Why Genic and Multilevel Selection Theories Are Here to Stay*

C. Kenneth Waters†‡

I clarify the difference between pluralist and monist interpretations of levels of selection disputes. Lloyd has challenged my claim that a plurality of models correctly accounts for situations such as maintenance of the sickle-cell trait, and I revisit this example to show that competing theories don't disagree about the existence of 'high-level' or 'low-level' causes; rather, they parse these causes differently. Applying Woodward's theory of causation, I analyze Sober's distinction between 'selection of' versus 'selection for'. My analysis shows that this distinction separates true causes from pseudocauses, but it also reveals that the distinction is irrelevant to the levels debate; it makes no sense to say true causes are at higher levels and not lower levels. The levels debate is not about separating real causes from pseudocauses; it's about finding useful ways to parse and disentangle causes.

1. Introduction. Everyone agrees there is a *plurality* of competing theories that model selection at different levels of biological organization. For example, some theories model evolutionary processes in terms of selection for individuals, and other theories generate models that represent the same processes in terms of selection for groups. An assumption many scientists and philosophers take for granted is that alternative model-types disagree about the facts. In levels of selection disputes, antagonists often assume

*Received April 2005.

†To contact the author, please write to: Minnesota Center for Philosophy of Science, Department of Philosophy, University of Minnesota, Minneapolis, MN 55408; e-mail: ckwaters@umn.edu.

‡I thank Mark Borello, Chris Hitchcock, Ben Kerr, Richard Lewontin, Helen Longino, Roberta Millstein, Bart Moffatt, Katie Plaisance, Alex Rosenberg, Elliott Sober, and Jim Woodward for reading and commenting on a draft of this paper. Participants in the Minnesota Center for Philosophy of Science pluralism project—and especially Stephen Kellert and Helen Longino—have helped me think about pluralism more clearly. Discussions with Jim Woodward and Chris Hitchcock on causation were very helpful. For a fuller account of scientific pluralism, see Kellert, Longino, and Waters forthcoming. I would also like to thank Lisa Lloyd for a stimulating correspondence and paper.

Philosophy of Science, 72 (April 2005) pp. 311–333. 0031-8248/2005/7202-0002$10.00

that for a given selection process, there must a model that identifies the *real* level or levels of selection in that process. No one denies that in many cases competing models disagree about important matters of fact, but there is no *a priori* reason to assume this is true in every case.

In some cases, alternative theories that model the same situations differently might not disagree about matters of fact. A model couched in terms of individual selection and a different model couched in terms of group selection might both represent an evolutionary situation truthfully. Of course, many models misrepresent evolutionary processes, and biologists rightly reject them on empirical grounds. But this leaves open an important question: Are there cases in the levels dispute where models competing for attention are not disagreeing about facts? I, and a number of philosophers and biologists, have argued that the answer to this question is *yes*.[1] Lisa Lloyd challenges this answer. Her spirited critique brings out important issues and raises a second question: If some situations can be individually accounted for by a plurality of true models, as pluralists contend, could scientists nevertheless have rational grounds for choosing some true models over others? I believe pragmatic reasons can exist for preferring one true model over another. Lloyd assumes the opposite.

Pluralism about levels of selection is epistemologically significant because it calls for a less strident form of scientific realism (at least with respect to evolutionary biology). The best models are, the evidence indicates, true representations. Scientific realism holds. But, a true model is not necessarily *uniquely* true. In some evolutionary situations, no model provides the single right way to represent the causal processes because there is no single right way; instead, there is a plurality of right ways each associated with—to use Helen Longino's (forthcoming) apt phrase—a *different parsing of causes*. Scientists and philosophers who believe the proper aim of science is to discover the comprehensive theory that provides the single, correct way to represent the causal structure of the world (or part of the world) need to temper their realism.

In this paper, I develop and defend pluralism about evolutionary biology. In Section 2 and Section 3, I clarify the difference between pluralism and monism. Lloyd challenges my claim that a plurality of models can correctly account for situations, such as the maintenance of the sickle-cell trait, and I defend my claim in Section 4. Using the sickle-cell example, I show that competing theories don't disagree about the existence of 'high-level' or 'low-level' causes; rather, they parse the causes differently. In Section 5, I explain that pluralism acknowledges the existence of genuine empirical issues in levels debates, but holds that questions about these

1. E.g., Cassidy 1978; Waters 1985, 1991; Sterelny and Kitcher 1988; Dugatkin and Reeve 1994; Kerr and Godfrey-Smith 2002.

issues are framed differently depending on the way theorists parse the causes. In Section 6, I argue that not all parsings suit particular epistemic interests equally well, and that the merits of competing parsings can be assessed on pragmatic grounds. In Section 7, I use Jim Woodward's recent work on causation to analyze Sober's distinction between 'selection of' and 'selection for'. The result concurs with Robert Wilson's (2003) claim: The distinction many assumed was central to the levels debate is irrelevant. I go further and argue that it makes no sense to say the causes are at higher levels and not at lower levels. I conclude by offering a pluralist interpretation of recent work on levels of selection and making a suggestion for future philosophical research.

2. The Difference between Pluralism and Monism.[2] Lloyd refers to the position advanced in Kim Sterelny and Philip Kitcher's "Return of the Gene" (1988), our jointly authored response to Sober (1990)—called KSW—and my dissertation and subsequent paper, "Tempered Realism about the Force of Selection" (1991) as 'genic pluralism'. I understand why Lloyd suspects that Sterelny and Kitcher (1988) wanted to have it both ways. On the one hand, they seem to argue for a strong version of pluralism according to which the genic and group theories are equivalent, while on the other hand, they seemed to argue that the genic theory is superior. I assume she uses the term *genic pluralism* to emphasize this contradiction. Although Lloyd acknowledges important differences between the position I advanced in "Tempered Realism" and conclusions Sterelny and Kitcher reached in "Return of the Gene," I think her lumping our views together under the oxymoronic label 'genic pluralism' is misleading. Tempered realism is a form of pluralism; it is not a Trojan horse for genic selectionism.

