Causal selection in biochemistry: Making things by making things happen

Lauren N. Ross

Causal selection has to do with a distinction between mere background conditions and the "true" causes of some outcome of interest. Mainstream philosophical views claim that causal selection is "groundless" in the sense that it lacks any type of principled rationale (Schaffer 2016; Mill 1874; Lewis 1986). I argue against this position in the context of biochemistry where causal factors are selected in explanations of metabolic processes. These factors are selected on the basis of a principled rationale, which is best understood in terms of the causal control that they provide over an outcome of interest.

1 Introduction. In the philosophical literature it is commonly claimed that to provide a causal explanation of some phenomenon of interest involves citing its causes. Of course, many phenomena appear to involve a large or seemingly infinite number of causal factors. The causal history of a rocket's launch into space or a caterpillar's metamorphosis into a butterfly seem to include numerous factors that interact in complex ways. Traditional philosophical views claim that "the" explanation or "the whole" explanation of some outcome involves accounting for the complete set of factors in its causal history (Lewis 1986, 218-9).

Proponents of these views admit that we can't feasibly identify, much less cite, this complete set of causal factors in our explanations. In fact, we rarely seem compelled to do this. We are usually highly selective about the causes that we view as explanatorily relevant to an outcome. In explaining a brush fire we are likely to cite the arsonist or her fire-starting implement, while backgrounding other factors like the presence of oxygen, dry brush, or lack of recent rainfall. This involves causal selection, which has to do with distinguishing the "true" causes of an outcome from mere background conditions. Mainstream philosophical views claim that causal selection lacks any objective rationale and claim that it is guided by "pragmatic" considerations, which are arbitrary, subjective, and audience-relative (Lewis 1986; Mill 1874). Mill famously claimed that causal selection lacks any "scientific ground" as it is guided by "capricious" considerations (Mill 1874, 238). Similarly, Lewis claims that causal selection is driven by "invidious principles" (Lewis 1986, 162). He states that we select causes because we find them to be good or bad, because we have control over them, or simply because we want to talk about them. These views are widely accepted as having "won the field" and causal selection is "now generally dismissed as groundless" (Schaffer 2016).

These mainstream claims raise a number of puzzles. We often view ourselves as providing successful causal explanations in ordinary life and scientific contexts. We think we can explain how rockets are launched into space, how caterpillars develop into butterflies, and how brush fires start. In our explanations of these phenomena we often cite an economy of carefully chosen factors. If providing a proper or complete causal explanation involves citing all of the factors in the causal

*Word count: 9,312

†Acknowledgements: I would like to thank Jim Woodward, Ken Schaffner, Ken Waters, Marcel Weber, Oliver Lean, and Marie Kaiser for helpful feedback on this paper.

†The research and writing of this paper was supported in part by a grant from the John Templeton Foundation (ID 50191, From Biological Practice to Scientific Metaphysics).

history of some outcome, why do we rarely (if ever) strive to do this? Why do we often view explanations that cite few causal factors as legitimate and successful? Finally, if causal selection is inherently subjective, why do we view the adequacy of an explanation as having to do with more than personal preferences?

One way to address these questions is to examine the causal concepts, reasoning, and methods that figure in our explanations. The biological sciences provide helpful case studies for this project, because biologists have a strong interest in providing causal explanations, the phenomena they want to explain often involve numerous causal factors, and many of their explanations are viewed as highly successful. One causal concept that commonly figures in biological explanation and which has received little philosophical attention is the notion of a causal pathway.[1] Examples of this concept include metabolic pathways, biosynthetic pathways, cell differentiation pathways, and developmental pathways. In this paper I focus on metabolic pathways, which capture processes by which living systems manufacture molecular materials. These processes involve a sequence of steps, much like a factory assembly line, where an upstream substrate goes through a sequence of changes before forming a final downstream product. The pathway concept represents this sequence of causal steps and the limited set of causal factors that are selected as explanatorily relevant for these processes. Biologists reach significant agreement on exactly which causal factors make up these pathways, which ones do not, and what they explain. These pathway explanations are associated with significant scientific achievements and they raise similar questions to those suggested above. If providing a complete causal explanation of some outcome involves citing all of the factors in its causal history, why do biologists view themselves as providing legitimate explanations without doing this? Why are they so selective in the factors that they cite? If their selection of these factors is arbitrary or unsystematic, how are they able to reach agreement on these causes and why do their explanations seem so successful?

