

DATA GRAPHS AND MECHANISTIC EXPLANATION

Daniel C. Burnston, Tulane University

PENULTIMATE DRAFT; Please cite final version:

<http://www.sciencedirect.com/science/article/pii/S1369848616000169?via%3Dihub>

Abstract

It is a widespread assumption in philosophy of science that representations of data are not explanatory—that they are mere stepping stones towards an explanation, such as a representation of a mechanism. I draw on instances of representational and explanatory practice from mammalian chronobiology to suggest that this assumption is unsustainable. In many instances, biologists employ representations of data in explanatory ways that are not reducible to constraints on or evidence for representations of mechanisms. Data graphs are used to represent relationships between quantities across conditions, and often these representations are necessary for explaining particular aspects of the phenomena under study. The benefit of the analysis is two-fold. First, it provides a more accurate account of explanatory practice in broadly mechanistic investigation in biology. Second, it suggests that there is not an explanatorily “fundamental” type of representation in biology. Rather, the practice of explanation consists in the construction of different types of representations and their employment for distinct explanatory purposes.

Keywords: Data graphs; Explanation; Explanatory Relations; Mechanism Diagrams; Scientific Representation

1. Introduction

Explanation in biology consists in showing how a biological system produces a phenomenon of interest. According to the new mechanistic philosophers of science (Bechtel & Richardson, 1993; Machamer, Darden, & Craver, 2000), explanations are couched as descriptions of mechanisms—organized systems of biological parts performing specific operations, which interact causally to produce the phenomenon. This perspective is a major departure from traditional, deductive-nomological views of explanation, and some mechanists (among others less directly concerned with mechanisms) make a further departure in arguing that visual representations can be a key vehicle for explanation (Bechtel & Abrahamsen, 2005; Griesemer, 1991; Machamer et al., 2000; Perini, 2005; Ruse, 1990; Sheredos, Burnston, Abrahamsen, & Bechtel, 2013). Generally, mechanists have focused on “mechanism diagrams,” which show the parts, operations, and organization of the mechanism. Indeed, biologists often present explanations in mechanism diagrams.

I claim that mechanism diagrams, while important for explaining biological phenomena, are not the only kind of visualization serving this goal. Often, in initial decomposition of a system, biologists pursue the kinds of epistemic activities that mechanists have rightly characterized as positing parts and operations, and that culminate in mechanism diagrams. However, mechanists have not paid attention to some other common explanatory practices, which more frequently

employ data graphs in explanatory ways, and which biologists often pursue following initial division of a system into parts and operations. In particular, data graphs show relations between quantities over time and/or across conditions, and representing these relations is often required to show how some aspect of the phenomenon is produced. Even if the parts, operations, and organization of the mechanism are known, an explanation is often insufficient unless quantitative relationships are also represented. When quantitative relationships are represented in an explanatory way, I call them “explanatory relations.” Looking at how data graphs are employed in active biological research highlights these elements of practice.

Mechanists, with few exceptions (Burnston et al., 2014), have missed or misconstrued the explanatory import of data graphs. It is not obvious why this has been the case, since there is relatively little overall discussion of data representation in the mechanist literature. It may be due in part to an overreliance on textbook accounts of biological discovery, which often paint an accepted mechanism diagram as the end-stage of explanation. Another possible reason is a tendency towards a kind of explanatory fundamentalism—the assumption that some particular kind of description, representation, or entity is the fundamentally explanatory kind, and that all epistemic work in research must be geared towards establishing the type that does the explaining. Some mechanists argue that mechanistic models are explanatorily fundamental, and that representations other than those describing mechanisms, be they verbal descriptions, computational models, network models, or what-have-you, are explanatory only to the extent that they map directly to a mechanistic explanation (for examples, see Craver, 2006; Craver & Kaplan, 2013). There is resistance to explanatory fundamentalism—Bechtel and Abrahamsen (2010), for instance, suggest in several places that mechanistic models must be conjoined with models representing the mechanism as a dynamical system. Pluralists of different stripes suggest that multiple models are necessary for explanation (e.g., Green, 2013).

Clearly, there are some complex issues at foot here. This is largely because these debates often discuss, in one go, the representations that scientists use to explain, the role of models in the explanation, and/or the metaphysics underlying the explanation. Here, I am only interested in representation (see section 4 for a discussion of how the representational question relates to the others). Explanatory fundamentalism vis-à-vis the question of representation argues that there is a type of representation that is fundamentally explanatory, and that other kinds of representation play subsidiary roles in establishing the genuinely explanatory kind. This kind of “representational fundamentalism” would explain the relationship that mechanists have posited between data representation and representation of a mechanism—data graphs are thought to provide *evidence for* or to *constrain* mechanistic hypotheses, but are not thought to themselves be explanatory. I argue that these views underdescribe the use of data graphs in mechanistic research, and that in some cases data graphs serve genuinely explanatory roles.

As mentioned, it is difficult to pin representational fundamentalism on anyone directly. As such, I have articulated it more as a foil than in an attempt to criticize any particular theorist. Arguing against representational fundamentalism points the way to a positive view of how representations are employed in explanation. The view I propose claims that giving an explanation in a given case consists in coordinating distinct representations with distinct explanatory roles. My account is thoroughly pragmatic—I am not offering an analysis of either “representation” or “explanation”

as such. However, I think we can learn something deep about the nature of explanation by understanding the relative roles that distinct representations play in accounting for phenomena.

I will discuss a detailed case study from mammalian chronobiology, which I claim can only be adequately described if data graphs are taken to be playing a key explanatory role in the investigation. I will argue that the role of the data graphs is *dissociable* from those played by those of mechanism diagrams and causal graphs in the investigation. That is, the specific content of the data graphs serves an explanatory role that is needed for the explanation, not present in other representations, and not dependent on their specific content for fulfilling its role.

I will pursue only one case study in order to thoroughly distinguish the representations and the relationships between them. However, I take the case to be exemplary of a wide range of explanatory uses of data representations, and I discuss other examples elsewhere (Burnston, in prep). I will start in section 2 by describing the representational differences between data graphs and mechanism diagrams in the context of mammalian chronobiology. I will then give the case study in section 3 and argue for the distinctive explanatory role of data graphs. In section 4, I will give a general view of the relative explanatory roles of data graphs and mechanism diagrams, and situate the view amongst the related debates in the mechanist literature. Section 5 concludes.

2. Diagrammatic Representation

Mechanistic explanation consists in describing the physical system responsible for an *explanandum* phenomenon (Craver, 2007). In a successful explanation, standardly construed, the system is decomposed into parts and operations, and the explanation shows how those parts and operations are organized to produce the phenomenon in question. A part is an entity or type of entity. An operation is a type of interaction between entities—standard examples include *binding*, *activating*, *regulating*, and *inhibiting*. Organization includes both the spatial locations of parts and operations, and the temporal ordering of operations. It is often important, for instance, that one operation occur before another, at a certain place, so that the second operation can occur. I will use the phrase “mechanistic posits” to refer to hypotheses about and representations of parts, operations, and organization in these traditional senses.

Mechanistic understanding in the field of chronobiology has advanced rapidly in the last 15 years. One of the field’s primary explanandum phenomena is circadian rhythms—roughly 24-hour, endogenously generated physiological rhythms occurring in a wide range of organisms, which regulate an array of processes ranging from sleep and activity patterns, to metabolism, to gene transcription across the entire genome. Rhythmicity, in mammals and many other organisms, is due to internal “clock” mechanisms, which in eukaryotes operate at several mechanistic levels. The most basic timekeeping mechanisms are molecular clocks within individual cells, which are composed of interlocking feedback loops amongst gene products (mRNAs and proteins). Circadian “time” is kept in the oscillating quantities of gene products—when the clock is working properly, these oscillate over a roughly 24 hour period, thus providing a timing signal that can regulate other processes.

