway. Because of their important role in determining the rate of manufacture, biological systems often target pacemaker enzymes in an effort to control flux through a pathway. In order for the pacemaker enzyme to determine total rate of product manufacture, all other pathway steps should have their respective enzymes, which is why all enzymes in a pathway are often cited in these explanations.

This analysis has characterized causal selection as a three-step procedure, which clarifies why particular factors are selected as the main causes of some outcome. What about background factors that are not selected? What does this analysis say about the distinction between these factors and the selected causes? Consider a set of factors that seem to be causally relevant to meta-

https://doi.org/10.1086/699022 Published online by Cambridge University Press

bolic outcomes but that are not typically cited in explaining these outcomes: temperature, pH, and genes. All of these factors might seem causally relevant to these metabolic processes in the sense that their presence is necessary for them to occur. Metabolic outcomes require particular ranges of temperature and pH, and they require genes, which produce the enzymes that operate at each step. Why are these factors not typically cited as important causes of these processes? My analysis suggests that they are not selected because they either fail to (*a*) provide any causal control over the outcome of interest or fail to provide (*b*) particular kinds of causal control. This is captured by the fact that these factors fail to advance through the causal selection procedure. Temperature and pH are backgrounded (or not selected) because they do not pass the second step—they do not provide control over the final metabolic outcome of interest. Both need to be present in a particular range for these processes to take place, but alternations of these conditions do not significantly influence metabolic outcomes in living organisms.[14] Genes, however, are backgrounded when they do not pass the final step—in particular, when they fail to provide causal control that meets the particular speed criterion. However, genes are sometimes cited in explanations of metabolic outcomes. Why is this?

Scientists sometimes appeal to genes in explaining metabolic processes, but they distinguish their explanatory relevance from enzymes. Scientists claim that genes have "coarse control" over metabolic processes in the sense that their control is "less immediately responsive" (Rogers 2011, 64, 178). Enzymes, however, have "fine control" over these processes because they have "immediate effects" on product manufacture (63, 170). This distinction can be understood with the notion of the particular speed of a factor's causal control, and it clarifies why genes are sometimes cited in explanations of metabolic processes and sometimes not. Relative to the speed at which enzymes catalyze reactions, genes take a much longer time to produce proteins. If the goal is to explain the speed of a metabolic manufacturing process, genes are unlikely to be cited because they produce enzymes at slow rate that does not capture or explain the speed of these processes. More specifically, it is often assumed that genes have already made enzymes that are present and prepared to catalyze a reaction. In this case, the faster "fine control" of enzymes explains the responsive nature of these processes. That being said, genes do control the amount of available enzyme, which does influence the rate of manufacture in a longer-term sense. If the goal is to explain the changes in a metabolic process over a longer time frame, genes are more likely to be cited because their slower "coarse control" operates at this scale and explains these changes.

14. In fact, altering these conditions outside very narrow ranges shuts these processes down full stop, as such ranges are required for life.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

This section has clarified additional reasons that guide the selection of enzymes as causes of metabolic processes. These reasons have to do with particular types of causal control that these factors have over an explanatory target. Enzymes are cited in addition to substrates because they have particular types of causal control that explain important features of this target. The specificity of enzymes and the particular speed of their causal control explain how metabolic processes create specific products and the rate at which these products are manufactured.

6. Conclusion. Most philosophical work on causal selection examines cases in which single causes are cited, and it aims to clarify what grounds, if any, justify this selection. Such analyses overlook the fact that in scientific (and ordinary life) contexts, multiple factors are often selected as the important causes of some outcome. This article has examined one such multicausal case in which factors in causal pathways are cited in biological explanations. This article provides a novel analysis of the pathway concept, its role in causal selection, and the rationale behind this selection. In the context of biochemistry, metabolic pathways refer to a sequence of factors with (i) causal control, (ii) material continuity, and (iii) a fixed order, which represents a manufacturing process that is universal or shared across some large domain of systems. These features are not just characteristic of pathways in biochemistry—they are defining features of the pathway concept in other areas of biology and even ordinary life contexts. This applicability to ordinary life cases is easily seen in scientists's comparison of metabolic pathways to manufacturing processes such as factory assembly lines. Scientists exploit our familiarity with these ordinary cases to easily communicate and highlight distinct features of these complex biological systems. Furthermore, the same use of the pathway concept is found throughout biology as seen in cases such as cell division (e.g., the cell cycle), cellular differentiation (e.g., stem cell pathways), and organism development (e.g., developmental pathways). In these examples scientists use the pathway concept to refer to a sequence of changes that these material substances undergo as they develop into some final, mature system.

This article has elucidated features of the pathway concept in order to provide an analysis of the role of this concept in causal selection. In these cases, causal selection is well understood in terms of *causal control*—factors are selected as causes of metabolic outcomes when they provide particular types of causal control over them. These types of control relate to features of these processes that scientists want to explain, such as why a particular product is formed (specific control) and the rate at which it is manufactured (speed of control). An interest in causal control and particular features of the explanandum uncovers the principled reasoning that guides this selection process. My analysis of this reasoning contrasts with mainstream accounts of causal selection, which view it as "groundless" and guided by considerations that

https://doi.org/10.1086/699022 Published online by Cambridge University Press

are audience relative, subjective, and arbitrary. Mainstream accounts label these considerations as "pragmatic" in a pejorative sense that is distinguished from objective or principled reasons. This article suggests a different way to understand how causal selection is guided by both pragmatic and objective considerations. It suggests that causal selection is pragmatic in the creditable sense of being relative to a practical and useful goal (i.e., the goal of control). Identifying factors with control is useful—it provides a means of changing and regulating an outcome of interest. Once this pragmatic goal is set there are objective facts and considerations about what means serve this goal and which do not. Whether a factor has causal control over an outcome and the particular types of control that it has are not determined by arbitrary or capricious decisions but by objective features about the nature of systems and their operation in the world. Uncovering these principled reasons and the common use of the pathway concept suggests that this causal selection procedure is not limited to biochemistry but that it generalizes to causal selection in various contexts.