In this paper I argue that there is a principled rationale that guides causal selection for explanations of metabolic processes. In these cases, causal factors are selected on the basis of considerations that have to do with causal control. I provide an account of the types of causal control that guide this selection and I suggest that it involves both pragmatic and objective considerations, which have been overlooked in the current literature. In my analysis, 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 with respect to this target. In the second step, factors are selected when they have at least some causal control over the contrastive focus. In the third step they are selected on the basis of having very particular kinds of causal control.

This paper is organized as follows. The second section reviews an example of the pathway concept in biochemistry, where a metabolic pathway is used to explain glycolysis. Section three introduces the first two steps of causal selection: specifying the explanatory target and identifying factors that have at least some causal control over this target. The fourth section clarifies an additional feature of the second step, which has to do with the general speed of a factor's causal control. The fifth section discusses the third step of causal selection, which involves assessing whether factors have particular types of causal control. Throughout these sections I discuss unique features of the pathway concept and how this concept is used to organize causal factors that are selected through this process.

[1]For discussions of the pathway concept in biology, see: (Schaffner 2008; Thagard 2003).

2 Pathways in biology: an example from biochemistry. Biochemistry and molecular biology focus on the chemicals, small molecules, and macromolecules that figure in 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. These manufacturing processes are usually divided into those that are synthetic and degradative, which involve the assembly and disassembly of various materials, respectively. Biological systems synthesize materials in order to grow and replace structural components, while they degrade materials to eliminate waste and acquire energy. Scientists are interested in explaining these manufacturing processes and the pathway concept is frequently cited in their explanations. How exactly does the pathway concept figure in these explanations?

Consider glycolysis, a fundamental biological process where 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) 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 ten sequential steps that are represented along a pathway, called the glycolytic pathway, shown in figure 1. Each step of the 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.[2] The enzymes at each step in the pathway are specific in the sense that they facilitate the conversion of a specific substrate into a specific product. 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 ordinary life manufacturing contexts. 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. This involves a sequence of steps where 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 (Chiras 2015, 45, 66). Just like 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 line. This captures a kind of material continuity along the pathway.

The glycolytic pathway is considered a "model" metabolic pathway: it was the first metabolic

[2]These nine intermediates include: glucose-6-phosphate (G6P), fructose 6-phosphate (F6P), fructose 1,6-bisphosphate (F1,6BP), dihyroxydacetone-phosphate (DAP), glyceraldehyde 3-phosphate (GAP), 1,3-bisphosphoglycerate (1,3BPG), 3-phosphoglycerate, 2-phosphoglycerate, and phosphoenolpyruvate (PEP).

