Explanation in two dimensions: diagrams and biological explanation

LAURA PERINI
Philosophy Department (0126), Virginia Tech, Blacksburg, VA 24061, USA (e-mail: lperini@vt.edu; phone: 540-231-8489; fax: 540-231-6367)

Abstract. Molecular biologists and biochemists often use diagrams to present hypotheses. Analysis of diagrams shows that their content can be expressed with linguistic representations. Why do biologists use visual representations instead? One reason is simple comprehensibility: some diagrams present information which is readily understood from the diagram format, but which would not be comprehensible if the same information was expressed linguistically. But often diagrams are used even when concise, comprehensible linguistic alternatives are available. I explain this phenomenon by showing why diagrammatic representation is especially well suited for a particular kind of explanation common in molecular biology and biochemistry: namely, functional analysis, in which a capacity of the system is explained in terms of capacities of its component parts.

Introduction

In 1953 James Watson and Francis Crick made history with a diagram of a double helix: their proposal for the structure of DNA (Watson and Crick 1953). This is not the only case in which an important scientific hypothesis is expressed with a figure rather than a linguistic representation. Many biological models were introduced with visual representations and continue to be presented in diagrammatic form, such as the Krebs (citric acid) cycle, and the chemiosmotic hypothesis. These models are all prized for their explanatory value. Why do scientists use diagrams to present them? Could a diagram play a genuine role in scientific explanation? If so, how does the visual format contribute to such a role?

Philosophy of science has few resources to address these questions. The literature on scientific explanation focuses on linguistic forms of representation. The most famous model of scientific explanation, the deductive-nomological (DN) model, conceives of explanation as an argument consisting of a universal generalization and statements of initial conditions, from which a statement of the phenomenon to be explained is deduced. This account of scientific explanation has been critiqued extensively; it does not fit all cases of genuine scientific explanation, and there are explanatorily irrelevant arguments that meet all the DN criteria for scientific explanation. But the problem with the DN model goes beyond mere failure to capture the particular logical structure of scientific

explanation. The tactic of looking for an account of scientific explanation in terms of logical relations among scientific claims no longer appears promising. Contemporary theories suggest several alternatives: explanatory relevance is determined by pragmatic considerations, or by causal relations, or patterns of reasoning (for respective examples, see van Fraassen 1980; Salmon 1984; Kitcher 1981). The search for a *general* theory of scientific explanation has not produced a consensus, however. An alternative approach to scientific explanation is based on investigation of *types* of explanations, such as the developing literature on mechanisms (for example, Machamer et al. 2000). The tactic of studying a type of explanation, rather than aiming to account for all scientific explanation, also allows for discipline-specific investigations of explanatory types. Philosophers of biology, for example, have had an extensive discussion of functions and the nature of their explanatory significance.

While some studies investigating types of scientific explanations have reproduced scientific diagrams used to convey explanatory hypotheses, these studies have not analyzed those diagrams as visual representations as part of their investigation.[1] The legacy of the DN model, and the focus on linguistic representations in general, is that philosophers have studied scientific explanation without thinking about the contribution of visual representations – as if guided by the assumption that only linguistic (or mathematical) representations *could* be relevant to scientific explanation. Such an assumption, of course, would be fatal to a study of *whether* diagrams can play a role in scientific explanation.

Avoiding the influence of assumptions about the nature of linguistic and visual representations and the potential contribution of diagrams to scientific explanation requires an understanding of diagrams as visual representations. Below I identify the key representational features of diagrams and use them to compare diagrams with linguistic representations. The results are then used to explain the advantages of using diagrams to present explanatory content. I will show that there are at least two reasons biologists might use diagrams to express hypotheses. One has to do with human cognition: diagrammatic representation can make extremely complex information comprehensible by humans, who could not utilize the same information presented in linguistic form. The second reason has to do with the distinctive explanatory role these models play, and how the visual format contributes to this explanatory role.