In abstract terms, the basic mechanism works as follows: proteins from the “positive loop” cause transcriptional activation of the “negative loop” genes, leading to increasing quantities of

negative loop gene products. The negative loop proteins then feed back and inhibit their own transcription. When negative loop proteins degrade, the inhibition is released, and activation can resume, causing levels to increase again. With the appropriate rates of transcription, translation, and degradation, the mechanism can produce oscillations of roughly 24 hours. In mammals, the positive loop genes are *Bmal1* and *Clock*, and the negative loop genes are *Cryptochrome* (*Cry*) and *Period* (*Per*). The standard “mechanism schema” (Machamer, Darden, & Craver, 2000) of the clock is diagrammed in the figure below.

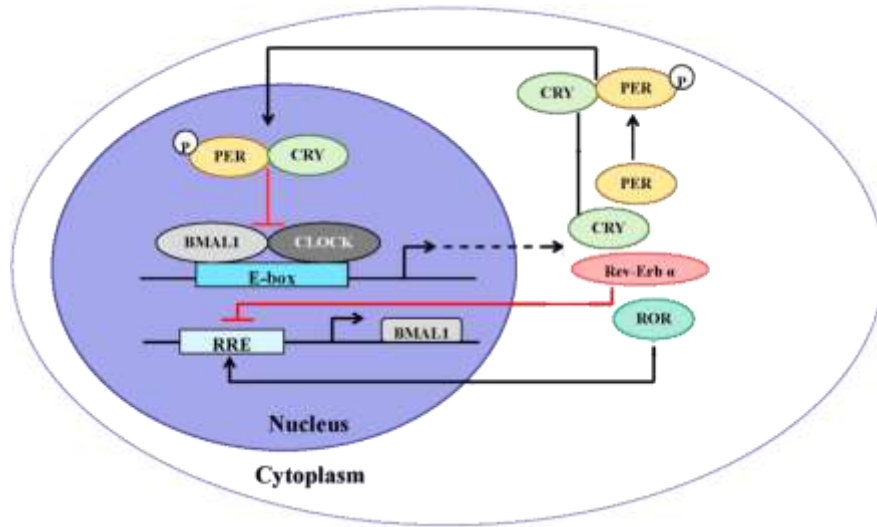


Figure 1. The mammalian intracellular clock mechanism. Modified from Wang, Zhang, Xu, and Tischkau (2014).

Genes and gene promoters (the sites where proteins bind and produce their effects) are shown as rectangles; proteins are shown as ovals. The positive loop proteins, after transcription, dimerize and are translocated into the nucleus, where they bind to the E-boxes of *Cry* and *Per*, activating their transcription. The negative loop proteins undergo a similar process before binding to the positive loop proteins at the E-box, halting their activation of the negative loop genes. The diagram, in addition, shows the other elements of the positive loop—of particular interest is the *Rev-Erba* gene, which is activated by the positive loop proteins, and whose protein feeds back to inhibit *Bmal1* transcription at the RRE promoter, leading to oscillating quantities of *Bmal1* products. (The ROR protein plays a supporting role in the positive loop, and I will not discuss it here.) The operation of activation is shown via straight arrows, transcription by bent arrows, translation by dashed arrows, and inhibition by flat-headed lines; the operations of dimerization and binding are shown via spatial contiguity.

There are several more aspects to the molecular clock in mammals, and more complete mechanism diagrams include more parts and operations. One particularly important aspect is that each of the clock genes has paralogs—structurally similar genes with similar overall functions (e.g., in the positive or negative loops), but which can play slightly different roles in certain circumstances. The paralog of *Clock* is *Npas2*, of *Rev-erba* is *Rev-erb β* , of *Cry1* is *Cry2*, etc. *Per1* has two paralogs, *Per2* and *Per3*. Moreover, while I will only discuss intracellular clock mechanisms here, it should be kept in mind that in mammals the clock is in fact a multi-

level mechanism in which a central pacemaker (the suprachiasmatic nucleus in the hypothalamus) regulates biological time throughout the body.

A representation is “explanatory” when it shows, perhaps in part, how a phenomenon is produced. Mechanists often claim that such diagrams as Figure 1 are explanatory, because they show the organization of parts and operations that leads to the phenomenon of rhythmicity. The resources the diagram provides are due to its particular format and means of representation—contrary to linguistic representations, for example, the diagram retains a spatial format to represent spatial relations within the mechanism, and uses this property to represent the organization of operations. As such, one can mentally simulate the sequence of operations in the mechanism, moving from one to another to see how the phenomenon comes about due to the “productive continuity” (Machamer et al., 2000) of the parts and operations represented.

One of the strengths of the mechanist approach is that it can account for some important aspects of explanatory practice in biology. Biologists do often in fact often offer diagrams like Figure 1 as explanatory representations. Mechanists, however, have historically not granted any explanatory role to data graphs, despite the fact that they vastly outnumber mechanism diagrams in biological research literature, and despite their very different structure, which suggests that they might provide other resources for understanding how phenomena are produced. Consider the diagrams in Figure 2.

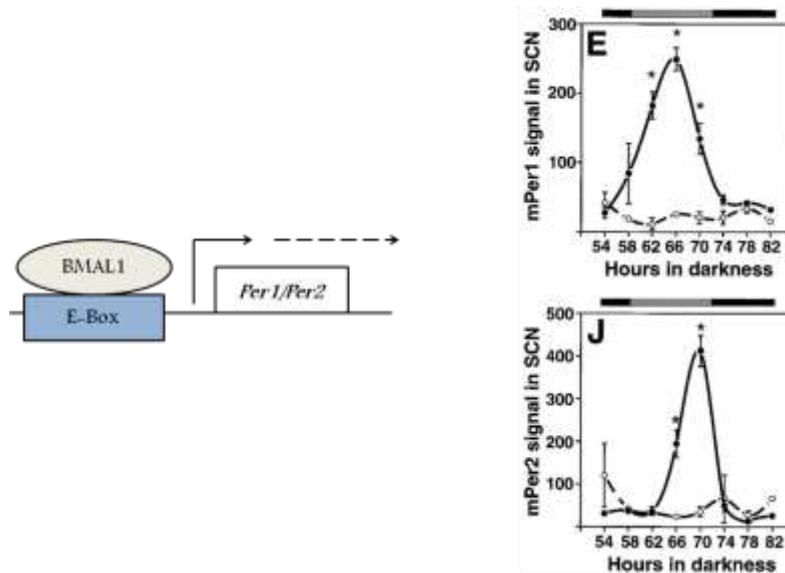


Figure 2. (Left) A mechanism diagram portraying the operation performed by BMAL1 on the *Per* paralogs. (Right) Bunger et al.’s (2000) data, which established the relationship.

The left panel is a diagram of the activation of *Per1* and its paralog *Per 2* by BMAL1, the protein product of the positive loop gene *Bmal1*. The right panel includes two data graphs from Bunger et al. (2000), the paper which is standardly credited with having established the activation relationship. What are the relevant differences?

The mechanism diagram on the left uses symbols to explicitly classify the interaction between BMAL1 and the *Per* paralogs as an *activation* operation, and doing so has a variety of benefits. For instance, it shows the interaction to be alike with other activation operations performed by other parts on other targets. However, symbolizing in this way means that the diagram cannot convey quantitative properties regarding the processes it represents. No information, for instance, about the rate or levels of activation, the quantities of *Per* products transcribed, or the timing of the rhythms is included in the diagram. Data graphs are the opposite: they dedicate their spatial dimensions to showing particular quantitative values, and therefore can only indicate types of parts or operations via labels or captions. The data graphs in the figure above plot particular results from a knockout condition compared to wild type. In the graph, each point represents a measured value of *Per* mRNA quantities in clock cells (*Per1* at top, *Per2* at bottom). Filled dots, connected by solid lines, represent quantities taken from wild type SCN cells. Open dots, connected by dashed lines, represent *Per* mRNA quantities (I will generally refer to mRNAs as the gene name preceded by ‘m’, as in ‘mPer’) in cells from *Bmal1* knockout mice. Each measurement is represented as occurring at a specific time—measured in the number of hours after the animal was exposed to constant darkness (this is done generally to dissociate the internal properties of the clock mechanism from its responses to light).¹

A first pass at the distinction, then, is that the data graphs on the right display specific quantitative values, while the ones on the left don’t. Arguing that the latter are explanatory, whereas the former aren’t, then, means that something explanatory is gained, and nothing explanatory lost, by representing the system in terms of parts and operations rather than in terms of specific quantities. And conversely, arguing against the claim, or at least against its universality, requires pointing to explanatory contexts in which the quantitative aspects of the representation cannot be omitted without explanatory loss. In the next sections I will discuss such a use.

