Exporting Causal Knowledge in Evolutionary and Developmental Biology

Sandra D. Mitchell†

In this article I consider the challenges for exporting causal knowledge raised by complex biological systems. In particular, James Woodward's interventionist approach to causality identified three types of stability in causal explanation: invariance, modularity, and insensitivity. I consider an example of robust degeneracy in genetic regulatory networks and knockout experimental practice to pose methodological and conceptual questions for our understanding of causal explanation in biology.

1. Introduction. For some time I have been exploring the consequences of characteristic features of complex evolved biological systems for scientific explanation (Mitchell 2003, 2008). Systems like a honeybee colony, the human brain, or the slime mold *Dictyostelim discoideum* display compositional, dynamical, and evolved complexity. Their behavior depends on multilevel organization, multicomponent causal interactions either within or among levels of organization, as well as evolutionary and other forms of contingency. Each of these features affects the kinds of explanations that will be adequate to account for specific aspects of their behavior.

In this article I will explore the ways in which two particular characters of an evolved complex system might be vexing to the investigation of its causal structure and hence to the explanatory resources we can bring to bear on accounting for system behavior. Those features are modularity and robustness.

Recently, there has been an explosion of interest in modularity among biologists (Schlosser and Wagner 2003; Callebaut and Rasskin-Gutman 2005). Modularity in biological organization has been identified as a feature of complex systems that allows for adaptability (Lewontin 1978) and

†To contact the author, please write to: Department of History and Philosophy of Science, University of Pittsburgh, 1017 Cathedral of Learning, Pittsburgh, PA 15260; e-mail: smitchel@pitt.edu.

Philosophy of Science, 75 (December 2008) pp. 697–706. 0031-8248/2008/7505-0017$10.00

stable development (Simon 1962). The discovery of conserved causal machinery across different subsystems in the developing organism and in widely diverse evolved species and higher taxa has been taken to hold the promise of providing a theoretical unification of evolutionary and developmental biology. However, the individual evolutionary modules or developmental modules do not always, or perhaps even often, coincide and hence the inferences one might draw from biological modularity are variable and context sensitive (Mitchell 2006).

Robustness is the ability of a biological system to maintain normal functioning in the face of internal or external perturbations (see Kitano 2004; Wagner 2005). The capacity for robustness appears to be both ubiquitous and significant for biological systems. But robustness raises challenges for some of our standard notions of causation, that is, the modularity or separability of component causes in generating complex effects and the generality of the causal interactions they display. Robustness may require explanations to shift causal agency to other levels of organization to preserve features that are both explanatory and exportable. But, I will argue, this shift of level is both variable and context sensitive.

What do we require of a causal relation for it to be exportable to contexts outside the one in which it is identified? The regularity of association, the capacity to contribute the same effort in all contexts, the universality of the functional relationship between cause and effect, the natural necessity with which the effect follows the cause have all been at one time identified with what makes the causal relation between variables in one setting explanatory or predictive of what occurs in another. While strong ties between cause and effect make exportation easy, requiring explanation to be only by means of universal, exceptionless laws excludes most of biology from being explanatory (Mitchell 2000). A weaker account of what is required for what constitutes a causal explanation better captures the epistemic practices of biology.

2. Causal Modularity. An interventionist account of causal explanation developed by Woodward (2003) has features that make it an attractive place to start in understanding causation, explanation, and the exportability of causal knowledge in biology. On Woodward's account, what is needed for causation and explanation is not universality but invariance or stability of the causal relationship in various, but not necessarily all, contexts. Three features capture this idea: invariance, modularity, and insensitivity. Invariance describes the relationship between the cause and effect or the dependent and independent variable $F(X,Y)$, such that under certain "ideal interventions" where the value of X changes, the function will describe the resulting value of Y. Hence X explains Y for those ranges

where the functional relationship is invariant under intervention. Modularity refers to a different sort of stability and applies not to causal relations among individual variables, say between X and Y, but to a causal system. Modularity identifies the separability of different causal contributions to an overall effect. In its representational form modularity says of the structural model that correctly represents the causal system that if one intervenes on one equation, the values of the variables that occur in the other equations in the model do not change. The third type of stability that is taken by Woodward to be important for attributions of causality is insensitivity (Woodward 2006). This feature describes the invariance of the causal relation across variations in background, context, or conditions external to system and not directly represented in any of the equations.