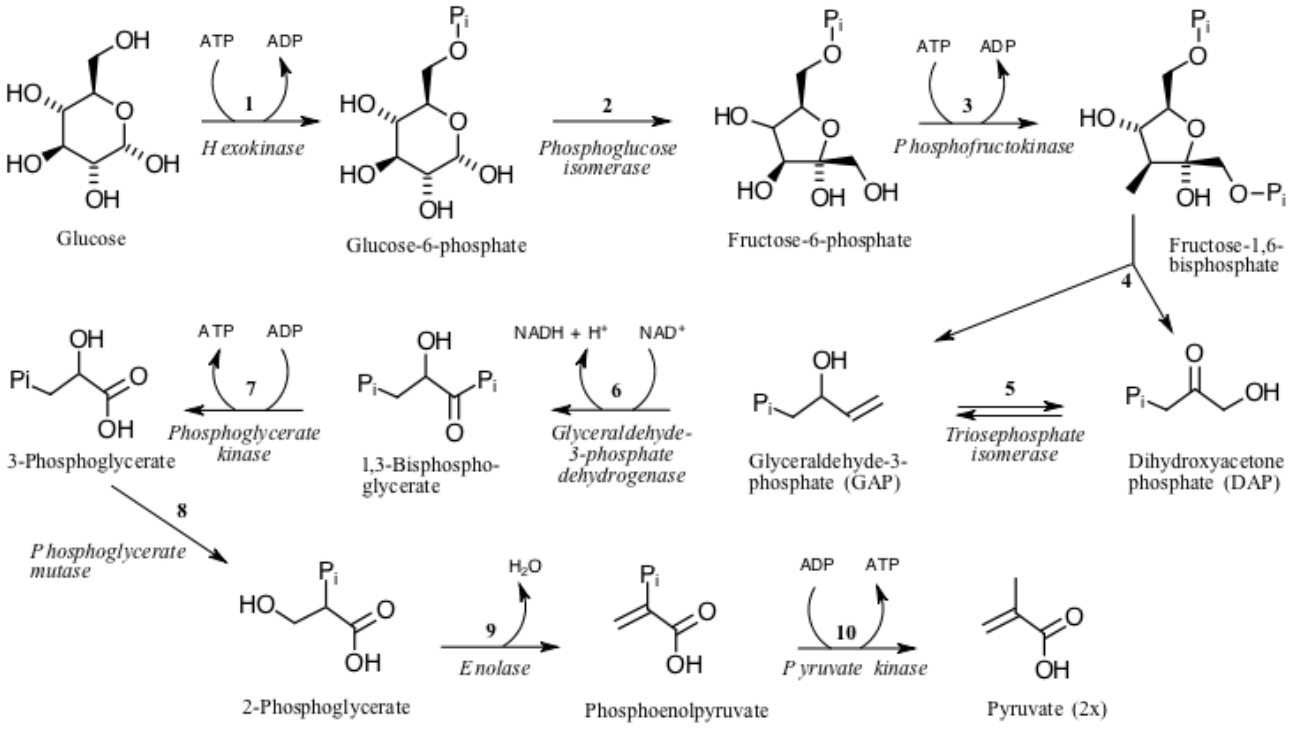

Figure 1: Metabolic pathway: glycolysis

pathway to be elucidated and it is the best understood (Nelson and Cox 2008, 528). An interesting 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 is nearly identical in all species. Thus, scientists state that "[t]he 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" (Nelson and Cox 2008, 25). For these reasons, metabolic pathways are said to capture a "unity" of life at the biochemical level (Berg, Tymoczko, and Stryer 2012, 3), (Rogers 2011, 49). In other words, "[f]rom 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 how a final product (pyruvate) is manufactured from some starting material. Thus, one main explanatory target in this case is the successful manufacture of a final product. 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. Thus, one common explanatory target in cases of metabolic manufacturing processes is this binary contrastive focus of the presence and absence of a final product.

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 motivated by Woodward's (2003) interventionist account of causation:

(i.c.) interventionist criterion: 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. Suppose 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 interventionist criterion (i.c.) relies on the notion of an ideal intervention. An ideal intervention involves an unconfounded experimental manipulation of X with respect to Y, where the changes in Y are produced by changes in X and not through any other variable. In other words, this intervention: (i) is not correlated with another variable W that has causal control over Y, (ii) it does not directly have causal control over Y, and (iii) 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. When we explain why the light is on or off we cite the light switch, because it provides causal control over this outcome.

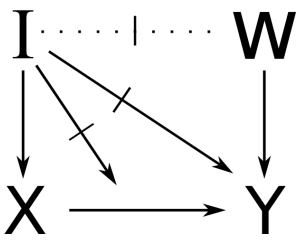

Figure 2: Causal control and ideal intervention (Woodward 2003)

3.1 Selecting causes: metabolites. In the glycolysis example there appear to be numerous factors with causal control over the manufacture of a final product. 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 that underlies this assessment. One set of factors that are cited in explanations of glycolysis are 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 (nine) metabolic intermediates. Each of these intermediates involves a relationship of direct causal control[3] where 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 (i.c.)–changing the value of glucose from present to absent provides causal control over whether G6P is formed or not. 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 substrate controls whether the immediately downstream product is formed or not. The same relationship of direct causal control holds for the second step of glycolysis. At this step G6P has causal control over F6P in the sense that manipulating this upstream metabolite, in the presence of the relevant enzyme (phosphoglucose isomerase) provides control over whether F6P is produced or not. The product of this 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, in a similar way as the diagram in figure 2. This is 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. The following 6 steps of glycolysis could be represented in the same way. If the enzymes, cofactors, and other regulatory molecules for this process were provided, 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 or not. 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 or not.