There are a few more things to note about the data graphs such as those in Figure 2. First, they represent specific data patterns *over time* and *across conditions*. These allow for comparisons of quantities to reveal properties such as relative *peaks*, *periods*, and *phases*, as well as to track how these comparisons vary depending on the condition. While temporal relationships are highly important in chronobiology, I will primarily focus on diagrams that represent quantities across conditions in this paper (see Burnston, in prep], for a discussion of the importance of data graphs in representing phase relationships). Second, there are a variety of statistical techniques that go into the construction and employment of data graphs—one will notice the error bars superimposed on the dots in the data graphs in figure 2, and the stars representing statistically significant differences between values in the knockout and wild type conditions. A full treatment of the relationship between statistical measures and data graphs is outside the scope of my project here; for now, provisionally, the statistical tests can be seen as a justificatory tool for employing the representation of the quantitative relationships shown in the data graph.

¹ Data were collected via in-situ hybridization, a method in which a fluorescent tracer is affixed to certain mRNAs. The units on the X-axis reflect detectable fluorescence, thus signaling the quantity of mRNA. This is similar in the data graphs in other sections, although different techniques involve different units. The light/dark bars in between the graphs represent the animals’ standard activity period during their “subjective day” in the period covered by the light bar, and the inactive period in the subjective night in the times covered by the dark bar. Light/dark bars can also be used to show the external light schedule.

In sum, I take data graphs to be used to represent *quantitative relations* (QRs) in the mechanism being studied—particular values of measureable quantities, where often these are compared over time and/or across conditions. For instance, the phase relationship between mPer1 and mPer2 is a property of the *relative* quantitative patterns shown in the graph, and is thus a paradigm example of a QR. The relationship, of course, only holds in the WT condition—it does not occur in the knockout. Moreover, the relationship between mPer1 quantities across the two conditions is itself a QR. That is, the quantity of mPer1 exhibits the pattern shown by the solid line in the WT condition, and the dashed line in the knockout condition. A representation of a QR is explanatory when the quantitative, temporal, and/or contrastive properties it conveys are needed to give a full account of how an explanandum phenomenon comes about—that is, when scientists must represent the QR to sufficiently meet an explanatory goal. When a QR is explanatory in this sense, I call it an *explanatory relation* (ER).² An ER can, potentially, hold between any number of quantities in a system, and often discovering ERs involves creative decisions about what quantities to measure in what conditions (Burnston et al., 2014). Often, ERs fall into types—some of these include being “in-phase” or being “proportional.” However, in order to be an ER, a type must be implemented in particular system components, occurring across particular conditions and/or particular times. Data graphs are used to represent the relationships that hold between components in the mechanism under particular conditions.

Unsurprisingly, given the little attention to these representations in the literature, QRs are a bit of an odd duck when it comes to the explanatory categories that mechanists have discussed. On the one hand, on a broadly interventionist account of causation, the relationships conveyed in data graphs are often causal—there is a systematic variation between the manipulated quantities and the observed quantities. On the other hand, simply this observation doesn’t meet many of the conditions that mechanists take to be important for giving explanations. For instance, Woodward (2010, 2013) adds on the conditions that explanatory causal relationships should be *stable* and *invariant*. There should be a range of conditions in which the observed relationship holds, and they should hold stably in those conditions. Explanatory causal relationships should be *specific* and *direct*. Both Woodward (2010) and Craver (2007) stress the norm that full explanations should fill in gaps between indirect causal influences (cf. Glennan, 2002), and determine whether the effect is due solely to the posited cause. The data graph, due to its representing particular quantities across particular conditions, doesn’t convey any of these potentially relevant properties of the relationship. Moreover, one might notice QR’s simply by contrasting conditions of observation, without physically manipulating the system (although showing that a certain pattern is necessary for a phenomenon to occur often involves manipulating it). So, it seems like any potentially explanatory role for QRs doesn’t depend primarily on their being causal relationships (see section 4 for more discussion). More importantly, focusing the properties of causal relationships distracts attention from what I *do* take to be the most important part of the representation, namely the specific *pattern* in the represented

² I will be somewhat loose in talking about a representation of a relation providing an explanation, versus the relation itself providing it. Some mechanists are proponents of the *ontic* view of explanation (e.g., Craver, 2007), which holds that, ultimately, explanations are themselves in the world. Others (e.g., Bechtel & Abrahamsen, 2005) hold the *epistemic* view, on which representations explain, and the world itself does not. I take no stand on this debate here. An ontic theorist will have to hold that it is only the ER as it obtains in the world that explains, while an epistemicist will have to hold that only the representations do. I will speak both ways as convenient; both ontic and epistemic theorists should feel free to regiment the meaning as they see fit.

data. My preferred construal of these relationships is thus along the lines of Dennett's (1983) "real patterns"—they are recognizable patterns occurring in certain observable conditions, where the types of patterns and conditions that are relevant depend on what one is trying to explain.³

The shortcomings of data graphs for representing causal relationships have perhaps informed the limited roles that mechanists have granted to them. Representations of data are standardly posited to *provide evidence for* or *constrain* mechanistic posits. Bechtel and Abrahamsen (2005) suggest that the data given in the results section of a paper provide evidence for the explanatory mechanism diagram advanced in the discussion section.⁴ Machamer, Darden, and Craver (2000) make a similar claim. In Craver's (2007) view, the results of experiments provide evidence that specifies the explanatorily relevant causal relationships in the mechanism, which are described as operations. Darden and Craver (2002) suggest a variety of ways in which quantitative details can constrain mechanistic positing, including by specifying the rate and duration in which operations occur. For instance, if a phenomenon involving two mechanistic operations occurring sequentially takes one second, and it is then established that the first operation occurs in 800ms, posits regarding the second operation had better pick one that occurs within 200ms. The rates, in this case, are not explanatory in themselves, but only constraints on positing specific operations and what order they stand in. Similarly, Craver (2007) stresses the use of contrasts between conditions in a variety of cases, but only for pragmatic purposes. Submitting a proposed mechanism to different conditions, for instance, can help tease apart "how-possibly" from "how-actually" models, but for Craver the contrasts between conditions are not themselves explanatory.

The evidence and constraint views of data representations presuppose that data graphs play a role in establishing the representations—namely, mechanism diagrams—which explain phenomena, but are not themselves part of the explanation. This is certainly right in *some* cases. Data graphs are often used for evidential or constraining purposes (as in the graph in Figure 2), as well as for a variety of other non-explanatory purposes (such as to validate a method or rule out an artifact). However, claiming that the evidence and constraint view exhausts the function of data graphs entails a view of explanatory practice which I will claim is false. On the evidence and constraint view, data representations (e.g., in new research papers) should be employed primarily in either generating a mechanism schema or modifying an existent one—for instance by challenging accepted posits of parts, operations, and organization, or by adding new elements to existing mechanism schemas to more thoroughly fill out the explanation (Craver, 2007; Woodward, 2010). In the case study below, I argue that in active research, new data is often represented without the goal of modifying a mechanism schema, but instead in order to show QRs between properties of the mechanism that are necessary for explaining particular aspects of the phenomenon. That is, they are used to represent ERs.⁵

³ I don't want to bring all of Dennett's analysis on board, however. Dennett identifies his real patterns with regularities or generalizations, and I think this is by no means an obvious or necessarily desirable identification to make. See section 4, and [Burnston, in prep].

⁴ Bechtel and Abrahamsen have recently modified their view to give a greater explanatory role to data graphs (Burnston et al., 2014), and I aim to give further support to this move here.

⁵ I will discuss a case in which the research practice was geared specifically towards establishing ERs, in order to clearly distinguish their explanatory roles. In a lot of scientific practice, however, the interplay between data and mechanism representations is extremely close and multi-faceted. See Burnston et al. (2014) for a detailed case study of how the two forms of representation contribute to the development of an explanation in a single research paper.