Note how each form of stability permits a different type of exportation of a causal relationship observed between X and Y to other situations. If X causes Y then invariance allows the prediction and explanation of the value of Y in a range of cases in which the value of X varies. If modularity holds then if X causes Y and Z causes W and Y and W cause effect E, then intervening on X should not change the relationship between Z and W. This is admittedly an oversimplified account of modularity in structural equations, but it will be sufficient for the argument below. And if the causal relationship between X and Y is relatively insensitive, then varied background conditions will continue to support the explanation of the value of Y by appeal to the value of X.

My concern in this article is with the second type of invariance, that is, modularity. Since the behavior of complex biological systems depends on multiple causes, there is a genuine question about their separability. The worry is about how strong to take the requirement of modularity in order to ascribe causality to a component of a complex causal network. In the past, Hausman and Woodward have claimed that independent disruptability is essential to our understanding of causal contribution: "this is implicit in the way people think about causation . . . this sort of independence is *essential* to the notion of causation. Causation is connected to manipulability and that connection entails that *separate mechanisms are in principle independently disruptable*" (Hausman and Woodward 1999, 550, my emphasis).

In Woodward 2003 (see 48ff.), the claim is less strong, in that modularity is viewed as only one type of causal invariance, that is, formally, the invariance of the causal dependence described in one equation relative to changes in other equations where the modular equations are taken to represent distinct, autonomous, context-insensitive causal mechanisms subject to independent disruptability. I will illustrate below a case where a causal network has elements that appear to behave in nonmodular ways.

What are we to make of such a case? I believe there are three possible responses:

1. We can infer that the elements are NOT causes because they fail to satisfy the condition of modularity.
2. We can redescribe the network in finer or coarser granularity to satisfy the modularity/independence condition.
3. We can infer that modular causes do not exhaust all the types of causality found in nature.

3. Case of the Flexible Genome. What is the causal relationship between genes and phenotypic traits? Undoubtedly this is an extremely difficult question to answer, yet technological developments in genetic engineering continue to provide new tools of intervention. In the early 1980s the technology of gene transfer was developed that permitted direct genetic intervention into a developing embryo. Early genetic engineering did not control where in the genome the new gene would attach, and since position is important to function, they were not good indicators of gene function. With genetic knockout techniques more precise controls became possible (Müller 1999). Scientists could replace, or knock out, a specific gene in a mouse with an inactive or mutated allele. By then letting the mutant mouse develop and then crossbreed, the next generation would issue in double mutants that lacked the targeted gene in both allelic positions. Differences between the normal organism and the double knockout mutant, in particular the presence in one and absence in the other of the expected phenotype, indicated the function of the normal gene.

The logic of these experiments is typical of controlled experiments. That is, if the gene that is knocked out is a component cause of the phenotypic trait, successfully eliminating its contribution should issue in a change in the trait. The change reflects the contribution or function of the knocked-out gene. Many of the results of knockout experiments are difficult to interpret. About 15% of knockout procedures are lethal. Thus the gene may be causally relevant to phenotypic expression, but its specific function cannot be isolated. In other cases, specific phenotypic effects are produced and the difference between the mutant and normal case is attributed to the presence or absence of that gene. However, in up to 30% of double knockouts there is little or no evident phenotypic consequence of knocking out a gene (Ihle 2000; Edelman and Gally 2001). The cases where the knockout produces no substantive phenotypic difference may point to the dynamical plasticity of the genetic network. Robustness due to redundancy or degeneracy will make it difficult to make inferences about the normal causal structure from an intervention or perturbation of the system (Edelman and Gally 2001; Greenspan 2001).