Representational features of Diagrams

Diagrams, like photographs, printed sentences, uttered sentences, and numerical formulae are symbols: physical things interpreted to refer to

---

[1]Some, such as Machamer et al. (2000), even provide some description of what *parts* of the diagram represent (e.g. arrows refer to activities), but a more fundamental understanding of *how* diagrams convey content is needed in order to understand their role in scientific explanation.

something else.[2] There are two important features that distinguish diagrams as symbols. First of all, they are visual representations. Biologists use several types of visual representations, including graphs, electron micrographs, photographs and drawings as well as diagrams. There are significant differences among visual representations, but they all share one important feature: some spatial relations among parts of the symbol refer to some features of the fact represented. Spatial relations in a molecular diagram refer to spatial relations among atoms; spatial relations in a graph refer to relations among properties, etc. The significance of spatial properties of the visible symbol is the defining feature of visual representations, and it is what differentiates them all from linguistic representations. Text and mathematical formulas are *serial* representations, in which the *sequence* of characters determines the meaning of the representation, but the spatial relations within the symbol (such as distance between two words in a printed sentence) are not meaningful.[3]

The distinction between serial representations and visual representations (in terms of sequential versus two-dimensional format) does not capture all the relevant features of various types of representations. This distinction does not account for important differences among visual representations: diagram systems have syntactic features that other visual representations – those that look more like pictures, like electron micrographs – do not. More importantly for the topic at hand, the visual-serial distinction does not account for an important *similarity* between diagrams and textual representations. Diagrams have a different kind of format from text (spatial vs. serial), but diagrams also share syntactic features with textual and numerical representations. These features are important to understanding why diagrams would be used to present scientific models.

Like numerical and textual representations, diagrams are composed of articulate characters (Goodman 1976).[4] This means that for any particular diagram in a diagram system, each of its component characters can be identified, and the meaning of the whole symbol is determined by the identity and arrangement of the atomic characters forming the complex character. Text has

---

[2]'Symbol' is often used in a much more narrow sense, to refer to representation-bearers like sentences and numerical formulae, but investigation of the similarities and differences among different kinds of representations like text and diagrams requires a more general approach in order to avoid merely presupposing differences. Goodman's (1976) study of symbol systems establishes a precedent for the broader use of the term.

[3]Letters are identified by their two-dimensional shapes, but those shapes, and thus the forms of the words and sentences they comprise, are not interpreted to mean something about the referent. Serial forms of representation can use spatial features to add meaning to text, for example by using extra spaces between words in a poem to convey a sense of hesitancy. But such use of space is not *necessary* for serial representations to be meaningful, in contrast to the fundamental referential use of spatial relations by visual representations.

[4]Diagrams are just one type of visual representation. There is another broad category of visual representations (pictorial representations) which have a different syntactic character from diagrams (for example, electron micrographs are pictorial). These visual representations cannot be decomposed into discrete syntactic elements.

a similar syntactic character: letters and punctuation symbols can be identified, and the meaning of a sentence is a function of the series of letters, spaces and punctuation marks. The difference between textual representation and diagrammatic representation is that the meaning of text is determined by the serial arrangement of atomic characters, and the meaning of diagrams is determined by the two-dimensional array of atomic characters.

The articulate syntax of diagram systems allows for the complete dissection of each diagrammatic representation into its component parts. In addition to this type of syntax, the semantics of diagram systems assign referents to these unambiguously identifiable characters. As a result the meaning of a diagram is a function of (1) the meaning of the atomic characters and (2) how they are arranged (Perini 2002). A diagram can thus be given a linguistic translation as long as the language contains terms for what the atomic characters refer to, and terms for the relations the spatial relations among the atomic characters refer to. So models given a diagrammatic representation could be expressed with a linguistic symbol system.[5]

This makes the question of why biologists use visual representations to convey important ideas even more pressing – why not use a linguistic expression of the same content instead? In the next section of the paper I will discuss two models from biochemistry and the diagrams used to convey them. Clarification of their explanatory roles and how diagrammatic representation contributes to those roles will ground an explanation of the use of diagrams even when a linguistic expression of the same explanatory content is available.