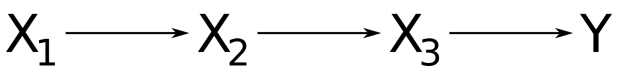

Figure 3: Sequence or chain of factors with causal control

In addition to causal control, the sequence of metabolites along the glycolytic pathway have a second important feature. These factors have a type of material continuity, where a significant amount of material from each upstream step carries over to the immediately downstream product. It makes sense that the 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 product as it undergoes this process. In fact,

[3]The notions of direct and indirect causal control bear similarities to Woodward's discussion of direct and indirect causal influence (Woodward 2003).

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 some tagged starting material and watching it flow through the pathway. By stopping the 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 tagged carbon isotopes ($^{14}CO_2$) and the tag was identified, successively, in a chain of metabolic intermediates.[4] Tracer experiments provide other types of information about metabolic pathways[5] and they are viewed as "[p]erhaps 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 need to take place in a very 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," because they identify a fixed ordering that is "required as a part of the nature of the work" and cannot be changed (Kerzner 2009, 501) (Newell 2005, 49). These differ from "discretionary dependencies" which hold between outcomes that are ordered on the basis of personal preferences and that are not fixed by the nature of the project itself.[6] 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 a ordering that is fixed by the nature of a biological system. This fixed order is sometimes referred to as a "processing chain" which 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. If the order of these steps is not followed, the process comes to a halt. While it is possible to imagine other routes through which glucose could be converted into F1,6BP,[7] the only possible 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.[8] In these contexts, the pathway concept is useful because its linear character

[4]As Israel 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' "(Zelitich 1971, 90-91).

[5]For example, they are also used to study the flow or flux of material through the pathways, which provides important information on the rate of product manufacture.

[6]For example, we may prefer to install the windows before painting the inside of the house (so as to not scratch the paint), but nothing prevents us from doing these in reverse order. In other words, painting the inside of the house does not require that the windows are already installed. In the earlier case, we are clearly prevented from roofing a house before the framing is up.

[7]For example, glucose could first be converted into G1P, then rearranged into F1P, which could be converted into F1,6BP. However, this order of steps is not found in biological systems.

[8]Engineers use arrow-on-activity (AOA) diagrams and precedence diagrams to depict the order of mandatory

captures an ordered sequence of steps. When these processes are explained, it's important that these factors are cited in a particular order that captures this fixed sequence (I discuss this more in section 4). 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 all organisms, because these factors provide causal control over this process in all of these systems.

Scientists appeal to glucose and the glycolytic intermediates in explaining glycolysis, because these factors provide causal control over this process. How exactly should we understand the rationale behind causal selection in this case? Recall that mainstream philosophical accounts view causal selection as lacking any objective rationale and as guided by "pragmatic" considerations, which are problematic influences that are arbitrary, subjective, and audience-relative. This analysis suggests that causal reasoning in this case is "pragmatic" in a different and non-denigratory sense. In the case of glycolysis, causal selection is pragmatic in the sense that it is relative to a practical or useful goal–the goal of control. Identifying factors with causal control is useful in the sense that it provides a way to potentially change the outcome of interest. Once this goal is specified, there are objective and empirical facts about what means conduce to it. For example, once an explanatory target is specified their are objective facts about what factors provide control over it and which ones do not. This isn't decided by our personal or subjective presences, but by the nature of these systems. In the following sections, I further examine the role of pragmatic and objective considerations in causal selection.

4 The second step: clarifications. Providing an explanation of glycolysis involves more than just appealing to a sequence of metabolites. These factors alone don't 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 (if any) 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.[9] This can be clarified with an ordinary life example where we want to explain what causes a flashlight to turn on and off (Y). 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 of interest. 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

dependency relations that characterize various manufacturing processes(Newell 2005, 49-51).

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

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 (viz. the on state), gaining complete causal control over this target requires both factors.

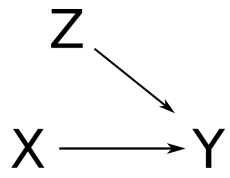

Figure 4: Directed acyclic graph (DAG) of flashlight bulb (Y), battery (X), and switch (Z).

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. 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.[10] Likewise, when the substrate is present changing whether the enzyme is available or not, controls whether or not 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.