3. Data Graphs and Explanatory Relations

The canonical model of the clock for mammals, including all of the major components and operations, as well as the standard organization shown in Figure 1, was in place by 2005 (Zhang & Kay, 2010). While there are some exceptions (see, e.g., Ye, Selby, Ozturk, Annayev, & Sancar, 2011), the standard mechanism schema has largely gone unchallenged. In rather short order, however, mammalian chronobiologists turned their attentions from establishing mechanistic posits to discovering ERs. Since these practices, revealed in explanatory usage of data graphs, *do not* involve challenging, modifying, or extending the mechanism schema of the clock, the evidence and constraint views leave out some genuinely explanatory activity.

A particularly clear example comes from Baggs et al. (2009), who used small-interfering RNA (siRNA) to perform fine-grained knockdowns of gene products in the clock. SiRNA degrades specifically targeted mRNAs. The effect is dose-dependent—the more siRNA inserted, the greater the decrease in the mRNA, and thus in the protein coded for by the mRNA. Quantitative measurements can then be taken to show the effect of these knockdowns on *other* components of the system.⁶

The authors specifically addressed the explanandum phenomenon of *compensation*. A variety of compensatory effects occur in gene regulation networks, in which genetic processes are maintained despite fluctuations in the quantities of key components of the system. In the circadian case, compensation means that rhythmic processes are maintained despite variation in internal or external conditions. At the molecular level, compensation is exhibited when rhythmic gene products retain rhythmicity despite variation in the levels of other components. Baggs et al. use siRNA manipulations, and the relationships uncovered, to explain how this phenomenon occurs. They use data graphs to show a variety of relationships between system components that, together, help to show how molecular rhythms are maintained across different knockdown levels. Three examples of key ERs are shown below. The bar graphs are color coded to reflect the increasing amounts of siRNA of a particular type inserted (as shown in the legends in the top left; the legend goes from least insertion at the top to greatest insertion at the bottom). The quantities are represented relative to a baseline quantity of mRNA for each type (the ‘1’ in the scales on the left).

⁶ Measurements of mRNA levels were taken with quantitative PCR. Gene expression levels were measured with bioluminescence reporters. Both techniques involve the insertion of light-emitting biological reactants into the DNA of a particular gene, although they differ in a variety of respects. I will not discuss these techniques in detail here.

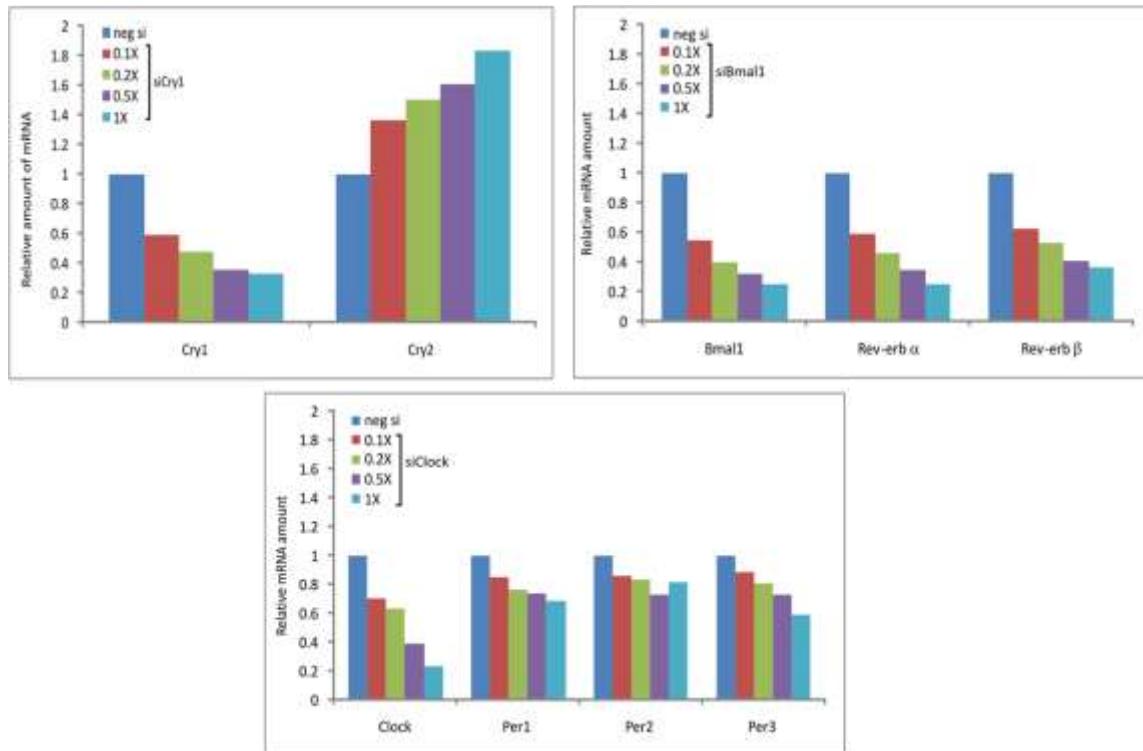


Figure 3. Data showing varieties of relationships between mRNA quantities in the mammalian clock. From Baggs et al. (2009).

On the top left is the result of siCry1 insertion on the levels of both mCry1 and mCry2. As can be seen, increasing insertion decreases (as expected) levels of mCry1. It also, however, *increases* levels of mCry2. Moreover, this increase is *proportional* to the decrease of mCry1. *Cry2* is a paralog of *Cry1*, and it is easy to see why proportional increase of paralogs in response to specific gene knockdowns could contribute to compensation. Since paralogs have similar effects on similar targets, a proportional increase of one in response to knockdown of the other could produce an overall effect similar to normal conditions. Proportional responses also occur between non-paralogous gene products. The top right panel of Figure 3 shows the effect of knockdowns of mBmal1 on *Rev-erba* and *Rev-erbb*. It would be expected from the standard mechanism schema given in Figure 1 that *Rev-erb* levels would decrease upon decrease of mBmal1, since BMAL1 proteins activate the *Rev-erb* genes. However, the proportionality of the relationship is an important further fact. To see this further, consider the bottom panel, where the effect of mClock knockdowns on the *Per* paralogs is shown. Despite the fact that *Clock* products *activate* the *Per* genes—the same type of operation that *Bmal1* performs on the *Rev-erb* genes—the quantitative relationship is different in this case. While it is proportional, it is “fractionally proportional”—the slope of the relationship is fractional, rather than being close to one (I discuss the use of equations to characterize proportionality below).

Baggs et al. discuss a variety of further types of relationships, but these will be enough to show how the representation of ERs allows them to explain compensation. Consider the figure below:

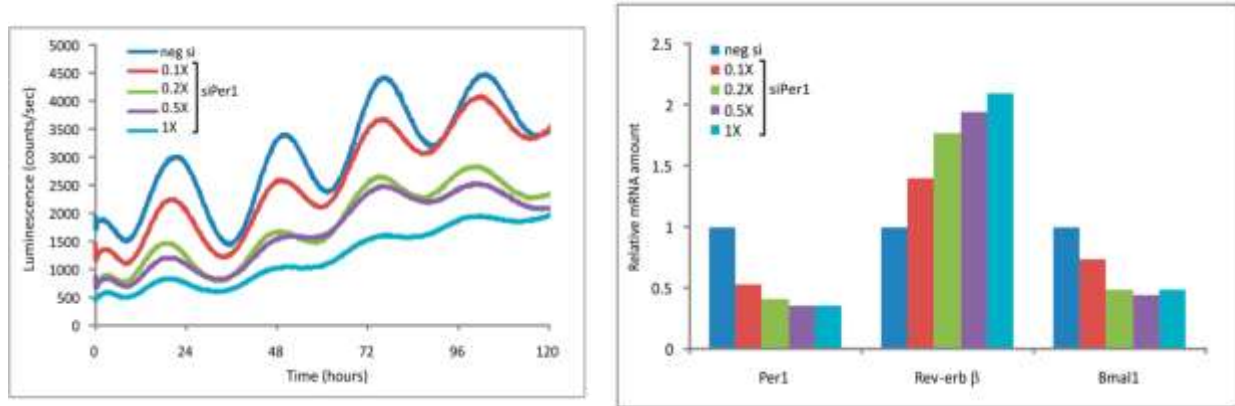


Figure 4. Compensation for mPer1 knockdown. From Baggs et al. (2009).