Examples: models of F_1 ATPase function and structure

ATP (adenosine triphosphate) is the immediate energy source for many cellular processes. ATP is produced by a multi-step process called oxidative phosphorylation that occurs within mitochondria. Electrons are extracted from carbon compounds and transferred to proteins associated with the mitochondrial membranes. The successive transfer of electrons results in a chemical and electrical gradient across mitochondrial membranes, which, in turn, is used as an energy source to make ATP out of precursors ADP(adenosine diphosphate) and Pi (phosphate). This last step is catalyzed by the phosphorylation, F_1 ATPase complex. In the early 1960's scientists discovered that F_1ATPase catalyzes ATP formation, but it took several decades of research to produce a consensus on *how* the enzyme catalyzes the production of ATP. Two models are now central to understanding how the F_1 ATPase works. The first is a

---

[5]Since any diagram has a finite number of atomic characters, it is possible to introduce linguistic terms into the language that refer to the referents of the atomic characters (if the language does not already include such terms). This expanded language can then express the content of the diagram.

model which describes the mechanism of the ATP-producing reaction. The second is a structural model of the F_1 ATPase complex. The structural model represents features of the complex essential to the process described by the mechanistic model. Together these models explain how the complex catalyzes synthesis of ATP.

I'll discuss the model of the reaction mechanism first. Paul Boyer's group at UCLA developed the binding-change model of the mechanism by which F_1 catalyzes ATP production. One key question the model addressed has to do with how energy is involved in the process of ATP formation. Production of ATP (from precursors ADP and Pi) cannot occur without an input of energy. Enzymes function in chemical reactions by lowering the activation energy – the amount of energy needed to destabilize the starting materials to the point where they can react and form products. Enzymes lower this energetic barrier, but they do not provide a net energy input to the reactions. Some enzymes, however, do link reactions to an energy source (e.g. with another, energy-providing reaction) thus making it possible for an energy-requiring chemical reaction to occur. Since it was known that ATP synthesis requires a net energy input, an explanation of F_1's role in ATP production would have to include not only a claim about its capacity to lower the activation energy of the reaction, but also an explanation of how energy is contributed to the phosphorylation reaction.

The binding change model of the enzyme is that the complex has three chemically identical catalytic subunits, each of which is in a different one of three distinct conformations. Each subunit cycles through the same sequence of states, changing from the loose conformation, to tight, to open. Each state has a different affinity for reactants and products. Transitions between states thus cause binding/release of components. The loose conformation has a moderate affinity for precursors ADP and Pi, and a low affinity for ATP; it tends to bind the precursors if any are available. The tight conformation binds both ADP and ATP so tightly that neither is released. When bound in the tight conformation, ADP and Pi react to form ATP without requiring energy input. However, energy is required to change the ATP-containing subunit from the tight to the open conformation. All three subunits change conformation at once, or none do. So when energy is available, the tight subunit shifts to the open conformation, which has such a low affinity for ATP that it is released. The precursor-binding subunit in the loose conformation moves into the tight conformation (in this state ATP forms from the precursors), and the open, empty subunit changes to the loose conformation, in which state it can bind precursors ADP and Pi. Without energy input the tight subunit will not shift to the open conformation and release ATP, and the other two subunits will be locked into their respective conformations. Although the loose subunit may have ADP and Pi bound they will not react to form ATP when complexed with the loosely conformed subunit. So without energy input, ATP production stalls even when the other starting materials are present.
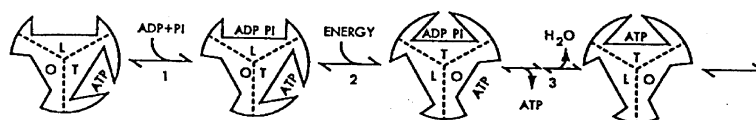
Figure 1. The binding-change mechanism (Abrahams et al. 1994). Reprinted with permission from Nature.

The binding-change mechanism is represented by the diagram in Figure 1.[6] This figure represents the role of energy in ATP production, and the number and sequential activity of the catalytic sites. The diagram shows clearly how these three basic features of the mechanism work together. The diagram accomplishes all this through the use of a small set of atomic characters. There are three arced shapes which refer to subunits in the loose, tight, and open conformations. Horizontal arrows refer to transitions between states. The names of chemical inputs to the reaction have their usual referents, and their position at the end of curved arrows means that they are added or removed from the complex at that point in the process. Note that this figure is not representing a rotation of the complex in space. The top subunit is the same throughout the figure, it is just shifting into a different conformation.