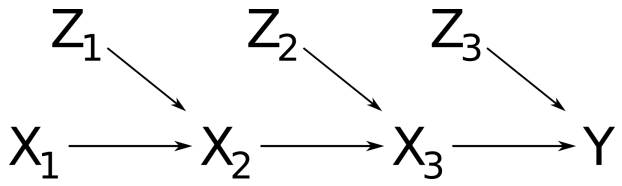

Figure 5: Directed acyclic graph (DAG) of metabolites (X_1, X_2, X_3, X_4 and enzymes (Z_1, Z_2, Z_3)

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 non-functional, this blocks a step in the pathway, resulting in a build-up of upstream material. The build up of a particular types of material can be used to identify the metabolites along the pathway and the enzymes they are associated with.[11] In fact, the initial discovery of many metabolic

[10] These interacting causes also have uneven control like the flashlight example–either factor alone can prevent product formation, both are required to allow for product formation. As both these values of the outcome are the focus of the explanatory target, both factors are required to gain complete causal control over it.

[11] These disorders have a similar pathology (excessive accumulation of material) and similar treatment strategies, including: (i) reducing the inflow of substrate, (ii) re-routing the flow of material, or (ii) overcoming the blockage by supplying a functional version of the defective enzyme.

disorders has involved identifying a pathological accumulation of metabolic materials.[12] Scientists compare these disorders to disrupted factory assembly lines. They state that non-functional 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.[13] This experimental strategy was used in groundbreaking 20th century research that helped establish the fields of molecular biology and biochemistry. For example, these experimental methods were used in Beadle and Tatum's noble prize work on metabolic pathways in *Neurospora* (Beadle and Tatum 1941) and Ephrussi and Beadle's research on biosynthetic eye color pathways in *Drosophila* (Beadle 1937; Beadle and Ephrussi 1937). These studies led to the one gene-one enzyme hypothesis[14] and provided one of the first and most convincing explanations for how genes influence phenotypes, viz. through enzymes, which regulate metabolic pathways. Hull's 1974 diagram in figure 6 depicts these relationships between genes, enzymes, and metabolites and bears some similarities to the causal graph in figure 5 (Hull 1974). (I discuss the role of genes in these explanations in section 5.) In these experiments metabolic pathways are studied with strategies that exploit: the causal control of enzymes, the location of enzymes and substrates in a pathway, and the flow of material along the pathway.

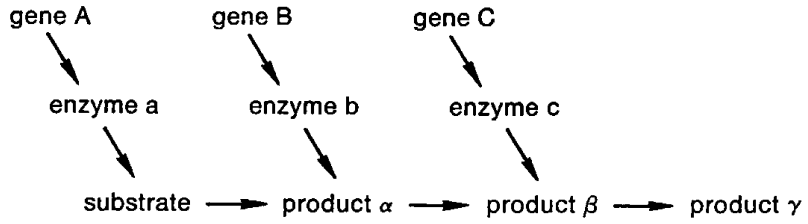

Figure 6: Metabolic pathway (Hull 1974, p. 29)

Metabolites and enzymes receive the most attention in explanations of metabolic processes. However, other factor do frequently figure in these explanations, for example enzyme cofactors and accessory substrates (e.g. ATP, NADPH, etc.). The interacting cause framework provides a promising approach to include many of these other factors–in many cases, they are an additional interacting causes at a given step which allow for causal control over product formation.

4.2 General speed. I have used the notion of interacting causes to suggest that enzymes and substrates are selected as causes of metabolic processes. This is because intervening on both of these factors is required to gain causal control over the manufacture of a final product. There is a puzzle

[12]Examples of metabolic diseases discovered in this way are alcaptonuria, phenylketonuria, and maple syrup urine disease.

[13]This is related to Waters' discussion of the "genetic approach" (Waters 2004).

[14]This hypothesis was suggested in (Beadle and Tatum 1941) and maintained that single enzymes were produced by single genes. Although the hypothesis is now considered an oversimplification, it is viewed as the "first major discovery of molecular biology" (Morange 1998, 21).

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 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 time-scale set by the context of inquiry. If a factor fails to meet this standard, it won't be selected because its control doesn't show up at the time scale of interest and, thus, is of little (to no) use in controlling and explaining the outcome at this scale. In the case of metabolic processes, the relevant time-scale is milliseconds to minutes. Metabolic processes need to operate at this time scale in order for living organisms to sustain life and functionality. Organisms need to react to their environments, break down materials, and grow (or develop) 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 time-scale 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 metabolites alone. It isn't just that both are required to gain causal control over pyruvate manufacture, but that their causal control meets this general speed principle.[15]

Another example helps to support these points and suggests their more general role in causal reasoning. Consider the conversion of diamond to graphite, which is another spontaneous process.