At this point, it would help to define terms. It is useful to have separate labels for the theory that models evolutionary situations in terms of selection for individual alleles, and for the claim that this theory, and this theory alone, provides the correct account of all situations involving natural selection. I call the first, the *genic theory*, and the second, *genic monism*. In parallel fashion, I distinguish between the *multilevel theory of selection* and *multilevel monism*. The multilevel theory models evolutionary situations in terms of selection processes that can occur at one or more levels. This theory might model one situation as selection for individuals, model another situation as selection for groups, and model a third situation as selection for both individuals and groups. Multilevel monism is not an oxymoron. One might hold for a given situation, there is a uniquely

2. For a fuller account of scientific pluralism, see Kellert, Longino, and Waters forthcoming.

correct model of any given selection situation that identifies the true level(s) of selection in that situation. According to multilevel monism, in one situation the *real* target of selection might be genes, in another situation the *real* target could be individual organisms, and in yet another situation, there might be genuine targets at multiple levels. The multilevel monist insists that for any given situation, there is a single correct answer to the question, What is the *real* level(s) of selection?

Pluralism about levels of selection is the view that monisms about the targets of selection are false. It says that in *some* situations, an evolutionary process can be modeled correctly as selection for units at one level, and alternatively modeled correctly as selection for units at another level (or multiple levels). This is the view I advanced in "Tempered Realism." My basic thesis was, and still is, that the causes in evolutionary situations are sufficiently complicated that they can be described truthfully in multiple ways. I argued that the sickle-cell case provided an example of such a situation; the causal processes that maintain the sickle-cell allele can be truthfully modeled as selection for alleles (via the genic theory), and they can be truthfully modeled as selection for allele pairs (via the diploid-genotypic model). Tempered realism is not genic pluralism; it is pluralism plain and simple.[3]

Although I was a consistent pluralist in "Tempered Realism," Lloyd found passages in KSW that betray a preference for the genic view. The most incriminating passage, which Lloyd astutely identified, but graciously refrained from quoting, stated: "All selective episodes (or, perhaps, almost all) can be interpreted in terms of genic selection. That is an important fact about natural selection" (KSW 1990, 160). I believe the first sentence quoted here is true, but I hereby retract the second sentence. In addition to confessing my careless slip into monism, I wish to point out that pluralism about levels of selection does not depend on the first of the quoted claims any more than it would depend on the claim that all selective episodes can be modeled as selection for groups. Pluralism depends only on the idea that in at least some situations at issue, the evolutionary process can be modeled accurately by alternative theories.

It is useful to have a label for the evolutionary situations that pluralists claim can be represented by a *p*lurality of *e*mpirically *r*easonable *m*odels,

3. My understanding of pluralism is quite different than Robert Wilson's (2003). What Wilson calls 'unit pluralism' is multilevel monism. What he calls 'model pluralism' is a contradictory fusion of two opposing views: pluralism and fundamentalism (see Kellert, Longino, and Waters forthcoming). Hence, my defense of pluralism has more in common with Wilson's criticisms of 'pluralism' than it may appear. Nevertheless, as I explain in Section 6, there are significant differences in our approaches and conclusions.

and I call them *PERM situations*. As I explain in the next section, one could be a pluralist with respect to the question of whether the plurality of models for a specified PERM situation is empirically sufficient in a bookkeeping sense, and a monist with respect to whether the models get something deeper (e.g., the causality) right. The position I'm defending in this paper is pluralism on both bookkeeping and causal accounts.

One might wonder, if my real aim in "Tempered Realism" was to advocate pluralism, why did I devote so much argumentation to defending genic models of selection? Shouldn't a pluralist be evenhanded? Of course, in one sense, a pluralist should not play favorites. But, advancing pluralism to an audience of monists requires showing that some applications of models (which monists reject) can, contrary to what monists believe, accurately represent some of the situations at issue. It was my impression that among philosophers of science, genic selection models needed to be defended because William Wimsatt (1980), Robert Brandon (1982), Elliot Sober (1984), Lloyd ([1988] 1994), and others had convinced potential readers that the genic theory was hopeless. If I had been writing "Tempered Realism" for *Science* instead of *Philosophy of Science*, and using Maynard Smith (1976) as a foil rather than Sober and Lewontin (1982), then I would have emphasized that group selection models can truthfully represent some PERM situations. Perhaps biologists would have smeared me with the label 'group pluralist'.

I reject the label 'genic pluralism' and the implication that tempered realism represents some form of genic selectionism or a weaker form of pluralism than that advocated by Lloyd. In fact, it is not clear to me whether Lloyd is a pluralist at all. She seems to be defending multilevel monism.

3. A Different Look at the Basics. Although it is not clear where Lloyd stands on the epistemological issues that motivated me to join the levels debate, prominent advocates of the kind of theory she champions (the multilevel selection theory) are pluralists of sorts. Sober, for example, has consistently maintained that there is a plurality of alternative models that are empirically adequate in the sense that they can do the 'bookkeeping'. He has acknowledged that genic models can accurately trace changing frequencies and correctly predict what multilevel selection models predict. With respect to bookkeeping, Sober has always been a pluralist.

Sober has recently taken a deeper turn towards pluralism, so it is important to distinguish between the position he advanced with Lewontin (1982) and in *The Nature of Selection* (1984), from the position he has championed with Wilson in *Unto Others* (1998). Although Sober's position and argumentation have shifted in important ways, he has been consistent on a fundamental distinction between success in bookkeeping and a deeper

kind of success (see also Wimsatt 1980). Success in the bookkeeping sense involves correctly tracing and predicting changing frequencies. The deeper kind of success involves identifying the causes that change those frequencies. Some participants in the blossoming corner of population biology, which Lloyd favorably cites, are (like the early Sober) pluralists with respect to bookkeeping and monists with respect to something deeper. For example, John Maynard Smith made this kind of distinction when he admitted that "equivalent mathematical descriptions are often, although not always, possible" but added that he found "the gene-centered approach both mathematically simpler and causally more appropriate" (Maynard Smith 2002, 524).

I suspect that if empirical success in the bookkeeping sense were the only issue, we would all be pluralists. Hence, for the remainder of this paper I will use *monism* and *pluralism* to designate positions about whether a plurality of models in PERM cases represents the causality truthfully.

The issues of serious contention regarding pluralism, then, involve whether a plurality of models aimed at different levels get the causality right, not whether a plurality succeeds at bookkeeping. What is the difference between successful bookkeeping and getting the causality right? David Hull (1980) made an important contribution by clarifying a technical distinction between 'interactors' and 'replicators' in selection processes. But, what is the difference between identifying a pseudointeractor that merely provides a basis for bookkeeping, and a genuine interactor that is a causal target of selection? Sober offered a clear answer to this question: Good bookkeeping only requires tracing the selection *of* objects. But getting the causality right requires identifying the properties that are selected *for*. He used a toy to illustrate the distinction between *selection of* and *selection for*.