The binding-change mechanism involves many complex relationships. The figure is an effective way to communicate the model because it does more than represent multiple features. The atomic characters used to represent different components of the model are utilized in a system whose semantics assign relations *among* different components based on the spatial arrangement of the atomic characters. So the contiguous arced shapes together refer to a three-subunit complex in a particular combination of subunit states, and the position of chemical terms in the hollows of the arced shapes indicate that the chemical referent of the term is bound to that subunit. When connected by double arrows, these symbols for the complex refer to sequential changes between one combination of subunit-states and another. The relations between the conformation state of the complex, energy input, and ATP formation are clearly expressed.

While the binding-change model provides a general mechanism for ATP catalysis, it does not convey any specific information about how the F_1 complex goes through the proposed transitions. Such information would have to come from a model of the enzyme's structure. In 1994 Walker's group published a detailed and fairly precise (to 2.8 Å) model of the structure of F_1 ATPase (Abrahams et al. 1994). This model is the result of a study in which X-ray

---

[6]Different *kinds* of diagrams were used to represent this model throughout its development. This is a distinct phenomena from using one kind of diagram – one symbol system – to represent different models. The diagram system used to express the binding-change hypothesis could be used to express alternative hypotheses: for example, a figure with L–L–L subunits in the first stage, followed by a T–T–T complex, would be properly interpreted (for this symbol system) as claiming that a three-subunit complex with all subunits in the loose conformation shifts into a complex with all subunits in the tight conformation.

diffraction data collected from crystallized F_1 were analyzed to determine the spatial structure of the protein. In such a study, analysis begins from a digitized representation of the diffraction pattern digitized, and proceeds mostly in the form of operations on mathematical representations, and, if successful, yields coordinates which represent the location of individual atoms of the enzyme.

Such studies are highly valued because they yield the structural information necessary to understand how molecules carry out their functions. Proteins consist of chains of amino acids. The side groups of the amino acids vary, and positioning several amino acids close together will put their side groups into close proximity. This could result in a functionally critical feature of the molecule; for example, a patch of positively charged side groups on the surface of a protein will tend to bind a negatively charged molecule like a phosphate ion. What makes structural models so important is the fact that amino acid chains do not simply stretch out in a line. They wind around in complicated formations. This means that side groups on amino acids that are very far from one another in terms of position on the amino acid chain might be located right next to one another in the protein.

Figure 2 is a representation of the three-dimensional structure of the catalytic subunit of the F_1 ATPase. The straight lines are atomic characters that refer to chemical bonds and atoms between the central carbon of amino acids adjacent in the linear sequence. Spatial relations among these atomic characters refer to spatial relations among amino acids. So the diagram as a whole represents the arrangement of amino acids in the subunit. The two-dimensional arrangement of atomic characters on the page does not have the same spatial form as the three-dimensional molecule, but because we interpret the figure as a representation of a three-dimensional object, visual perception of the figure allows humans to comprehend the shape of the molecule. These diagrams are presented in pairs, because when viewed in a stereoscope, or when "free fused",
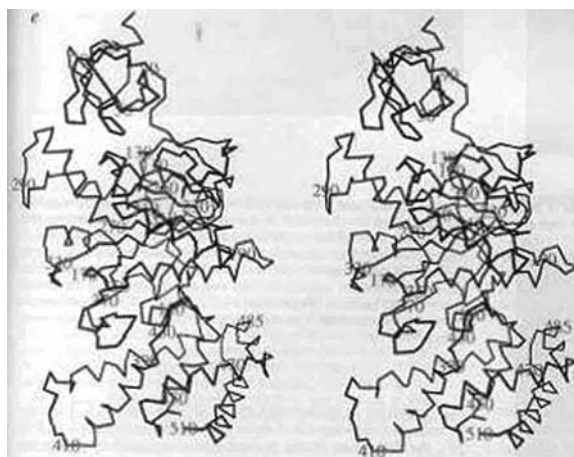

Figure 2. Structural diagram (Abrahams et al. 1994). Reprinted with permission from Nature.

they appear as a three-dimensional image of the structure in which the perception of its parts does correspond to their relative spatial positions.