[15]Thus, 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).

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...[t]he rate of conversion...is so slow, that *for all practical purposes, it does not occur...*" (Woodbury 2012, 10). This reaction is viewed as "*negligible* at standard conditions"(Jespersen 2017) because "it cannot be observed over any *practical* period of time" (OpenStax 2016). We tend to deny that diamond alone decomposes into graphite, because this decomposition doesn't occur on a time-scale relevant to human life, which is our default time-scale in these situations. The only way to see this decomposition at this time-scale is to subject diamond to extremely high heat or high pressure. When we want to explain this process on a time-scale 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 time-scale–e.g. one at the geological or cosmological level–we would be more likely to just cite just diamond as the cause of this conversion, because its causal control does manifest at these larger time-frames. 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 an 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 may be an "in principle" 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 isn't useful to us because it doesn't occur within a "practical period of time," given various constraints (like the length of our lifespans). This is not a "merely 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 in the world. Causal selection is relative to a practical goal, where causes need to meet some general time-frame specified by the context of interest. One this time-frame is specified, there are objective facts about what factors meet it and which factors don't. Again, this isn't dictated by "merely pragmatic" preferences. Our preferences don't determine how fast diamond converts into graphite, or glucose into pyruvate. The speed of causal control is determined by objective facts about the nature of these systems.

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 some have causal control, it needs to meet the interventionist criterion (i.c.) 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.[16] In the case of glycolysis, these steps help clarify why

[16]Exploring this suggestion in detail requires a further paper, but it is supported by the applicability of the interventionist criterion (i.c.) and the general speed principle to causal selection in non-biological cases like light switches, flashlights, diamond decomposition, and factory assembly lines.

metabolites and enzymes are selected as causes. However, there are further reasons for why these factors are selected as causes, 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 three, 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 answered 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 make-up of a product can completely change its functional capacity or render it non-functional in ways that lead to deleterious outcomes and even death. Minor changes in the stereochemical structure of amino acids and the amino acid sequence of proteins are associated with various diseases.[17] For these reasons it matters that a very particular product is produced and scientists want to explain what causes the manufacture of one particular 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, 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. Both of these notions are suggested by scientists' claims that "[e]nzymes are highly specific both in the reactions they catalyze and in their choice of reactants which are called substrates" (Berg, Tymoczko, and Stryer 2012, 220). I'll 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, viz. 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 non-sugars. Out of all these types of molecules, it only interacts with one specific type. On the other hand, (b) has to do with specificity in the sense of producing a particular outcome. Enzymes exhibit specificity 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,

[17] A minor change in one amino acid of hemoglobin results in the blood disorder sickle cell anemia. Changes in the particular geometrical configuration of amino acids are associated with conditions like chronic kidney disease.

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.[18] 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.

Enzymes are selected in explanations of metabolic manufacturing processes, not just because they have causal control over product formation (along with substrates), 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 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.

Notice that specific causal control is *attributed* to enzymes and *seen* in the specific metabolic intermediates along the pathway. It will help to discuss this briefly, before moving on. 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–the sequence of specific glycolytic intermediates is nearly universal across all life forms. 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. On the other hand, the enzymes in these pathways are not universal in this sense and not represented in this 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, where each enzyme in the family catalyzes the same chemical reaction. The enzymes in these families are called isozymes or isoforms and they are "different proteins that catalyze the same reaction" (Nelson and Cox 2008, 583). The criteria 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 is a particular system (e.g. a tissue, cell, etc.) the enzyme variable represents a particular enzyme from the family. 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.

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

[18]This notion of specific causal control relates to a form of one-to-one specificity discussed by (Woodward 2010).

extremely fast nerve signaling, which is facilitated by the speed of neurotransmitter degradation.[19] 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 destroyed before it could figure in transcription or other activities (Berg, Tymoczko, and Stryer 2012, a9). For these reasons the particular rate of production of a biological manufacturing process 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.[20] 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. These enzymes are singled out as important because they explain the rate of manufacture: the speed of their causal control explains the speed at which the final product is made.