Redundancy in a genetic regulatory system describes a situation where there are multiple copies of a functional gene. Redundancy is widespread in biological systems, though it is somewhat puzzling how it could have evolved. To carry around an extra copy of a functional component incurs a cost. However, there is no correlative adaptive benefit unless it is at least occasionally called upon to operate. It has been argued that there are scenarios in which redundancy is nevertheless evolutionarily stable (Nowak et al. 1997). Degeneracy, as defined by Edelman and Gally (2001), is distinguished from copy redundancy. It refers to an organization where alternative components and structures with distinct functions may nevertheless produce the effect of a targeted component when the component is no longer operative. Degeneracy has been identified in 22 different levels of biological systems, including protein folding, biosynthetic and catalytic pathways, immune response, neural circuitry, and sensory modalities (Edelman and Gally 2001). Redundancy and degeneracy are ways to instantiate robustness: "Robustness is a property that allows a system to maintain its functions despite external and internal perturbations. It is one of the fundamental and ubiquitously observed systems-level phenomena that cannot be understood by looking at the individual components. A system must be robust to function in unpredictable environments using unreliable components" (Kitano 2004, 826).

The absence of a phenotypic change even when all the redundant copies of a single genetic component are knocked out may indicate that the network itself has reorganized. If so, parts of the network that in the normal state would be described by one set of functional relationships respond to the experimental perturbation and change their interactions to produce some product similar to, if not identical to, that of the unperturbed system. Greenspan (2001) offers the following figure (see Figure 1) to display what may be happening in robust causal networks under single node intervention. If *a* is the normal case and *b* is the case where one nodal gene is knocked out, then the it appears that the functions describing the causal contribution of the various components changes in both connectivity in their functional relationships. "The relationships that have been described as pathways are no doubt real, but they need not be invariant. Their relationships are embedded in broader and more plastic networks that can be reconfigured depending on the immediate circumstances" (Greenspan 2001, 386).

How are the anomalous knockout experiments interpreted by Woodward's account of casual explanation? It appears that in degenerate systems like that described by Greenspan there is not independent disruptability of the causal functions describing the relationships between the genetic nodes. Thus, Woodward's condition of modularity is not met. And yet, would we want to conclude that because the normal genetic

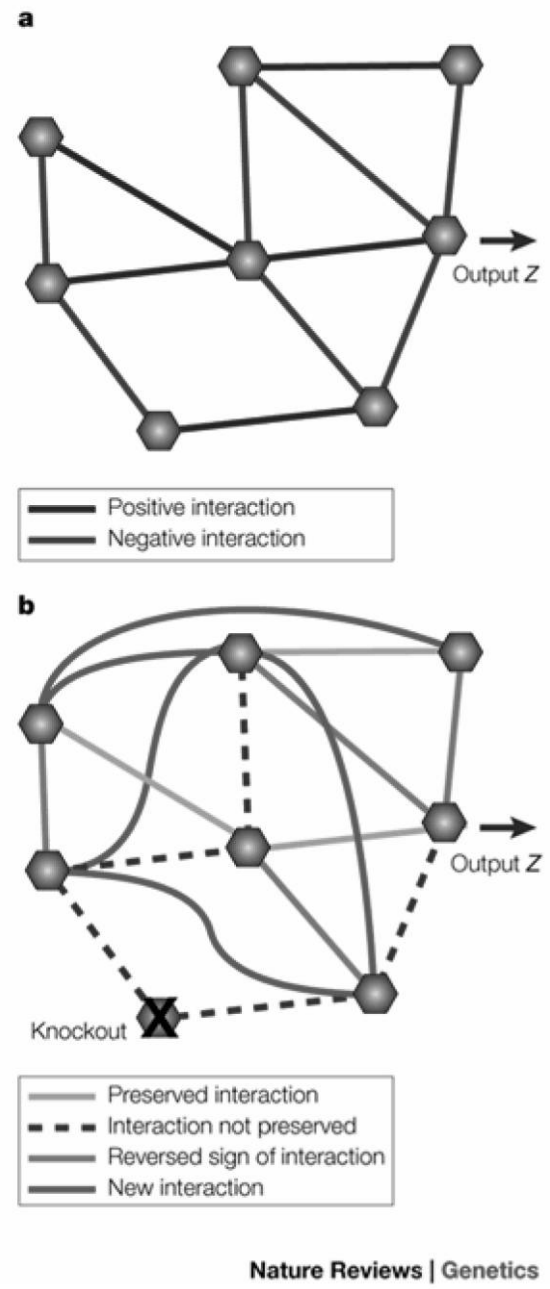

Figure 1. From "The Flexible Genome" (Greenspan 2001).

structure could not be independently disrupted that the knocked-out gene did not causally contribute to the production of the phenotype in the normal organism? Obviously not. While some causally invariant relationships are modular in this strong sense, others are not, and yet I believe we would still want to identify them as both causal and explanatory.