While the two F_1 models have important differences, both are examples of a type of hypothesis which has been studied by philosophers of biology *because* of its recognized explanatory import. They are examples of the type of functional explanation described by Cummins (1975).[7] Cummins argues that a functional explanation is an analytic account in which some capacity of a system is explained by the capacities of components of the system. This type of explanation is common in the biological sciences. In the binding-change model, the capacity of the complex to produce ATP is explained by the capacities of the subunits to bind precursors, the capacity of the complex as a whole to respond to energy input by changing the conformations of all the subunits, and the different ligand binding capacities of the various subunit conformation states.

Cummins' account of functional analysis stresses the relation between capacities of system components and the capacity of the system as a whole. It does not stress the importance of relations *among* the features of components, which are often critical to explanations in molecular biology and biochemistry. As Bechtel and Richardson (1992) show, hypotheses that made important explanatory contributions to biochemistry, such as Kreb's model of the citric acid cycle and Mitchell's chemiosmotic hypothesis, involve relations among the components as well as features of the components themselves.[8] The binding change mechanism is complicated due to the functional interdependencies among parts of the enzyme complex: catalysis depends on the coordinated transition of each subunit from one conformation to the next. The model is explanatory because it shows how the capacities of and interrelations among components of the complex cause ATP synthesis.

The binding-change model is abstract in the sense that only a limited number of possible features of the reaction are included.[9] The binding-change model includes some very general structural claims: changes in subunit capacities to bind ATP, ADP and Pi are explained by appeal to conformational differences among chemically identical subunits. But it does not provide an answer to more specific questions. What are the conformations identified functionally as

---

[7]There is a continuing dispute about *which* concept of function is correct: Cummins' proposal, or one that defines the function of a trait in terms of its selection history. Biologists do explain the capacities of systems in terms of the capacities and relations among component parts. Whether or not this type of explanation deserves the label of a functional explanation does not matter for the analysis presented here. My conclusions in this paper do, however, depend on the claim that Cummins-style accounts are genuinely explanatory.

[8]Bechtel and Richardson (1993) identify this as a research strategy – called decomposition – that involves explaining the capacity in terms of the features of the entity's component parts (explaining *how* it carries out the capacity).

[9]This abstraction is reflected in the diagram of the mechanism: that does not represent the *shapes* of the subunits for example, it merely conveys the information that the complex consists of subunits in different conformational states.

loose, tight, and open? How is the structure of their binding sites related to their different affinities for the reaction components? How do the three subunits of the complex coordinate their transitions with energy input? These are questions that would have to be answered by the chemical structure of the enzyme complex.

The model of F_1's structure provides information about the structure of the subunits and their relations to one another. This is another example of a Cummins/Bechtel and Richardson type of explanation. The structure of a subunit includes both its component parts (amino acids) and their relative spatial locations. The paper in which this structure is presented includes discussion of structural features of enzyme subunits, and assessment of how well they account for functional features of F_1. For example, a region of a subunit that has the right shape and chemical character explains the capacity of the complex to bind ADP. The identity and relative locations of the amino acids determine the structural character of the subunit; its binding capacity is a result of those features. So the structural model is explanatory because it shows how the identity and spatial relations among the amino acids along the protein account for its capacities to bind the reactants and/or products of the reaction. The model provides the same type of explanation as the diagram of the binding-change mechanism: it provides a functional analysis of a system.

The visual format and explanatory content

At this point we can tackle the question of why scientists use diagrams to convey these models. In both the examples discussed, the models were explanatory because they provide a complex functional analysis of the phenomenon in question. The binding-change model is a composite of the relationships between interacting subunits and their different capacities for ligand binding. The mechanism of ATP production includes properties that arise out of the organization of the component parts of the F_1 complex; the mechanism is not simply a result of individual capacities of the parts of the complex. Similarly, the structural model is explanatory because the spatial relations among different atoms in the complex account for the capacities of the complex. In both cases then, relations among the component parts are key to the model. But the explanatory benefit of functional analysis is not simply an assertion about the existence of higher order relations relevant to the phenomenon of interest. The explanation is not composed of just the pattern of relations among components; rather, the identity of the components in those relations is part of the explanation. The goal of the functional analysis is to understand a capacity of the whole in terms of capacities of its component parts. So the identity (and some features) of those components are integral to the functional explanation. In complex functional analysis, the relations holding among components as well as some capacities of the components themselves are critical.