Pacemaker enzymes are also called "bottleneck enzymes," because their slower speed results in congestion along the pathway, as upstream material accumulates (El-Mansi and Stephanopoulos 2007). Bottlenecks are often considered problematic, because their deceleration of the manufacturing process (relative to the speed of other enzymes in the pathway) is viewed as inefficient and wasteful. In this context, efficiency is valued and associated with faster manufacturing rates, while inefficiency is less valued and associated with slower manufacturing rates. The preference for faster production and the association of this with efficiency is seen in scientists' discussion of the catalytic speed of enzymes. Consider two dimensions along which the catalytic speed of an enzyme is assessed: the enzyme's (a) turnover rate, or number of catalytic reactions per unit time, and (b) rate enhancement, or how much faster a reaction is with the enzyme than without it. Consider how scientists refer to enzymes that score higher along these dimensions in the sense of having faster turnover rates and larger rate enhancements. They refer to these enzymes as "perfect enzymes," that have achieved "catalytic prowess," and are "proficient," "efficient," and "highly evolved."[21] On the other hand, enzymes that score lower along these dimensions are viewed as "wasteful," "surprisingly inefficient," "painfully slow," representative of "unintelligent design," and lead to processes being "hamstrung by slow catalysis."[22] This preference for faster product manufacture can be understood in the following way. If a biological system needs to make (or degrade) a par-

[19]The speed at which enzymes breakdown neurotransmitter (e.g. enzymes like acytelcholinesterase) is viewed as "physiologically crucial" for proper nerve signaling (Berg, Tymoczko, and Stryer 2012, 398).

[20]The pacemaker enzyme of glycolysis is phosphofructokinase (PFK), which catalyzes the third step (Berg, Tymoczko, and Stryer 2012, 474).

[21]Examples of these enzymes include acetylcholinesterase, which has a turnover rate of 25,000 reactions per second and carbonic anhydrase, which has a turn over rate of 1 million reactions per second (Berg, Tymoczko, and Stryer 2012, 398, 285).

[22](Parry et al. 2013, 717; Suboski 2011; Ellis 200, 164; Tcherkez et al. 2006, 7246.)

ticular material, it will typically be advantageous for this to happen as quickly as possible. There seem to be fewer potential disadvantages with producing a product "too" quickly, and many more with producing the product "too" slowly.

My analysis has focused on why metabolites and enzymes are selected as causes of glycolysis. What about genes? 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, on the other hand, have "fine control" over these processes, because they have "immediate effects" on product manufacture (Rogers 2011, 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 genes produce enzymes at slow rate that doesn't capture or explain the speed of metabolic processes. More specifically, its assumed that genes have already manufactured 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.[23] 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.

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. This paper has provided an analysis of the principled rationale that guides causal selection in biochemistry, particularly, in cases where scientists explain metabolic processes. Contrary to mainstream philosophical views, factors are not selected as causes on the basis of considerations that are arbitrary, subjective, and audience-relative. In this domain, causes are selected on the basis of principled considerations that have to do with causal control. Identifying factors with causal control over an explanatory target is useful–it provides a means of controlling, changing, and regulating an outcome of interest. This suggests that causal selection is pragmatic in the sense that it is relative to a practical goal–the goal of control. Once this goal is specified, there are objective facts about which factors provide this control and which do not.

A special feature of this analysis is that it clarifies how selected factors are organized with the pathway concept and how this concept plays a important role in biological explanation. This

[23]One way to increase the rate of product manufacture is by increasing enzyme concentration. Biological systems exploit this strategy to overcome slow manufacturing rates caused by enzymes with slow catalytic speeds. An example of this is the enzyme Rubisco, which is an extremely slow photosynthetic enzyme (Suboski, 2011). Plants work to overcome this slow enzyme by producing it in large quantities. In fact, they produce so much of this enzyme that it is the most abundant protein in the biosphere (Berg, Tymoczko, and Stryer 2012, 591).

basic explanatory pattern is not just found in biochemistry and ordinary life examples of factory assembly lines–it is present all throughout biology. When scientists explain biological processes like cell division (e.g. meiosis and mitosis), cellular differentiation (e.g. stem cell differentiation), and organism development (e.g. embryological stages) they use the pathway concept to represent the sequence of changes that these material substances undergo as they develop into some final, mature system.[24] Similar to metabolic pathways, these cases identify a sequence of factors with (i) causal control, (ii) material continuity, and a (iii) fixed order that represents a (iv) manufacturing process, which is (v) universal or shared across some large domain of systems. Further analysis of these examples may reveal whether they are guided by a similar rationale as is found in the case of metabolic pathways.