On the left is a line graph representing quantities of *Bmal1* transcription under several levels of mPer1 knockdown, color coded according to the amount of siPer1 inserted. The graph shows that *Bmal1* expression retains rhythmicity (albeit with decreased amplitude) for several increasing increments of siPer1 insertion, thus exhibiting the explanandum phenomenon of compensation. The explanation for compensation in this case is given in the diagram to the right, showing that, in response to *Per1* knockdowns, both *Bmal1* and *Rev-erb* transcription show proportional effects—*Rev-erb* increasing, *Bmal1* decreasing. The net result of these combined proportional relationships is to “balance” the quantity of *Per1* products in the negative loop with those of *Bmal1* in the positive loop. Baggs et al. offer this balance as a key to understanding compensation in this case. Recall that rhythmicity in the clock is the result of interlocking activities of the positive and negative loops. As such, quantities in one loop being disproportionately larger than the other would interrupt the precise patterns of activation and inhibition that are needed to produce rhythmicity. Thus a means for retaining balance between positive and negative loop products is important for compensation, and Baggs et al. offer the proportional relationships between mPer1 and mRev-erb β , and between mRev-erb β and mBmal1, as the means by which the balance is maintained despite the knockdown. Compensation is explained due to the particular patterns in quantitative relationships across conditions. The proportional patterns of variation between mPer1 and mRev-erb β , as well as between mRev-erb β and mBmal1 help to explain how the pattern of balance between mPer1 and mBmal1 occurs. This pattern in turn helps explain compensation via balancing quantities in the positive and negative loops.

Other instances of compensation are explained in a similar fashion—one slightly more complicated example is given below, in showing how *Bmal1* rhythms are maintained in response to several levels of mClock knockdowns.

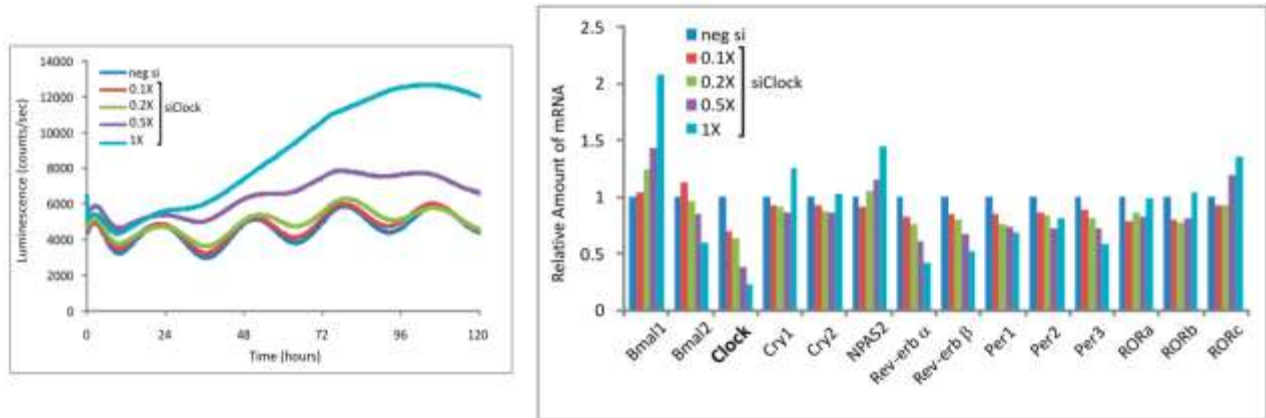


Figure 5. Compensation for mClock knockdown. From Baggs et al. (2009).

In this case, the figure at the left shows that mBmal1 rhythms are maintained for a range of siClock levels (the wavy lines at bottom) until eventually increasing and becoming arrhythmic for very large quantities inserted. While many gene responses are shown in the graph on the right, the important ones are *Npas1*, *Bmal1*, and the *Per* paralogs. *Npas1* is the paralog of *Clock*, and its mRNA shows proportional increase as mClock is depleted. This corresponds with a proportional increase of *Bmal1*, mediated by proportional decrease of the *Rev-erb* genes which normally inhibit it. The *Per* paralogs show the same fractional-proportional decrease shown in Figure 3 (bottom panel). In this case, the “balance” that occurs between the positive and negative loop genes is between the overall positive loop genes and overall negative loop ones—the increase of *Bmal1* and *Npas2*, combined with the fractional-proportional decrease of the *Per* genes, keeps the overall quantities in balance. Up to 70% decrease of mClock can be compensated for in this way.

Details aside, what is important is the kind of reasoning displayed here, and how the data graphs contribute to the explanation. Baggs et al. present the discovery of proportional relationships (of different types) between paralogous and non-paralogous genes as set of *additional* explanatory facts—i.e., facts going beyond the accepted mechanism schema of the clock—about how the mammalian clock can exhibit compensation. That is, it is not *just* the fact that the parts exhibit the operations and organization standardly attributed to them, but also that they respond to perturbations in the particular quantitative patterns that they do, that explains how compensation comes about. In their summation, Baggs et al. “propose that the clock ... combines ... activator and repressor modules with various forms of proportionality to construct relays that generate complex gene expression responses to single gene perturbations” (2009, p. 0570). That is, the proportional relationships must be represented in addition to facts about the mechanism schema in order to explain compensation at the molecular level.

Nor is it coincidental that the assorted ERs are represented in data graphs. Proportionality is an *inherently* quantitative and contrastive property of the relationships between clock mechanism components. As such, showing it via a representation which conveys the quantities and the conditions across which the proportional relationship holds is appropriate. The data graphs are used here to exemplify the kinds of relationships that particular system components have in knockdown conditions—e.g., proportional, fractional-proportional, etc.—and, as in Figures 4 and

5 above, these relationships are offered as part of the explanation for how compensation occurs in particular cases.

The evidence and constraint views, which argue that data graphs play only the roles of questioning or expanding a mechanism schema, do not describe the case. Baggs et al. do not offer proportionality and balance as new parts or operations to be added to the standard mechanism schema. Indeed, it is hard to even make sense of them in terms of mechanistic posits, since they are adverbial (two quantities change *proportionally* or in a balanced way).⁷ Nor are any of the current parts and operations questioned. So, the role of the data graphs is not to modify the extant mechanism schema. There is a sense in which the data graphs *could* be a spur to an extension of the schema. Baggs et al., drawing on the broader literature in gene regulatory systems (Kafka, Levy, & Pilpel, 2006), note that in the particular instance of paralog compensation, there are several mechanisms that could underlie proportional relationships. Drawing on an analogy to developmental systems, they even hypothesize that there may be direct inhibitory influence by *Per1* and *Cry1* on their paralogs, which is lessened as mPer1 and mCry1 quantities decrease. On the evidence and constraint view, the value of discovering proportional paralog responses *depends on* and is subsidiary to discovering the causal structure at work. But Baggs et al. do not suggest at any point that the explanatory value of their data graphs depends on the specifics of the eventual “complete” mechanistic description. They do not perform experiments to determine whether their hypothesis about the lower-level mechanism is correct, simply leaving it open as another hypothesis that could be tested. And given the way the data graphs are employed in the investigation, it seems clear that proportionality and balancing will remain necessary for explaining compensation *regardless of how* this additional question comes out—whichever possible lower level causal structure turns out to be at work, the ERs of proportionality and balancing will play the exact same role in the explanation. Moreover, for non-paralogous relationships, Baggs et al. do not even offer further mechanistic positing as a relevant consideration. The relevant mechanistic organization is already known.

So, the best reading of the case is that the data graphs are used to represent ERs, rather than as providing evidence or constraint for mechanistic posits. In fact, it is arguable that in this case the standard *mechanism* diagram is what is providing the constraint. Given what the standard schema says about the parts and operations in the mechanism, what patterns do the relative quantities of those parts exhibit such that they can implement compensation? If this is right, then sorting some representations as falling solely into the “explanatory” box and others as falling solely into the “evidence/constraint” box is a bad description of practice.