REFERENCES

Beadle, G. W., and B. Ephrussi. 1937. "Development of Eye Colors in *Drosophila*: Transplantation Experiments on the Interaction of Vermilion with Other Eye Colors." *Genetics* 22 (1): 65–75.

Beadle, G. W., and E. L. Tatum. 1941. "Genetic Control of Biochemical Reactions in *Neurospora*." *Proceedings of the National Academy of Sciences of the USA* 27 (11): 499–506.

Bell, G. 2008. *Selection: The Mechanism of Evolution*. 2nd ed. Oxford: Oxford University Press.

Berg, J. M., J. L. Tymoczko, and L. Stryer. 2012. *Biochemistry*. 7th ed. New York: Freeman.

Bunney, T. D., H. S. van Walraven, and A. H. de Boer. 2001. "14-3-3 Protein Is a Regulatorchloroplast ATP Synthase." *Proceedings of the National Academy of Sciences of the USA* 98:4249–54.

Chiras, D. D. 2015. *Human Biology*. 8th ed. Burlington, MA: Jones & Bartlett Learning.

Fecht, H. J., K. Brühne, and P. Gluche. 2014. *Carbon-Based Nanomaterials and Hybrids*. Boca Raton, FL: Pan Stanford.

Franklin-Hall, L. R. 2015. *Explaining Causal Selection with Explanatory Causal Economy: Biology and Beyond*. Dordrecht: Springer.

Goodwin, W. 2012. *Mechanisms and Chemical Reactions*. New York: Elsevier.

Hull, D. L. 1974. *Philosophy of Biological Science*. Upper Saddle River, NJ: Prentice-Hall.

Kerzner, H. 2009. *Project Management*. Hoboken, NJ: Wiley.

Lewis, D. A. 1986. "Causal Explanation." In *Philosophical Papers*. Oxford: Oxford University Press.

Mill, J. S. 1874. *A System of Logic*. 8th ed. New York: Harper & Bros.

Morange, M. 1998. *A History of Molecular Biology*. Cambridge, MA: Harvard University Press.

Nelson, D. L., and M. M. Cox. 2008. *Principles of Biochemistry*. 5th ed. London: Freeman.

Newell, M. W. 2005. *Preparing for the Project Management Professional (PMP) Certification Exam*. 3rd ed. New York: American Management Association.

OpenStax. 2016. *Chemistry*. Houston: OpenStax.

Rogers, K. 2011. *The Chemical Reactions of Life*. New York: Britannica.

Ross, L. N. Forthcoming. "Causal Control: A Rationale for Causal Selection." In Minnesota Studies in Philosophy of Science. Minneapolis: University of Minnesota Press.

Schaffer, J. 2016. "Metaphysics of Causation." In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. https://plato.stanford.edu/entries/causation-metaphysics/.

https://doi.org/10.1086/699022 Published online by Cambridge University Press

Schaffner, K. F. 2008. "Etiological Models in Psychiatry." In *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, ed. K. S. Kendler and J. Parnas, 1–49. Baltimore: Johns Hopkins University Press.
Spirtes, P., C. Glymour, and R. Scheines. 2000. *Causation, Prediction, and Search*. 2nd ed. Cambridge, MA: MIT Press.
Thagard, P. 2003. "Pathways to Biomedical Discovery." *Philosophy of Science* 70 (2): 235–54.
van Frassen, B. 1980. *The Scientific Image*. Oxford: Oxford University Press.
Waters, C. K. 2004. "What Was Classical Genetics?" *Studies in History and Philosophy of Science* 35 (4): 783–809.
———. 2007. "Causes That Make a Difference." *Journal of Philosophy* 104 (11): 551–79.
Weber, M. Forthcoming. "Causal Selection versus Causal Parity in Biology: Relevant Counterfactuals and Biologically Normal Interventions." In *Philosophical Perspectives on Causal Reasoning in Biology*, ed. C. K. Waters and J. Woodward. Minnesota Studies in Philosophy of Science 21. Minneapolis: University of Minnesota Press.
Windelspecht, M. 2007. *Genetics 101*. Westport, CT: Greenwood.
Woodbury, C. P. 2012. *Biochemistry for the Pharmaceutical Sciences*. Burlington, MA: Jones & Bartlett Learning.
Woodward, J. 2003. *Making Things Happen*. Oxford: Oxford University Press.
———. 2010. "Causation in Biology: Stability, Specificity, and the Choice of Levels of Explanation." *Biology and Philosophy* 25 (3): 287–318.
———. 2016. "Causation and Manipulability." In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. https://plato.stanford.edu/entries/causation-mani/.
Zelitich, I. 1971. *Photosynthesis, Photorespiration, and Plant Productivity*. New York: Academic Press.

https://doi.org/10.1086/699022 Published online by Cambridge University Press