How can this example be brought into compliance with modularity? Instead of claiming that the new causal relations came into being as a result of the intervention on a gene by knocking it out, we might instead say that the network described in Figure 1*a* is not complete. Not all the causal relations are portrayed, at least not all the potential causal relations that could occur between the nodes in different situations. In that case it would no longer be a counterexample, because it would not be an example of a completely described causal network.

This solution misconstrues the character of biological explanation. Many, if not most, explanatory projects in biology are directed not at what might be possible, but rather at what has actually evolved. Given the contingency of the forms and behaviors of biological objects, describing all the physically and chemically possible interactions does not zero in on the actual target of biological explanation. While the actual physical and chemical structures constrain what could have evolved they do not determine what has in fact evolved. Rather that is the result of contingent evolutionary history, replete with undirected randomness. Biologists seek explanations for the actual domain of biological forms and behaviors, and not the set of causal functions that are true of the larger set of the biologically possible forms. The explanandum is the particular process by which some effect, in fact, has been brought about. There are often many pathways by which a specified effect could come about and robustness ensures that if perturbed from its normal state the biological system can recover something close to its normal function. Since what a component of a complex genetic network does is dependent on the context, it may be too sensitive to be reliably discovered by means of comparing cases of its presence and its absence.

This leads to a second solution to the failure of modularity in the case described. Perhaps we are working at the wrong level of organization. If we move up to the network itself, then the interactive and connected set of nodes in Figure 1*a* would be deemed the cause of the trait Z and a second network, represented in Figure 1*b*, would be identified as a different cause of the trait Z. One doesn't detect the causal upshot of, say, striking a match by comparing it to not striking the match while at the same time flicking the lighter. In both cases, a flame appears. Does that tell us that striking the match did not cause the flame in the first case? Because the same phenotypic effect occurs in the normal and double mutant knockout mice, we cannot infer that the knocked-out gene did not causally con-

tribute in the normal genetic network even when the particular gene's functional relationship to the effect in the system in which other genes (and other conditions) are integratively composed is not independently disruptable.

Modularity is the mark of a type of independence from context. The same functional relationship between variables will hold in a given component of the contributing mechanisms whether or not there is a change in a different component. The total effect may change when different components contribute, but the operation of the modular mechanism will not be changed nor change them. In situations where the presence or absence of other contributing factors changes the behavior of a component, and not just the total effect, modularity will fail. But what about the knockout experiments that are successful? In these cases the single perturbation study did reveal a significant causal contribution of a single gene. There is not just one correct level of organization in which to analyze causal invariance and from which to generalize or export to other cases. When single knockout experiments work, when the system is well behaved, that is, it displays invariance, modularity, and insensitivity, then exporting causal knowledge may be relatively straightforward. The "gene for" PKU disease approximates this case where a naturally occurring double mutant at a specific site on chromosome 12 in humans prevents the production of the enzyme phenylalanine hydroxylase. In the normal state the enzyme is produced, in the mutated state it is not. In robust, reorganizing networks, knocking out a single gene will not issue in a change in the phenotypic effect. Genes are sometimes the right level to find modularity, but not always. Wagner has argued that if the behavior of a complex system is linear, then it will always be possible to find a representation such that the effects of the components become independent of the context of the other variables (Wagner 1999). However, if the system is nonlinear, which is often the case in biology, then one "may need to know the state of the entire system with all its state variables to make predictions" (Wagner 1999, 99).