The toy is a clear plastic cylinder with several different levels and colored marbles. When the cylinder is held right side up, the marbles are all on the bottom level of the cylinder. When the toy is turned upside down and shaken, marbles descend the levels of the cylinder by falling through circular holes in the floor of each level. When marbles descend no further, some marbles have come to rest at lower levels than others. Green marbles descend to the bottom. Yellow marbles rest one level up from the bottom, and so on. The marbles are of different sizes: marbles that reach the bottom level are smallest (and are green); marbles that rest at the next level up are of the next-to-smallest size (and are yellow); and so on. The circular holes in the floors of each level are also of different sizes. The hole in the floor leading to the bottom level has the smallest circumference; the holes in the levels above have increasingly larger circumferences.

The toy is a selection device. It selects green marbles to descend to the lowest level, yellow marbles to descend to the next level up, and so on.

One can model this selection in terms of color and derive conclusions about size (e.g., spherical circumference) from the correlation between the colors and sizes of marbles. This model, I'll call it M_C, gets the bookkeeping right. One can use M_C to trace descents and make predictions about where marbles of different colors (or *derivatively,* where marbles of different circumferences) will rest when the device is turned one way or the other. But Sober claims that M_C represents the causality wrong. Colored marbles are selected by the selection device, but color is not the causal target of selection. The causal target is circumference. I call the model based on selection for spherical circumference, M_{SC}. In Sober's useful terminology, there is only *selection of* colored objects whereas there is *selection for* spherical circumference. The model that gets the causality right is M_{SC}, not M_C. Partisans might defend M_C on the basis that one can use this model to derive information about selection of spherical circumference, but Sober could respond that the empirical success of this *derivation* does not show that the model gets the causality right. I agree. Both M_C and M_{SC} are empirically adequate in the sense that they get the bookkeeping right, but only M_{SC} identifies real causes.

Where I disagreed with Sober's earlier position[4] is on the question of whether the difference between M_C and M_{SC} illustrates the difference between, say, George Williams' (1966) genic account of the selective maintenance of the sickle-cell trait, and the conventional diploid-genotypic model of the same process. To clarify the point of contention, consider a third model of the toy's selection process, one based on the selection for spherical diameter (rather than spherical circumference or color). On the basis of this model, one can derive information about the selection of color (because of the correlation between spherical diameter and color), and one can derive information about the selection of spherical circumference (because of the equality $c = \pi d$). Consider the plurality of three models: M_C, M_{SC}, and M_{SD}. All three are empirically adequate in the sense that each gets the bookkeeping right. But M_C clearly misrepresents the causal process. What should we say about the difference between M_{SC} and M_{SD}? Does one represent the causes of selection truthfully and the other misrepresent the causes? Of course not. It is not as if the real target of selection is circumference, rather than diameter, or vice versa. M_{SC} and M_{SD} both model the causal situation accurately. In some contexts they might compete for attention, but it is not the case that one is true and the other is false. In some contexts there might be good pragmatic reasons to favor one model over the other (e.g., it might be easier to measure circumference than diameter), but this doesn't imply that one is seman-

4. Cf. Waters 1985, 1991, and Sober and Lewontin 1982. Lloyd (2005) now seems to disagree with me.

tically derivative of the other. I contend that the difference between the genic and diploid-genotypic models of the sickle-cell situation, to take a concrete example, is more akin to the difference between M_{SC} and M_{SD} than to the difference between M_{SC} and M_{C}.

4. How Alternative Theories of Selection Parse Causes Differently. Selection models draw conceptual divisions between environments and selected domains. On one end of two extremes, multilevel theories draw this division between conspecific groups of organisms and everything outside the groups. On the other end, genic theories of Williams' style (1966) draw the division between an individual allele and everything outside the individual allele. Which entities' properties are modeled as being selected for sometimes depends in part on where biologists draw this division. This is one way evolutionary theorists parse causes.

I illustrated this point in "Tempered Realism" by considering the maintenance of the sickle-cell trait. The conventional population model of this situation is the diploid-genotypic model, but Williams (1966) argued that it is correctly accounted for in terms of genic selection. Monists countered that the correct *causal* account is provided by the diploid-genotypic model, not the genic model (e.g., Sober and Lewontin 1982). I argued that both models described the causal processes accurately. Since Lloyd challenges my argument, it is necessary to review the case. I start with the genic model, not because it is uniquely right or better, but because it is the one that Lloyd attacks. Next, I describe the diploid-genotypic model that Lloyd associates with the multilevel theory. Then, I defend both models and show where Lloyd's critique goes astray.

4.1. The Genic Model of the Sickle-Cell Situation. Williams' genic theory draws the environmental divide between an individual allele and everything outside the individual allele.[5] Hence, if we assume there are two allele-types at the locus (designated below as *S* and *A*), then selection is occurring in two different genic environments, and there are two distinct selection processes.

One selection process occurs in $environment_S$, where there is an *S* allele at the corresponding locus. In $environment_S$, the form of hemoglobin causing the sickle-cell condition (which I call *S-hemoglobin*) is synthesized because the environmental allele is *S* and the presence of this allele causes the synthesis of *S*-hemoglobin.[6] In $environment_S$, *A* is viable because it leads to the synthesis of regular hemoglobin molecules and this molecule

5. Dawkins' conception of the genic environment is different than Williams' (see Waters 1991). Here, I deal with Williams' version only.

6. I analyze the notion of gene causation in Waters 2000, forthcoming.

participates in a host of processes in environment$_S$ that leads to A's survival and replication. The S allele is not viable in environment$_S$ because it causes the production of S-hemoglobin, and in environment$_S$ the synthesis of more S-hemoglobin leads to this allele's demise. Hence, there is selection *for* A and against S in environment$_S$. In the genic model

$$W_A^S > W_S^S,$$

where W_A^S is the selection coefficient for allele A in environment$_S$, and W_S^S is the selection coefficient for allele S in environment$_S$. In this model, A and S are selective interactors with respect to environment$_S$. Notice that explaining the selective advantage of A in environment$_S$ draws upon information about causal processes occurring in environment$_S$. Although application of this model depends on environmental information, the causal processes occurring in the genic environment are not modeled in the genic theory as selection processes. The genic theory parses the causes so that selective interactions are between the allele and its genic environment.