I am only claiming that data graphs play *a* representational role that is needed for explaining compensation. This claim needs to be immediately distinguished from the claim that the data graphs, on their own and as such, *are* or *constitute* an explanation of compensation. It is easiest to distinguish these points by extending a discussion of Craver’s regarding explanatory models to the current discussion of representation. Craver (2006) argues that models that do not convey

⁷ In a loose sense, of course, these are facts about organization, since the quantities represented have effects at particular places (e.g., on the promoters of the genes they activate) over time. While mechanists have not discussed ERs as part of organization, it is possible that the notion could be expanded to include representation of ERs. This move would, however, constitute an acceptance of my point, since it requires accepting an explanatory role for some representations of QRs, above and beyond the evidence and constraint roles.

mechanistic organization are not explanations, because they fail to show the causal structure that produces the phenomenon.⁸ However, he then moves from this to the more substantial claim that mechanistic models *are* explanations, and that any explanatory role for other models consists solely in their helping to establish mechanistic ones. This kind of argument makes a lot of sense on representational fundamentalism. If we assume that one type of representation is the basis for explanation, then showing that *some* explanatory task cannot be completed by a particular type of representation suggests that they cannot constitute explanations. But this kind of view simply does not describe the explanatory use of representations in the Baggs et al. case. Baggs et al. do not offer the data graphs as a *replacement* or *alternative* for the well-understood mechanistic model. Instead, they offer the graphs as representations conveying needed content not present in the representation of the accepted mechanism.

This result, I submit, is incompatible with representational fundamentalism. Rather than one type of representation being the basis of explanation, the current considerations suggest that distinct representations play distinct roles in explanation. If representational fundamentalism is false, then we need an account of the relationship between the explanatory roles of different types of representation that neither makes one type a self-sufficient explanation, nor makes the explanatory role of one type depend entirely on that of another. I offer the following definition of “dissociability” to fill this need:

Dissociability: The explanatory role of a representation-type R1 is dissociable from that of a representation-type R2, in explaining phenomenon P, if R1 (i) conveys content not representable in R2, and (ii) that content would remain necessary for explaining P even given changes to the content in R2.

The Baggs et al. case, I claim, establishes that the explanatory role of data graphs is dissociable from that of mechanistic posits in this sense. Condition (i) is evidenced in the fact that the standard mechanism schema does not convey the needed patterns. Condition (ii) is evident in the fact that while one often *could* make further mechanistic posits related to (e.g.) proportionality, the explanatory value of the ERs does not depend on the outcome of that investigation. That is, our detailed mechanistic understanding could change in several ways and several times without at all modifying the explanatory role and value of the ERs.

So, while it is right to point out that data graphs do not discriminate between distinct possible causal structures, concluding from this that representations are only explanatory to the extent that they help to establish mechanistic posits is a mistake. The ability of the data graphs to convey the needed pattern is based on their representational format, which is precisely what *prevents* them from representing mechanistic organization—they are simply playing a different role in the account. Moreover, the notion of dissociable representations suggests that *no* one type of representation is likely to give a “complete” explanation of some phenomenon. If the considerations given here are correct, we are likely to need multiple representations playing distinct roles to explain any particular phenomenon.

⁸ Craver is specifically talking about Hodgkin and Huxley’s dynamical model of the action potential, but it seems his argument would extend to any representation that does not represent mechanistic posits.

Dissociability is thus a weaker relation than either explanatory *subsumption* or explanatory *independence*. Representational fundamentalism says that the value of non-fundamental representations in the process of giving explanations is subsumed by the explanatory value of the fundamental type that they eventually help to establish—i.e., that non-fundamental representations are *only* valuable to the extent that they help establish the fundamental representation. But if the explanatory role of data graphs is dissociable from that of mechanistic posits, then their role is not subsumed by or reducible to the particular content provided by mechanistic posits. This is far from saying that the two are unrelated, or that the data graphs are a self-sufficient explanation, independent of mechanistic considerations. Obviously, the mRNAs discussed must be shown to be parts of the mechanism before ERs concerning them can be relevant. Similarly, if the mechanism did not include multiple feedback loops, then no one would be motivated to look for ERs that balanced their quantities under varying conditions. But this relationship is best construed as one of mutual constraint between data graphs and mechanistic posits, not one in which they are entirely independent or one in which all of the explanatory power is vested in one at the expense of the other (see section 4).

While I have argued that the explanatory use of data graphs is dissociable from that of mechanistic posits, this is not quite sufficient to establish that they play a distinct explanatory role as such. For instance, data representations might be dissociable from mechanistic posits but subsumed by some other type of representation. Here I briefly discuss one possibility—the relationship between data graphs and causal graphs (Pearl, 1995, 2000). Causal graphs comprise two elements: a qualitative network or path diagram symbolizing the causal connections between nodes, and a set of structural equations characterizing the nature of the influence between nodes. In fact, Baggs et al. *do* offer network diagrams that they extract from the relationships explored in the paper. Moreover, they also test for the proportionality of the relationships by fitting equations to the data. So, why are the data graphs themselves used in an explanatory way, rather than only being a stepping stone to a completed causal graph?

Here are examples of the network diagrams compiled by Baggs et al. The one on the left is for mClock knockdowns, the one on the right for mPer knockdowns (they only show graphs for one knockdown at a time, but a more complete graph could be compiled from the individual ones).

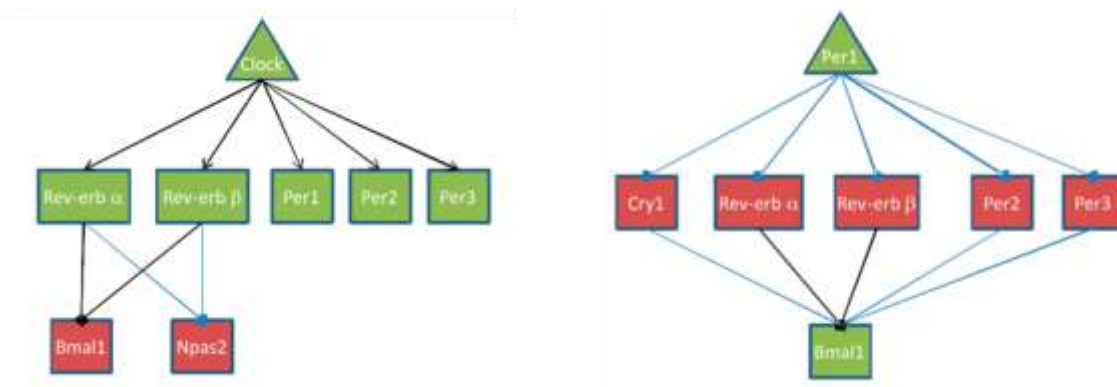


Figure 6. Network diagrams of responses to mRNA knockdowns. From Baggs et al. (2009).

In the graph the labeled rectangles signify the mRNAs. The mRNA labeled at the top is the one being studied via knockdown. The edges are causal relationships; green boxes represent that the mRNA labeled was decreased in response to the knockdown, red that the mRNA increased (the blue and black dots signal whether the relationship had been previously published, and don't concern us here).

Let's take the two aspects of a causal graph—the qualitative graph and the equations—in turn. First, the proportionality shown in the data graph is not included in the qualitative network diagram, meeting condition (i) of the notion of dissociability. Condition (ii) is also met. Suppose we discovered that the causal graphs were different—involving more steps between interacting mRNA quantities, for instance. This would certainly change the overall understanding of the system, but it would not modify the explanatory usage of the data graphs. Proportionality of compensatory relationships and balance between positive and negative loops would be important even if there were more or fewer steps in the pathways. Since the explanatory role that the data graphs are playing could be played even if the graph were different, their status as explanations does not depend on the specifics of the causal graph. The two are dissociable.