[24]Scientists also compare these biological processes to factory assembly lines, just as they do with metabolic processes (Ferrell, Tsai, and Yang 2011) (Bernard 2012, 33).

References

Beadle, G. W. (1937). The development of eye colors in Drosophila as studied by transplantation. *The American Naturalist 71*, 1–8.

Beadle, G. W. and B. Ephrussi (1937). Development of eye colors in drosophila: Transplantation experiments on the interaction of vermillion with other eye colors. *Genetics*, 65–75.

Beadle, G. W. and E. L. Tatum (1941). Genetic control of biochemical reactions in Neurospora. *Proceedings of the National Academy of Sciences 27*, 1–8.

Bell, G. (2008). *Selection: The mechanism of evolution* (2 ed.). Oxford University Press.

Berg, J. M., J. L. Tymoczko, and L. Stryer (2012). *Biochemistry* (7 ed.). W. H. Freeman and Comany.

Bernard, J. (2012). *The eukaryote genome in development and evolution*. B. John and G. Miklos.

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 the United States of America 98*, 4249–4254.

Chiras, D. D. (2015). *Human biology* (8 ed.). Jones & Bartlett Learning.

El-Mansi, E. M. T. and G. Stephanopoulos (2007). Flux control analysis: Basic principles and industrial applications. In *Fermentation Microbiology and Biotechnology*. Taylor and Francis Group.

Ellis, R. J. (2010). Biochemistry: Tackling unintelligent design. *Nature*.

Ferrell, Jr, J. E., T. Y.-C. Tsai, and Q. Yang (2011). Modeling the cell cycle:Why do certain circuits oscillate? *Cell 144*(6), 874–885.

Hull, D. L. (1974). *Philosophy of biological science*. Prentice Hall.

Jespersen, N. D. (2017). *Chemistry: The molecular nature of matter* (7 ed.). Content Technologies.

Kerzner, H. (2009). Project management. *John Wiley & Sons, Inc.*, 1–1121.

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

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

Morange, M. (1998). *A History of Molecular Biology*. Harvard University Press.

Nelson, D. L. and M. M. Cox (2008). *Principles of biochemistry* (5 ed.). W.H. Freeman and Company.

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

OpenStax (2016). *Chemistry*. OpenStax.

Parry, M. A. H., P. J. Andralojc, J. C. Scales, M. E. Salvucci, A. E. Carmo-Silva, H. Alonso, and S. M. Whitney (2013). Rubisco activity and regulation as targets for crop improvement. *Journal of Experimental Botany*, 1–14.

Rogers, K. (2011). *The chemical reactions of life*. Britannica Educational Publishing.

Schaffer, J. (2016). Metaphysics of causation.

Schaffner, K. F. (2008). Etiological models in psychiatry. In K. S. Kendler and J. Parnas (Eds.), *Philosophical Issues in Psychiatry: Explanation, Phenomenology, and Nosology*, pp. 1–49. Johns Hopkins University.

Spirtes, P., C. Glymour, and R. Scheines (2000). *Causation, prediction, and search* (2 ed.). Massachusetts Institute of Technology.

Suboski, W. (2011). The day of the kudzu. *Planet Magazine*, 1–4.

Tcherkez, G. G. B., G. D. Farquhar, and T. J. Andrews (2006). Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized. *Proceedings of the National Academy of Sciences of the United States of America 103*(19), 7246–7251.

Thagard, P. (2003). Pathways to biomedical discovery. *Philosophy of Science*, 1–20.

Waters, C. K. (2004). What was classical genetics? *Studies in History and Philosophy of Science*.

Windelspecht, M. (2007). *Genetics 101*. Greenwood Press.

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

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

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

Woodward, J. (2016). Causation and Manipulability. *Stanford Encyclopedia of Philosophy*.

Yu, O. and J. M. Jez (2008). Nature's assembly line: biosynthesis of simple phenylpropanoids and polyketides. *The Plant Journal 54*(4), 750–762.

Zelitich, I. (1971). *Photosynthesis, photorespiration, and plant productivity*. Academic Press.