Are there properties of diagrams that account for their use in this kind of explanatory context? Recall that diagrams are symbols with both articulate syntax and a visual format. They are comprised of atomic characters, which are used to refer to components of the biological system. Spatial relations holding among the atomic characters of the symbol are then used to refer to relations among the referents of the atomic characters. For example, the arced shapes in Figure 1 refer to enzyme subunits, and contiguity relations among the symbols for subunits refer to the relation of being in the same complex. Because spatial relations among atomic characters are interpreted as referring to relations among the referents of the atomic characters, a reference to the relations among the components is built into the representation. For this reason, the two kinds of information that make these models explanatorily relevant are presented in one figure: identification of important component parts of a system and of important relations among those components. The visual format allows a single representation to refer to both these kinds of features: the components and relations among components.

The diagrams are representations that refer to both the components of the system in question and the features and relations among components relevant to the capacity of the system. The two-dimensional format of visual representations is especially well-suited to this role, because the visual format of the diagram provides a way to refer to system components in such a way that their relationship to the system itself is also represented. Serial formats can indeed be used to express these models. A serially formatted representation, such as a printed sentence, could do so by naming the components of the system and the relations among those components. The representation would not be concise: in order to describe all the different kinds of relations involved in the model, a series of such statements would be required, or a long conjunction. More importantly, the visible form of such a linguistic representation bears no relation to the structure of the model.

The structural diagram illustrates an additional reason that scientists use diagrams to convey models. The list of individual atomic coordinates would do little for a human in terms of understanding how these locations add up to the functional capacities of the complex. A serial representation of the positions of amino acids is readily available; it can be printed from the same electronically stored file of atomic coordinates which was used to make the diagram of the structure. To make such a diagram, the printer plots lines on a page so that the spatial relations among the parts of the diagram correspond to spatial relations among parts of the molecule. A person who understands the interpretive conventions for this type of diagram (that the lines represent bonds between amino acids, for example) can then look at the diagram and through perception of that symbol, understand the shape of the molecule. In order to get this same understanding of the spatial properties of the molecule from a numerical list of atomic coordinates, a person would have to calculate the relative locations of all the amino acids. Because there are so many amino acids, the computational task is too complicated for humans to accomplish: people can't

comprehend the spatial features of the model in this way. Figure 2, however, makes the spatial relations among atoms easily comprehensible, because this figure uses visible spatial relations to represent the locations of individual atoms. Understanding the figure in terms of the position of individual atoms is bound up with perception of the figure, complete with the now-evident spatial relations among its component parts. Further computation of spatial relations among parts is not required, as it would be from a list of linguistic expressions describing each atom.

Figure 2 is thus a good example of one reason the visual format can function differently in scientific communication from linguistic representations. The human visual system works by integrating sensory cues into a perceptual whole. The result of this process is a perception in which the parts are related to the whole. It is not feasible for a person to infer equivalent content from linguistically encoded information. Machines can do the computations required to infer spatial relations from long lists of atomic coordinates; current computer modeling programs can rotate structures, changing the absolute position of atoms but retaining their relative positions. Humans just cannot do this, but they are excellent at synthesizing complex *visual* information. It should not be surprising that scientists take advantage of these capacities when they design symbols to convey important ideas.

This capacity to make higher order relations cognizable certainly contributes to the value of diagrams. Using spatial relations among parts of the diagram to represent spatial relations among parts of the enzyme structure is the only known way to make the structural model comprehensible: in such a case, a diagram is *cognitively essential*. The fact that the structural diagram makes the information it presents comprehensible is not, however, what makes this diagram *explanatory*. Recall that the content of the diagram of the binding-change mechanism (Figure 1) can be expressed linguistically, and understood from the verbal description. The diagram of the binding-change mechanism may be cognitively convenient (it leads to quick understanding) but it is *not* cognitively essential. There is another reason for scientists to use a diagram: the two-dimensional format conveys the content that makes the model explanatory in a way that is especially well suited to the explanatory relevance of that content. This is true for the structural diagram as well, which means that even for cases in which a diagram is essential for successful communication due to the nature of human cognitive capacities, there is still another reason for using a diagrammatic form of representation.