Again, as above, I am not denying the importance of the causal graph for providing needed resources. Baggs et al. characterize their network diagram as providing “a framework to uncover ... biological relay mechanisms,” (2009, p. 0568) and this is at least superficially similar to Pearl's characterization of causal graphs as “oracles for interventions” (2000, p. 22). Moreover, there *are* certain explanatory tasks for which data graphs would be subsumed under the explanation given in a causal graph. Consider Pearl's discussion of the relationship between causal graphs and structural equations. Pearl argues that the combination of the causal graph and the equation allows for predictive power under manipulations—if one has the causal graph and the equations right, one can predict the response of the effect variable under manipulations to the causal one. However, the explanatory task in which the data graphs are employed in the Baggs et al. case is *not* simply to predict what each variable will do in response to manipulations. Instead, the particular types of relationships shown in the data graphs are used to give an account of how compensation comes about in the system. This use is what renders the dissociability between the two kinds of representations. *If* the goal were simply to determine the causal graph that can aid prediction of the outcome of manipulations, then the primary task would be to develop the graph (e.g., by mediating between alternative hypotheses about the structure of the graph), and any represented data would be used in the service of that project. But since it is the particular quantitative patterns that are offered as explanatory representations in the Baggs case, this use is insulated from particularities in how the causal graph comes out. The fact that both explanatory tasks are important does not mean that one should be collapsed into the other.

What of the use of equations to test for linearity in the proportional relationships? Doesn't a linear equation provide a more powerful explanatory tool than the represented data? It is worth noting, although not conclusive, that Baggs et al. do not offer the equations as explanations. Instead, they are backgrounded as simply another way of characterizing the relationships shown in the data graph (they count a relationship as proportional if the slope of the equation characterizing the data is close to one, and fractionally proportional otherwise). Now, we might say that the data graphs are merely a visual convenience, standing in for the genuinely

explanatory equation. It would be hard to know what that claim was based on, however, if not some substantive view of what representations are explanatory—for instance, that quantitative ones are universally more explanatory than qualitative ones—which is exactly what is in question. How should we assess the issue?

There is analogy to be made here to Pearl's (2000) argument concerning why qualitative causal graphs are necessary, as opposed to simply representing causal relationships in equations. He argues that the qualitative graphs supply representational resources that are distinct from those given in the equations—i.e., they show causal paths that are difficult if not impossible to keep in one's head all at once, and that must be assumed anyway to make the equations meaningful. A similar account can be given for the data graphs, vis-à-vis the equations. They are used to show specific relationships between several components under specific conditions—i.e., it is not just proportionality between individual components under manipulation, but the way that those relative proportional relationships combine in specific situations to maintain balance between positive and negative loop gene products. This information is not only not explicit in the equations; it would be extremely difficult to extract simply from looking at the equations characterizing the relationships. But if we are willing to admit that causal diagrams—a qualitative type of representation—are explanatory in addition to the equations because they play a specific role, we should not balk at admitting an additional type of qualitative representation if it plays a distinct, clearly articulable role. This is what I have tried to establish.

There is further support for this picture in both recent philosophical work and results from cognitive science. Shah, Mayer, and Hegarty (1999) have established that well-designed data graphs afford the recognition of trends in numerical relationships. Stieff, Hegarty, and Deslongschamps (2011) have shown that, when reasoning about quantitative questions in chemistry, students make heavier use of data graphs showing patterns in quantitative relationships than equations modeling those patterns. Now, these studies looked at the performance of undergraduates rather than practicing researchers. However, in other work my colleagues and I (2014) have closely analyzed the use of data graphs during the development of mechanistic explanations by expert practitioners, and argue that looking for QRs and patterns in data is vital for developing an understanding of the organization of a system. If these accounts are suggestive, then there is something to be gained in Baggs et al.'s use of the data graphs other than visual convenience. The representations render the key relations of proportionality and balance discernible in a way not reducible to that given in the equations, which explains why Baggs et al. choose to represent the relationships in this format.

I have offered an analysis of the explanatory practice in the Baggs et al. paper, and posited a clear explanatory role for the data graphs, which is dissociable from those played by other representations. In the final section, I summarize the view of representation suggested here and discuss how it relates to other discussions in the mechanist literature.

4. Explanatory Relations and Mechanistic Explanation

I have argued that distinct types of representation play dissociable explanatory roles, and that multiple representations are often necessary in explaining a phenomenon such as compensation. I have also claimed, that, in explanatory practice, the relationship of constraint between

representations is not unidirectional. Sometimes data graphs *do* serve constraint and evidence roles. But in other situations it is more accurate to say that a mechanistic representation constrains the search for ERs, as in the Baggs et al. case. Which representation is providing the constraint versus which is representing a new explanatory fact depends both on what is known about the system and what question is being asked. Recognizing this inspires a kind of contextualism about how representation relates to explanation (for a contextualist view of explanation writ large, see Van Fraassen, 1983). This is exactly what we should expect if representational fundamentalism is false—if different representations can play different roles given what and how they represent, then different representations should come to the forefront in particular explanatory contexts. Here, I will attempt to summarize the relative explanatory roles of data graphs and mechanism diagrams. I will only discuss mechanistic diagrams and data graphs here, since the relationship between mechanism diagrams and causal graphs is both complex and slightly outside of the scope of this paper.⁹

In section 2, I outlined some conditions that many mechanists require before a causal relationship is of use in mechanistic explanation—such as stability, invariance, directness, and specificity. These properties, or some combination of them, go into most mechanists' characterizations of what makes something an operation, and in keeping with the distinction between mechanistic diagrams and data graphs, the latter don't represent anything about them. Data graphs are inherently bad at representing these properties, because their role of representing specific quantities in specific conditions renders them unable to generalize in the right way. I thus propose that mechanism diagrams represent relations in the system regarding which a *certain kind* of epistemic work has been done—namely, differentiating the relationships in the system that are invariant, direct, specific (or some subset of these) etc., from those that are not. Since representations of QRs are poor at representing these aspects of the system, mechanism diagrams are developed to convey them. That is, constructing and testing a mechanism diagram or schema involves pursuing a certain explanatory project with regards to the discovered entities and relationships in a mechanism. When the explanatory context shifts, for instance to showing how a phenomenon like compensation is produced by the mechanism, the quantitative patterns represented in data graphs portraying QRs become relevant, resulting in their being represented as ERs.

The foregoing makes good sense of the idea that the relationship of “constraint” between mechanism diagrams and data graphs is neither unidirectional nor constant across all explanatory contexts. If one wants to know whether a part performs a particular operation, and whether that operation is direct, specific, etc., with regards to the other parts affected by it, then one collects data as evidence for those posits (as, e.g., in the case of the Bunker et al. result shown in Figure 2). However, the mechanistic posits can themselves in turn constrain the search for new ERs, such as those of balance and proportionality. Sheredos et al. (2013) and Burnston (2013) have articulated a view of the cognitive role of mechanism diagrams for practicing scientists, on

⁹Mechanism diagrams and causal graphs share their usefulness for representing causal relations in the system being studied. One way in which they significantly differ is in mechanism diagrams' ability to represent spatial relationships. Another is that causal graphs do not say anything specific about *how* the causal influence is enacted. This is perhaps why Baggs et al. say that their network diagrams are “not intended to be biochemically mechanistic” (2009, p. 0574).

which mechanism diagrams guide the search for new hypotheses, and this role fits nicely with the current perspective.

I'd like to conclude by discussing the relationship between the points I've made here and the debates within mechanistic philosophy of science mentioned in the introduction. At several points, those familiar with the mechanist literature may have wondered about the relationship between my discussion of patterns in data graphs and debates surrounding the roles of laws and generalizations in mechanistic explanation. Some argue that operations are metaphysically and explanatorily fundamental in giving mechanistic explanations (e.g., Machamer, 2004); others that it is laws or generalizations (e.g., Leuridan, 2010). I have made a claim about representations and their role in explanation, *not* about the metaphysics of mechanisms. As such, I make no commitment as to the appropriate metaphysical category of ERs. If it turns out that ERs are best construed as generalizations, then the representational arguments made here might provide support for the view that they are often necessary parts of explanation. Note, however, that it would not provide an argument that they are *fundamentally* explanatory—i.e., more fundamentally explanatory than mechanisms. The whole thrust of my argument has been geared towards disavowing any sort of representational fundamentalism, vis-à-vis explanation. At best, the argument could support a view that representations of generalizations and representations of mechanisms are often both necessary in explaining a phenomenon (Tabery, 2004; cf. Weber's "explanatory heteronomy," Weber, 2008). I do not in fact think that ERs are best described as either laws or generalizations, but instead simply as patterns exemplified by the particulars observed. I cannot go into the argument now, however, and undertake it in detail elsewhere (Burnston, in prep).