Parallel considerations can be made with respect to the other genic environment, environment$_A$. Explaining why S is selected for, relative to A, in environment$_A$ requires taking into account causal processes in the genic environment involving the synthesis of regular hemoglobin, malaria viruses, mosquitoes, and human living conditions. The genic model does not deny the importance of such causal processes. It simply parses the causes such that these processes are modeled as processes within the environment, not as processes between selective interactors and their environment. In the genic model

$$W_S^A > W_A^A,$$

where W_S^A is the selection coefficient for allele S in environment$_A$, and W_A^A is the selection coefficient for allele A in environment$_A$.

Since A alleles participate in two separate selection processes, calculating the overall change in the number of alleles in a human population requires adding the changes of allele number in each environment. Williams does the math by taking a weighted average of the selection coefficients. As Williams explains:

> If the gene associated with anemia and malarial resistance is designated S, its selection coefficient in the genetic environment S [environment$_S$] would be very different from its coefficient in environment S' [environment$_A$]. Its effective (mean) coefficient would be the mean for these two environments, weighted by the frequencies of the environments. (Williams 1966, 60)

Although this is not the conventional way to model the situation, the

kind of 'spatially-dependent' modeling Williams is using here is part of conventional population genetics. Following Williams, we can designate the effective selection coefficients as follows:

$$\overline{W_A} = P^A W_A^A + P^S W_A^S,$$

where P^A is the relative frequency of environment$_A$, and P^S is the relative frequency of environment$_S$. Notice that $\overline{W_A}$ is not a basic causal parameter in this model; rather, it is a mathematical device for calculating the net effects of the basic parameters. The basic parameters are selection coefficients W_A^A and W_A^S.

4.2. The Diploid-Genotypic Model of the Sickle-Cell Situation. The diploid-genotypic theory draws the environmental divide between diploid genotypes (or perhaps organisms) and their ecological environments. Parsing the causes this way yields a model in which there is only one environment (instead of two) and three types of selective interactors (instead of two). There is a selection coefficient for each interactor type:

> W_{SS} is the selection coefficient for diploid genotype *SS* in the diploid environment.
> W_{SA} is the selection coefficient for diploid genotype *SA* in the diploid environment.
> W_{AA} is the selection coefficient for diploid genotype *AA* in the diploid environment.

The relative values of the selection coefficients are

$$W_{SA} > W_{AA} > W_{SS}.$$

Explaining why *SA* is selected for, relative to *AA*, requires taking into account causal processes in the ecological environment involving viruses, mosquitoes, and human living conditions. The diploid model does not deny the importance of such causal processes. I repeat the point I made with respect to the genic model: The diploid-genotypic model parses the causes such that these processes are modeled as processes within the environment, not as processes between an interactor and its environment.

Explaining why *SA* is selected for, relative to *SS*, in the diploid environment also invokes information about alleles and the syntheses of different forms of hemoglobin molecules. The diploid model does not deny the importance of these causal processes. It just parses the causes such that these processes are modeled as occurring within an interactor, rather than between an interactor and its environment.

Calculating the change in the relative number of diploid genotypes in the environment also depends on a mathematical device that takes a weighted average of the diploid selection coefficients. This mathematical

device does not represent a basic parameter in the model. The basic parameters are the selection coefficients W_{SS}, W_{SA}, and W_{AA}.

4.3. Where Lloyd's Critique of the Genic Model Goes Astray. Lloyd challenges pluralism by attacking the genic model of the sickle-cell situation. Recall that pluralism does not depend on the claim that the genic model is superior to the diploid model, or even that applications of the two models are equivalent. What it depends on is the idea that there are PERM situations that can be represented by more than one causal model. Pluralists have argued that the sickle-cell situation is a PERM situation, and that the underlying causes can be correctly represented by both genic and diploid models. Lloyd claims the genic model is 'derivative' because the value of W_S^A is allegedly derivative of the value of W_{SA}. In "Tempered Realism," I pointed out that W_S^A and W_{SA} may have the same numerical values, but they are different parameters. Lloyd counters:

> If the parameters are semantically distinct and you must use the higher-level information, then the pluralists' [i.e., genic selectionists'] models are parasitic, derivative, and hence, not independent. We still need the information about heterozygote fitnesses, and we need to get it the same way—by looking for interactors in a selection process. (2005, 295)

Lloyd is right that application of the genic model requires 'higher-level' (environmental) information, that is, information about what molecules are synthesized, because of the allele on the corresponding locus, the prevalence of malaria and mosquitoes, and the nature of human living conditions in environment$_S$ and environment$_A$. But, she is mistaken in assuming that this information must be modeled in terms of selection for diploid interactors. In effect, she is drawing an environmental divide where the diploid theory does, rather than where the genic theory does.

The pluralist view isn't that there are two adequate models, one of which contains only information about the causal interaction of alleles and their proximate environments, and the other of which contains only information about the causal interaction of diploid genotypes and their proximate environment. The pluralist view is that applications of both models draw upon information 'at many levels' in the complex biological situation. It is not as if the diploid model owns the information above the alleles and the genic model owns the information below the diploid genotype. The models represent the same causality, but they do so differently. Of course, the genic model is going to include information about causal interactions in the environment. The fact that it depends on such information doesn't make it derivative of the diploid-genotypic model.

Lloyd extends her critique of the sickle-cell cases to genic models of

other contested situations as well. She acknowledges that genic models partition genic environments and differentiate between different selection processes in different genic environments. But she asks: "How does Waters know that interactions at the group or organismic level will have an effect on genic fitness?" (Lloyd 2005, 296). My answer: Genic selection theoreticians learn that a genic environment needs to be partitioned in a model, if there are different selective pressures on the allele type in different spaces of the genic environment.