The analysis of diagrams in terms of their representational format has provided some insight into why a diagram rather than a linguistic representation is used to express a model. In both of the cases just considered, diagrams were used to represent models that provide a complex functional explanation. The unique capacity of visual representations to represent higher-order relations in virtue of visible relations holding among parts of the figure suits them to functional explanation in a way that representations with a linear format cannot match.

The two-dimensional format was doubly important in the case of the structural diagram, because it allows the model to be represented in a way comprehensible through visual perception, while human cognition is not capable of comprehending the spatial relations relevant to the structure from reading a list of coordinates. Linguistic representations will not be of any explanatory value if people cannot use them to understand the facts relevant to the explanation. But this does not imply that the explanatory contribution of visual representations is just the role they play in making complicated relationships comprehensible, and it does not explain why diagrams are used to express simpler ideas, like the binding-change model, which *can* be expressed in comprehensible linguistic representations.

Why are some explanations expressed with linguistic representations and others with figures?

While some biological models are conveyed with diagrams, there are many proposals that made important explanatory contributions to biology which are usually expressed with linguistic or mathematical representations, and there are some which are never expressed through diagrams. This poses no difficulty for the account of diagrams given above: it does not imply that diagrams will be the most effective means of expression for *any* hypothesis. That depends on the relation between a theory's content and the representational capacities of linguistic and visual formats. Structural and functional models relate the component parts of the system to the whole. Diagrams can do this with a representation that builds-in reference to higher-order relations between the referents of component elements of a character. Other theories do not seem to be explanatory in virtue of providing a functional analysis. For example, localization of a capacity to a particular entity or trait is explanatorily relevant (Bechtel and Richardson 1993). But localization hypotheses, such as the claim that F_1 is the complex that catalyzes ATP production, ascribe properties to objects and can generally be conveyed conveniently through serial representations. Causal claims are also examples of hypotheses amenable to serial representation. Examples include both simple causal claims (e.g. HIV infection causes AIDS), as well as more complex hypotheses, such as the theory of natural selection.

The analysis of diagrams as representations provides an explanation for why they are effective for expressing certain kinds of models and theories, and ineffective for others. Diagrams are a form of representation in which reference to elements of a model builds in an expression of the relationships holding among those elements. If the relations that make the theory explanatory cannot be conveyed through spatial relations or other visible features of a diagram, then diagrams offer no advantage over linguistic representations. For those proposals whose key features can be represented spatially, diagrams offer a cognitive advantage in communicating complex information, and make

effective vehicles for a particular kind of explanation: complex functional analysis.

References

Abrahams J.P., Leslie A.G., Lutter R. and Walker J. 1994. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. Nature 370(6491): 621–8.

Bechtel William and Richardson R. 1992. Emergent phenomena and complex systems. In: Emergence or Reduction? Essays on the prospects of nonreductive physicalism. Walter de Gruyter Verlag, New York.

Bechtel William. and Richardson R. 1993. Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research. Princeton University Press, Princeton.

Cummins Robert 1975. Functional Analysis. J. Phil. 72: 741–764.

Machamer Peter, Darden L. and Craver C. 2000. Thinking about mechanisms. Phil. Sci. 67: 1–25.

Goodman Nelson 1976. Languages of Art: An Approach to a Theory of Symbols. Hackett Publishing Company, Indiana.

Kitcher Philip 1981. Explanatory Unification. Phil. Sci. 48: 507–531.

Perini Laura 2002. Visual representations and scientific knowledge. Ph.D. Dissertation.

Salmon Wesley 1984. Scientific Explanation and the Causal Structure of the World. Princeton University Press, Princeton.

van Fraassen Bas 1980. The Scientific Image. Oxford University Press, New York.

Watson James and Crick F. 1953. A structure for deoxyribose nucleic acid. Nature 171: 737.