4. Conclusion. Biological systems are complex and contingent. These features present challenges, both methodologically and conceptually, to our ability to export causal knowledge from one case to another. Recognizing that the causal functions discoverable in biology are for the most part not universal has led some in the past to doubt that there could be laws in biology and cast a correlated shadow on biology's explanatory potential. Woodward has provided an alternative account of casual explanation that replaces the demand for universality with invariance, modularity, and insensitivity. His approach is better suited to the contingencies populating the biological world.

I have argued in this article that some forms of complex organization and dynamics present problems even for Woodward's more forgiving approach. In particular, when a system is compositionally integrative in the Bechtel and Richardson (1993) sense, in which the causal properties of the parts are themselves nonlinearly dependent on the whole network, then modularity or independent disruptability will not be satisfied. If we read the modularity requirement strongly, as defining what it is for a component to make a causal contribution to the system behavior, then we are left with the uncomfortable conclusion that a component may have a causal role in an unperturbed system, but loses its causal status if when perturbed the system reorganizes. One way out of this situation is to let the individuation of causes be constrained by modularity and accept that the appropriate level of organization to find causal explanations is itself context dependent. A third alternative is to take modularity to be a feature present in some but not required of all causal systems. Some complex structures harbor nonmodular, context-sensitive actual causes that can explain its behavior. On this view, exporting knowledge from one system to another will requires more than generalization from observed structure and simple instantiation in a new context.

REFERENCES

Bechtel, William, and Robert C. Richardson (1993), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Princeton, NJ: Princeton University Press.

Callebaut, Werner, and Diego Rasskin-Gutman, eds. (2005), *Modularity: Understanding the Development and Evolution of Natural Complex Systems*. Cambridge, MA: MIT Press.

Edelman, G. M., and Joseph A. Gally (2001), "Degeneracy and Complexity in Biological Systems", *PNAS* 98 (24): 13763–13768.

Greenspan, Ralph J. (2001), "The Flexible Genome", *Nature Reviews: Genetics* 2: 383–387.

Hausman, Daniel, and James Woodward (1999), "Independence, Invariance and the Causal Markov Condition", *British Journal of the Philosophy of Science* 50: 521–583.

Ihle, James N. (2000), "The Challenges of Translating Knockout Phenotypes into Gene Function", *Cell* 102: 131–134.

Kitano, Hiroaki (2004), "Biological Robustness", *Nature Genetics* 5: 826–837.

Lewontin, Richard (1978), "Adaptation", *Scientific American* 239 (3): 156–169.

Mitchell, Sandra D. (2000), "Dimensions of Scientific Law", *Philosophy of Science* 67: 242–265.

——— (2003), *Biological Complexity and Integrative Pluralism*. Cambridge: Cambridge University Press.

——— (2006), "Modularity—More than a Buzzword?" *Biological Theory* 1: 98–101

——— (2008), "Explaining Complex Behavior", in Kenneth Kendler and Josef Parnas (eds.), *Philosophical Issues in Psychiatry: Natural Kinds, Mental Taxonomy and Causation.* Baltimore: Johns Hopkins University Press, 41–64..

Müller, Ulrike (1999), "Ten Years of Gene Targeting: Targeted Mouse Mutants; From Vector Design to Phenotype Analysis", *Mechanisms of Development* 82: 3–21.

Nowak M. A., M. C. Boerlijst, J. Cooke, and J. Maynard Smith (1997), "Evolution of Genetic Redundancy", *Nature* 388: 167–171.

Schlosser, Gerhard, and Gunter P. Wagner, eds. (2003), *Modularity in Development and Evolution*. Chicago: University of Chicago Press.

Simon, Herbert (1962), "The Sciences of the Artificial", *Proceedings of the American Philosophical Society* 106: 467–482.
Wagner, Andreas (1999), "Causality in Complex Systems", *Biology and Philosophy* 14: 83–101.
——— (2005), *Robustness and Evolvability in Living Systems*. Princeton, NJ: Princeton University Press
Woodward, James (2003), *Making Things Happen.* Oxford: Oxford University Press.
——— (2006), "Sensitive and Insensitive Causation", *Philosophical Review* 115: 1–50.