Lloyd loads her questions by using the word 'interaction'. In the *generic sense*, interaction is any sort of interplay. But in the *technical sense* of selection theory, interaction means *selective interaction* between entities in a selected domain and their environment. The genic theory does not model the causal interplay to which Lloyd refers as selective interaction between a selective interactor and its environment. The genic theory acknowledges that there is causal interplay (interaction in the generic sense) within genic environments, but this causal interplay does not need to be modeled as selective interaction. Hence, when Lloyd writes "by looking for interactors in a selection process," (2005, 295) a genic theorist would say *by looking for causal interplay contained within genic environments.*

Lloyd's criticisms often sound convincing because they equivocate on the meaning of *interactor* and *interaction*, and because they implicitly draw environmental divides where the genic theory does not. Lloyd says, "This whole procedure of determining which level of allelic environment needs to be included [i.e., of partitioning the genic environment and identifying separate selection processes] looks suspiciously like those used to determine whether something is functioning as a hierarchical interactor" (2005, 297). I reply that the procedure looks similar when Lloyd reconceives it by redrawing the environmental divide at the diploid level and by trading on the technical meaning of selective interactor and generic meaning of causal interaction.

I am not arguing for genic selectionism; I am instead arguing against monism. If I were responding to a genic monist, I would argue in a parallel fashion. If a genic monist claimed that the diploid-genotypic model is redescribing information about alleles interacting in their genic environments, I would insist that the fact that diploid-genotypic models depend on information about causal interactions of alleles does not indicate that they are "parasitic, derivative, and hence, not independent of the genic theory" (Lloyd 2005, 295). Suppose the partisan argued:

> Lloyd's diploid-genotypic account of why the *SA* genotype is fitter than the *SS* genotype depends on lower-level information about interactions involving alleles and the syntheses of hemoglobin molecules. Where does she get such information? This information is

> obtained by looking for genic interactors in different genic environments.

This criticism of Lloyd's diploid model would be patently unfair. The genic monist would be correct in saying that application of the diploid model depends on 'lower-level' information, but wrong to assume that this information must be modeled in terms of selection for lower-level interactors. I would respond that Lloyd has a different way to parse the causes and that her parsing can yield a model of the complicated biological situation that is just as truthful as the genic model.

5. Two Different Kinds of Issues in the Units of Selection Debate. Lloyd gets lots of mileage quoting rhetorical flourishes in KSW, to the effect that "quibbling about the real unit of selection" (quoted in Lloyd 2005, 288) is a waste of time because it is "an exercise in muddled metaphysics" (quoted in Lloyd 2005, 290). Such remarks might make it sound as if we thought there were no empirical issues in the levels debates. This is misleading. Pluralism does *not* deny the existence of genuine, empirical issues. What it denies is the metaphysical interpretation that monists read into them.

The *t*-allele situation provides a clear example of a PERM situation in which important issues are at stake. (See Waters 1991 for details.) For genic theorists, the issue is whether the genic environment of the *t*-allele is heterogeneous with one selection process occurring in spaces where *t*-alleles are contained within female mice in groups with all sterile males, and another selection process occurring in spaces where *t*-alleles are not contained within such females. For multilevel theorists, the issue is whether there is selection against groups of females in demes containing all sterile males. In a way, it seems to be the same issue. This makes sense because the questions concern the same causal situation. When the causes in this situation are parsed one way, the issue emerges as a question about whether there are multiple selection processes in a heterogeneous environment. When the causes are parsed differently, the issue emerges as a question about whether group selection is occurring. In "Tempered Realism," I stressed that the empirical issues are as important to the genic theorist as they are to the multilevel theorist. The rhetorical flourishes were not intended to impugn these substantive questions.

Sober and Wilson insist that the following is a genuine question: "Can traits evolve by benefiting whole groups, despite being selectively neutral or disadvantageous within groups?" (2002, 530). I agree. But, we should keep in mind (as Sober and Wilson point out) that this question is framed within the multilevel theory. It is a question that *seems* to fall between the cracks of the genic framework. The genic theorist does not ask about

the causal situation by framing her question in terms of group selection; instead she asks her questions in terms of whether there are multiple selection processes in a heterogeneous genic environment. The exercise of answering such questions framed within the multilevel (or genic) framework is empirically meaningful. The exercise becomes muddled only when monists ask, with the exclusive sense of 'or': But is it *really* group selection in the ecological environment or multiple selection processes in a heterogeneous genic environment?

Some cases in dispute (e.g., the *t*-allele case) do involve important empirical issues, but others, such as the sickle-cell case, do not. Why do monists care about the sickle-cell case? Monists seem wedded to the Platonic ideal that science aims to cut nature at its joints. It appears that some philosophers joined the levels debate to clarify conceptual issues with the expectation that biologists could then use empirical methods to identify the real joints of natural selection. The pluralist view, on the other hand, is that there are multiple ways to divide nature, and that some questions are empirically meaningful only with respect to particular causal parsings.

6. The Grounds for Choosing among Competing Models in PERM Cases Are Pragmatic. Scientific models provide partial descriptions. They highlight some features and obscure others. In simple situations, models may highlight all causally relevant features and obscure the rest. In complicated situations, however, such as biological development and evolution, models typically make some causal features salient by obscuring others. This doesn't render models false, but depending upon scientists' interests, it makes them more or less suitable, pragmatically speaking. In some cases, different interests might favor different model-types; hence, there could be pragmatic reasons for retaining a plurality (see Kellert, Longino, and Waters forthcoming).

Readers might wonder whether this is consistent with the claim I made in KSW to the effect that alternative models in PERM situations are representationally equivalent. Lloyd rightly criticizes KSW for making such claims. Kerr and Godfrey-Smith (2002) have offered the clearest account of the relationship between contextual models (e.g., the genic model of heterozygotic superiority) and group models (e.g., the diploid-genotypic model of heterozygotic superiority), and they prove that the models are mathematically equivalent. But Lloyd is absolutely correct in saying that mathematically equivalent models are not necessarily representationally equivalent.[7] The mathematics does not exhaust the infor-

7. For a clear exposition of this point, see Wilson 2003.

mation of *applications* of these models (as I read Kerr and Godfrey-Smith, they wouldn't disagree). I retract all claims about representational equivalence in KSW.[8]

Lloyd suggests that pluralism implies that there can be no rational grounds for choosing among genic and higher-level models. I disagree. There can be pragmatic reasons for preferring one true model over another. I remained neutral on pragmatic issues in "Tempered Realism" because drawing conclusions about pragmatic advantages and disadvantages would require a different kind of analysis. But, since Lloyd's critique brings up pragmatic issues as well as semantic ones, I admit my bias: The genic theory has a number of pragmatic disadvantages. It often requires cumbersome conceptions of environments, obscures causal details, and conceals issues concerning adaptedness, which are of interest to biologists. These are pragmatic disadvantages. They do not indicate that genic models are semantically derivative of higher-level models, that genic models misrepresent causal processes, or that genic models fail to identify *the* true targets of selection. But depending upon biologists' goals, these pragmatic disadvantages could provide rational grounds for preferring higher-level models.