Finally, I noted in the last section that the relationship of dissociability between different representations is different from both subsumption and independence. This is relevant to the broader discussion in the literature about how mechanistic models relate to other types. While some favor mechanisms, and insist that all other representations are subsidiary to mechanistic models in explanation, objectors to this view often attempt to describe situations in which a different type of model—perhaps a dynamical system or a network model—explains *independently* of mechanistic representations (or, at least, representations intended to functionally decompose the system being studied; Silberstein & Chemero, 2013). I have stressed, however, that explanation is a function of combining different types of representations in different ways depending on the context. As such there is no notion here that data graphs are explanatory "independently" of mechanism diagrams (or vice versa). To the extent that models are representations, then, the current considerations suggest that the idea of looking *either* for fundamentally or independently explanatory models might be a bum steer. If the kinds of considerations I have evinced are right, then explanation is a cooperation between different ways of representing a system, not a competition between representations or models for explanatory primacy. Construing the question in the latter way does not describe the practice of explanation.

5. Conclusion

Mechanistic philosophy of science has made large strides in reconstruing scientific explanation, partially by attending to the actual practice of constructing explanations and the representations that constitute them. I have argued that a full account of explanatory and representational

practice in such fields as mammalian chronobiology requires admitting certain uses of data representations as genuinely explanatory—as representing explanatory relations. Doing so has at least the potential to shift the focus in discussion of mechanistic explanation away from debates about which types of explanations are fundamental, to analysis of the nature of and relationships between different ways of representing biological systems. Given that biology is representationally rich—there are many types of representations used in biology that I have not discussed—such a focus is a potentially fruitful direction for further research within the mechanistic perspective.

Acknowledgements

Part of this research was conducted with support from the National Science Foundation, grant no. 1127640. I would like to heartily thank the other members of the WORKing Group on Diagrams in Science (WORGODS), Adele Abrahamsen, William Bechtel, and Ben Sheredos, for many stimulating discussions of diagrams over the last several years. Thanks also to the editor and several anonymous reviewers for this journal, whose helpful reports prompted considerable improvement. A much earlier version of this paper was presented at the 2013 International Society for the History, Philosophy, and Social Science of Biology conference in Montpellier, France; a slightly newer version was presented at the 2015 Society for Philosophy of Science in Practice in Aarhus, Denmark. I am grateful to audiences at both conferences for helpful discussion.

References

- Baggs, J. E., Price, T. S., DiTacchio, L., Panda, S., Fitzgerald, G. A., & Hogenesch, J. B. (2009). Network features of the mammalian circadian clock. *PLoS Biology*, 7(3), e52-e52. doi: 10.1371/journal.pbio.1000052
- Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 36(2), 421-441.
- Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3), 321-333.
- Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity: Decomposition and localization as scientific research strategies*. Princeton, NJ: Princeton University Press.
- Bunger, M. K., Wilsbacher, L. D., Moran, S. M., Clendenin, C., Radcliffe, L. A., Hogenesch, J. B., . . . Hughes, H. (2000). Mop3 Is an essential component of the master circadian pacemaker in mammals. *Cell*, 103(7), 1009-1017.
- Burnston, D. C. (2013). *Mechanistic Diagrams as Search Organizers*. Paper presented at the The Annual Meeting of the Cognitive Science Society, Berlin.
- Burnston, D. C. (In prep). Patterns in biological explanation.
- Burnston, D. C., Sheredos, B., Abrahamsen, A., & Bechtel, W. (2014). Scientists' use of diagrams in developing mechanistic explanations: A case study from chronobiology. *Pragmatics and Cognition*.
- Craver, C. F. (2006). When mechanistic models explain. *Synthese*, 153(3), 355-376.
- Craver, C. F. (2007). *Explaining the brain*. Oxford: Oxford University Press.

- Darden, L., & Craver, C. (2002). Strategies in the interfield discovery of the mechanism of protein synthesis. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 33(1), 1-28.
- Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, 88(1), 27-51.
- Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science*, 69(S3), S342-S353.
- Goodman, N. (1976). *Languages of art: An approach to a theory of symbols*. Indianapolis, IN: Hackett.
- Green, S. (2013). When one model is not enough: Combining epistemic tools in systems biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(2), 170-180.
- Griesemer, J. R. (1991). Must scientific diagrams be eliminable? The case of path analysis. *Biology & Philosophy*, 6(2), 155-180. doi: 10.1007/BF02426836
- Kafri, R., Levy, M., & Pilpel, Y. (2006). The regulatory utilization of genetic redundancy through responsive backup circuits. *Proceedings of the National Academy of Sciences*, 103(31), 11653-11658.
- Kaplan, D. M., & Craver, C. F. (2011). The explanatory force of dynamical and mathematical models in neuroscience: A mechanistic perspective. *Philosophy of Science*, 78(4), 601-627.
- Leuridan, B. (2010). Can mechanisms really replace laws of nature? *Philosophy of Science*, 77(3), 317-340. doi: 10.1086/652959
- Machamer, P. (2004). Activities and causation: The metaphysics and epistemology of mechanisms. *International Studies in the Philosophy of Science*, 18(1), 27-39.
- Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67, 1-25.
- Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4), 669-688.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.
- Perini, L. (2005). Explanation in two dimensions: Diagrams and biological explanation. *Biology and Philosophy*, 20(2-3), 257-269. doi: 10.1007/s10539-005-2562-y
- Ruse, M. (1990). *Are Pictures Really Necessary? The Case of Sewell Wright's "Adaptive Landscapes"*. Paper presented at the PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association.
- Shah, P., Mayer, R. E., & Hegarty, M. (1999). Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. *Journal of Educational Psychology*, 91(4), 690-702
- Sheredos, B., Burnston, D., Abrahamsen, A., & Bechtel, W. (2013). Why do biologists use so many diagrams? *Philosophy of Science*, 80(5), 931-944.
- Silberstein, M., & Chemero, A. (2013). Constraints on localization and decomposition as explanatory strategies in the biological sciences. *Philosophy of Science*, 80(5), 958-970.
- Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying representational competence with multi-representational displays. *Cognition and Instruction*, 29(1), 123-145.
- Tabery, J. G. (2004). Synthesizing activities and interactions in the concept of a mechanism. *Philosophy of Science*, 71(1), 1-15.
- Van Fraassen, B. C. (1980). *The scientific image*. Oxford: Clarendon.

- Wang, C., Zhang, Z.-M., Xu, C.-X., & Tischkau, S. A. (2014). Interplay between dioxin-mediated signaling and circadian clock: A possible determinant in metabolic homeostasis. *International Journal of Molecular Sciences*, *15*(7), 11700-11712.
- Weber, M. (2008). Causes without mechanisms: Experimental regularities, physical laws, and neuroscientific explanation. *Philosophy of Science*, *75*(5), 995-1007.
- Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology & Philosophy*, *25*(3), 287-318.
- Woodward, J. (2013). Mechanistic explanation: Its scope and limits. *Aristotelian Society Supplementary Volume*, *87*(1), 39-65. doi: 10.1111/j.1467-8349.2013.00219.x
- Ye, R., Selby, C. P., Ozturk, N., Annayev, Y., & Sancar, A. (2011). Biochemical analysis of the canonical model for the mammalian circadian clock. *The Journal of Biological Chemistry*, *286*(29), 25891-25902. doi: 10.1074/jbc.M111.254680
- Zhang, E. E., & Kay, S. A. (2010). Clocks not winding down: Unravelling circadian networks. *Nature Reviews Molecular Cell Biology*, *11*(11), 764-776.