7. Why the Claim That Selection Is for Higher-Level Traits and Not for Lower-Level Traits Is Nonsense. Arguments about levels typically invoke causal claims. It's time to take a closer look at what it means to identify a cause. Sober argued that models of selection should identify genuine causal targets. In "Tempered Realism," I showed that the criterion that Sober and Lewontin (1982) applied to determine whether a model correctly pinpointed 'real' targets of selection could not establish their conclusions (Sober has subsequently abandoned his former line of argumentation). I also argued that the conception of probabilistic causation upon which their criterion rested didn't capture the sense of causation appropriate for understanding natural selection. At the time, I didn't have an alternative proposal. Now I do.

Over the past 15 years, philosophers have made significant progress developing a theory of causation that can help elucidate the idea that some selection models identify genuine causes and others identify pseudocauses. The theory I have in mind is the manipulability theory that has been set out in a wonderfully clear and careful manner in Woodward's recent book, *Making Things Happen* (2003). According to this theory, the distinguishing feature of causal claims is that they provide the basis for answering 'what if' questions about what would happen if certain prop-

8. I didn't make such erroneous claims in "Tempered Realism."

erties were manipulated in specified ways. Woodward formulates his account, not directly in terms of properties, but in terms of variables that take different values. He writes,

> The claim that X causes Y means that for at least some individuals, there is a possible manipulation of some value of X that they possess, which, given other appropriate conditions (perhaps including manipulations that fix other variables distinct from X at certain values), will change the value of Y or the probability distribution of Y for those individuals. (2003, 40)

According to this theory, causal relationships involve patterns of counterfactual dependencies.

The manipulability theory involves a number of subtleties that I will not discuss here. But, one subtlety is crucial for our purposes. It is not the case that any manipulation of the value of X that subsequently changes the value of Y implies that X causes Y. For example, a manipulation of the thermometer reading brought about by decreasing temperature would change the phase state of water, but this does not mean that thermometer readings cause water to freeze. To say that the thermometer reading causes water to take particular phase states is to say that a special kind of manipulation of the thermometer reading, a manipulation that didn't independently affect other causal variables such as temperature, would change the phase state. The technical term for the special kind of manipulation is 'intervention'. Woodward carefully spells out what counts as an intervention, but for our purposes we just need to keep in mind that the manipulability theory says that 'X, rather than V, causes Y' roughly means that an intervention in the value of X would change the value of Y, whereas an intervention in the value of V would not change the value of Y.

Although Woodward's theory is not reductive, and although there might be different senses of causation underlying different instances of reasoning that are typically called 'causal', Woodward's theory helps elucidate the difference between 'selection of' and 'selection for'.

Let's take a closer look at the toy illustration. I identified three different models of the selection process, each of which, modeled the process as selection for a different property: M_C (for color); M_{SC} (for spherical circumference); and M_{SD} (for spherical diameter). Following Sober, I appealed to a basic intuitive judgment when I claimed that M_{SC} represents the causes correctly, while M_C misrepresents the causes. But, what is the basis for our judgment? The answer, according to the manipulability theory of causation, is that the causal models disagree about what would happen if the system were manipulated in certain ways. M_C's answers to

the corresponding 'what if' questions are wrong, and M_{SC}'s answers are correct. Consider the following 'what if' question:

> What would happen if we changed the color of the green marbles to yellow?

According to M_C, color is the difference-making property, and the intervention on color (changing it from green to yellow) would result in the marbles descending to the level of the other yellow marbles. Changing the color would count as an intervention only if the change did not independently affect other causal factors. According to M_{SC}, however, a genuine intervention on color (i.e., a manipulation that did not also change spherical circumference) would not affect how far the marbles descend. This claim could be tested by painting the green marbles yellow (perhaps the procedure would include grinding down the marbles before painting them so the manipulation wouldn't change the marbles' circumference). Of course, the painted marbles would still descend to the bottom level. According to the manipulability theory of causation, this captures what it means to say that M_C provides a false account of the causal process.

The reason why we think M_{SC} represents the causes correctly is that we believe the level to which the green marbles descend by intervening on the circumferences of the marbles. This claim could be tested by manipulating the circumference of the green marbles (partisans of M_C might insist that we do so without changing their color). For example, we could increase the circumference of green balls by adding green clay to their surfaces. Now consider the following question:

> What would happen if we increased the circumference of green marbles to the circumference of the yellow marbles?

According to M_{SC}, circumference is the difference-making property and this intervention on circumference would result in green marbles coming to rest at the same level as the yellow marbles. According to M_C, the change in circumference will have no affect on descent and the marbles will descend to the bottom level. Of course, M_{SC} answers the question correctly.

The pair of experimental results relating to the pair of 'what if' questions would show that the selection process depends on circumference, not color. This illustrates how, according to Woodward's manipulability theory, the idea that M_{SC} gets the causal process right and M_C gets it wrong involves beliefs about counterfactual dependencies concerning what would happen if particular properties were manipulated in special ways.

The manipulability theory shows why it would not even make sense to claim that M_{SC} identifies the true causes and M_{SD} does not. For such a claim would mean that an intervention on circumference would affect

descent, but an intervention on diameter would not. But, it is impossible to intervene on circumference without also intervening on diameter. It is not that we lack the technological means to intervene on one without intervening on the other. The problem is that it is impossible to change one without changing the other, because of the geometrical relationship between circumference and diameter ($c = \pi d$). Hence, according to the manipulability theory, it simply makes no sense to claim that M_{SC} identifies the true cause and M_{SD} does not.

I claim that in PERM situations, the difference between individual models and multilevel models is more like the difference between M_{SC} and M_{SD} than the difference between M_C and M_{SC}. I explain why by returning to the case involving the sickle-cell trait. Recall that Williams claimed that the genic model M_G identifies the real cause of maintenance of the trait, while Sober and Lewontin (1982) argued that the diploid-genotypic model M_{DG} identifies the real cause (Sober has dropped this argument, but Lloyd persists). A monist would assume that one or the other partisan is correct. A pluralist would hold that both partisans might be correct and look for an argument to decide whether one, the other, or both models represent the causal situation truthfully.

Let's assume, for the sake of a reductio argument, that the monist is correct and either M_G or M_{DG} (but not both) gets the causality right. According to the manipulability theory, the meaning of the competing causal claims involves the following 'what if' questions:

> What would happen if we changed the contextual fitnesses of individual alleles? What would happen if we changed the fitnesses of diploid pairs of alleles?

The monist interpretation requires that the answer to first question be the outcome of the selection process would be different and the answer to the second be the outcome would not be altered (or vice versa). But, there is a problem. It is impossible to change fitnesses of the diploid allele pairs without also changing contextual fitnesses of the individual alleles (as shown by Equations 1–4 in Kerr and Godfrey-Smith 2002, 482–483). Likewise, contextual fitnesses of the individual alleles cannot be changed without changing fitnesses of the diploid allele pairs. These impossibilities do not arise because of contingent biological facts. They arise from the mathematical structures of the models. It is similar to the situation described above where spherical circumference cannot be changed without altering spherical diameter. The difference between M_G and M_{DG} is *akin* to the difference between M_{SC} and M_{SD}. (I say *akin* rather than *identical*, because applying M_G requires making explicit information that is not necessarily made explicit in applying M_{DG} and vice versa. Applying these different models requires bringing in different information, whereas ap-

plying M_{SC} and M_{SD} involves bringing in the same information.) Hence, claiming that M_G identifies properties that are selected for, while M_{DG} does not, (or vice versa) is, literally speaking, an exercise in muddled metaphysics.

Monists about levels of selection might resist this argument by denying that causation is the issue, or by claiming that Woodward's theory of causation is mistaken. To those who would deny that causation is the issue, I ask: What is the issue, if not causation? If the issue is just a matter of determining which models are empirically adequate and identifying practical advantages of different bookkeeping systems, then what Lloyd takes to be the key issue, the 'interactor issue', is a pseudoissue. Denying the importance of causal claims in this debate is a strongly antirealist position. As a causal realist, I wonder what such antirealists would say about Sober's selection device. Is M_C as true a description as M_{SC}? If not, why not?

Monists might acknowledge that causation is the issue, but claim that the manipulability theory is mistaken, and that my argument in this section is based on a false premise. (I have structured this paper so my earlier arguments do not depend on the manipulability theory.) One problem with the levels debate is that many disputants have been unclear about what they mean by cause, causal interactor, or causally appropriate (on this point, Lloyd and I agree). If monists seek to protect their epistemology by claiming the manipulability theory is mistaken, then the burden is now on monists to clarify causal notions such as 'interactor', 'force of selection', and 'selection of' versus 'selection for' in a way that doesn't undermine their position.

Wilson (2003) also draws on the linkage between properties at different levels to argue that Sober's distinction between 'selection of' and 'selection for' is irrelevant to the levels debate. Wilson's interests are mainly ontological, concerning whether causes in the world are entwined, fused, or leveled. My interests are epistemological, concerning whether we should interpret the best scientific theories, as (*a*) potentially capturing the one true way to represent a domain, or (*b*) potentially capturing one of perhaps a plurality of true ways. As I read Wilson, he is claiming that the selective causes are so entwined in PERM situations that they cannot be disentangled at distinct levels, as multilevel theorists seek to do. This may be true, and if so, it is a significant point. But it is not my point. My point is that PERM situations are sufficiently complicated (or entwined) that there are multiple ways to disentangle them. I think it is an open question whether multilevel theories can be developed to sort out causes at distinct levels. I believe that just as there is a plurality of (correct) ways to sort out selective forces at single levels for particular situations (e.g., at the genic level or at the diploid-genotype level for the sickle-cell situation),

there is a plurality of ways to sort causes at multiple levels. Different sortings may have different pragmatic advantages and disadvantages.

The manipulability theory provides a clear way to understand why the distinction between 'selection of' and 'selection for' is valid for distinguishing between selection models that identify true causes from those that do the bookkeeping by identifying pseudocauses. The theory also shows that the notion that selection can be for properties of groups—and not be for properties of the individuals making up the groups—does not make sense in situations where one could not, in principle, change a group property without changing an individual property (a basic condition of supervenience that all or nearly all philosophers of biology accept).

Sober's distinction between 'selection of' and 'selection for' is irrelevant to considerations about levels of selection. If the argument in this section is correct, notions widespread in philosophy of science (and especially philosophy of psychology) about identifying the real levels of causation are inherently confused.[9]

8. Conclusion: Guidance from the Past for Future Research. Pluralism denies neither the existence of empirical issues in the levels debate, nor the idea that there can be rational grounds for making decisions among a plurality of truthful models. What it denies is what monists read into those issues and decisions. Multilevel monists assume the debate hinges on identifying the uniquely correct model of any given situation. There is a plurality of models about some evolutionary situations, they believe, because biologists disagree about facts of the matter. Decisions among the different models of a given situation are decisions aimed at choosing the one true model and rejecting the false ones.

But not all decisions affecting the content of science reduce to decisions about choosing the true theory and rejecting the false ones. Hans Reichenbach put the point this way:

> That there are certain elements of knowledge, however, which are not governed by the idea of truth, but which are due to volitional resolutions, and though highly influencing the makeup of the whole system of knowledge, do not touch its truth-character, is less known to philosophical investigators. (1938, 9)

Reichenbach said that an important task of epistemology is to identify volitional stipulations. He distinguished between two classes of volitional decisions. He said 'conventional decisions' are between 'equivalent' con-

9. Discussions with Chris Hitchcock and Jim Woodward on this issue were helpful, but they should not be blamed for my simplifications.

ceptions, and do not influence the content of knowledge. He cited decisions to choose a particular system of weights and measures as conventional: The choice concerns "different ways leading to the same place" (1938, 10). Reichenbach said the other class of volitional decisions involve choices that can lead to divergent systems of knowledge. He called these 'volitional bifurcations'.

Some writers call pluralism about levels of selection 'conventionalism'. This label is misleading because it suggests that decisions between genic and multilevel theories have trivial consequences for the practice of evolutionary biology. The decisions between competing theories of selection actually fall under Reichenbach's category of volitional bifurcation. In cases of bifurcations, Reichenbach said philosophers should investigate the consequences of volitional decisions and elucidate entailments such as "If you choose this decision, then you are obliged to agree to this statement, or to this other decision" (1938, 16). It is time to stop quibbling about how to identify the *real* level of selection, and start investigating the pragmatic consequences of different volitional decisions about where to draw the environmental divide and how to parse and disentangle the causes.

Some of the most interesting current philosophical research on levels of selection issues can be understood in this way. Samir Okasha (2004), for example, has been exploring different statistical approaches for assessing models of PERM situations. His investigation, which presupposes a decision to draw the environmental divide where the multilevel selection theory does, explores the consequences of different volitional decisions about how to disentangle the causes. Sober and Wilson (1998) can be understood in a similar way. Okasha argues that one approach is 'theoretically preferable'. As a tempered realist, I interpret this conclusion in terms of pragmatics.

The manipulability theory suggests another way to assess alternative parsings. Kerr and Godfrey-Smith have started exploring the idea that some theories (or 'parameterizations') might turn out to be better than others in terms of accounting for perturbations.[10] As Kerr is developing this idea, some parameterizations might localize causation in certain perturbations more narrowly than others. Kerr's intuitions correspond to ideas of econometricians who believe that causal models are better insofar as they identify autonomous parameters. Roughly speaking, a parameter is said to 'autonomous', if it can be manipulated without changing the

10. Godfrey-Smith presented a paper on their research at the 2003 meeting of the International Society for the History, Philosophy, and Social Studies of Biology in Vienna. Kerr presented additional ideas to the SST colloquium at the University of Minnesota in 2005.

value of other parameters in the model (Woodward 1995). One might compare different selection models by examining whether the parameters within each model are autonomous in various situations. It is not clear where such comparisons will lead, but this approach (which is aligned with the manipulability theory of causation) might provide a practical basis for assessing models that parse causes differently, while still avoiding metaphysical muddles such as questioning whether maintenance of the sickle-cell trait is caused by selection for alleles in genic environments, or by selection for diploid genotypes in ecological environments.

REFERENCES

Brandon, Robert (1982), "The Levels of Selection", *Philosophy of Science* (Proceedings): 315–323.

Cassidy, John (1978), "Philosophical Aspects of the Group Selection Controversy", *Philosophy of Science* 45: 575–594.

Dugatkin, Lee Allan, and Hudson Kern Reeve (1994), "Behavioral Ecology and Levels of Selection: Dissolving the Group Selection Controversy", *Advances in the Study of Behavior* 23: 101–133.

Hull, David (1980), "Individuality and Selection", *Annual Review of Ecology and Systematics* 11: 311–332.

Kellert, Stephen, Helen Longino, and C. Kenneth Waters (forthcoming), "The Pluralist Stance", in Stephen Kellert, Helen Longino, and C. Kenneth Waters (eds.), *Scientific Pluralism,* Minnesota Studies in the Philosophy of Science, vol. 19. Minneapolis: University of Minnesota Press.

Kerr, Benjamin, and Peter Godfrey-Smith (2002), "Individualist and Multilevel Perspectives on Selection in Structured Populations", *Biology and Philosophy* 17: 477–517.

Kitcher, Philip, Kim Sterelny, and C. Kenneth Waters (1990), "The Illusory Riches of Sober's Monism", *Journal of Philosophy* 87: 158–160.

Lloyd, Elisabeth A. ([1988] 1994), *The Structure and Confirmation of Evolutionary Theory*. Reprint. Originally published by Greenwood Press. Princeton, NJ: Princeton University Press.

——— (2005) "Why the Gene Will Not Return", *Philosophy of Science* 72: 287–310.

Longino, Helen (forthcoming), "Theoretical Pluralism and the Sciences of Behavior", in Stephen Kellert, Helen Longino, and C. Kenneth Waters (eds.), *Scientific Pluralism,* Minnesota Studies in the Philosophy of Science, vol. 19. Minneapolis: University of Minnesota Press.

Maynard Smith, John (1976), "Group Selection", *Quarterly Review of Biology* 51: 277–283.

——— (2002), "Commentary on Kerr and Godfrey-Smith", *Biology and Philosophy* 17: 523–527.

Okasha, Samir (2004), "Multilevel Selection, Covariance, and Contextual Analysis", *British Journal for the Philosophy of Science* 55: 481–504.

Reichenbach, Hans (1938), *Experience and Prediction*. Chicago: University of Chicago Press.

Sober, Elliot (1984), *The Nature of Selection: Evolutionary Theory in Philosophical Focus*. Cambridge, MA: MIT Press.

——— (1990), "The Poverty of Pluralism: A Reply to Sterelny and Kitcher", *Journal of Philosophy* 87: 151–158.

Sober, Elliot, and Richard Lewontin (1982), "Artifact, Cause and Genic Selection", *Philosophy of Science* 47: 157–180.

Sober, Elliot, and David Sloan Wilson (1998), *Unto Others*. Cambridge, MA: Harvard University Press.

——— (2002), "Perspectives and Parameterizations", *Biology and Philosophy* 17: 529–537.

Sterelny, Kim, and Phillip Kitcher (1988), "The Return of the Gene", *Journal of Philosophy* 85: 339–361.

Waters, C. Kenneth (1985), "Environments, Pragmatics and Genic Selection." Address to the American Philosophical Association, Eastern Division.

——— (1991), "Tempered Realism about the Force of Selection", *Philosophy of Science* 58: 553–573.

——— (2000), "Molecules Made Biological", *Revue Internationale de Philosophie* 4: 539–564.

——— (forthcoming), "A Pluralist Interpretation of Gene-Centered Biology", in Stephen Kellert, Helen Longino, and C. Kenneth Waters (eds.), *Scientific Pluralism,* Minnesota Studies in the Philosophy of Science, vol. 19. Minneapolis: University of Minnesota Press.

Williams, George (1966), *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.

Wilson, Robert (2003), "Pluralism, Entwinement, and the Levels of Selection", *Philosophy of Science* 70: 531–552.

Wimsatt, William (1980), "Reductionist Research Strategies and Their Biases in the Units of Selection Controversy", in Thomas Nickles (ed.), *Scientific Discovery: Case Studies*. Dordrecht: Reidel, 213–259.

Woodward, Jim (1995), "Causality and Explanation in Econometrics", in Daniel Little (ed.), *On the Reliability of Economic Models: Essays in the Philosophy of Economics.* Dordrecht: Kluwer, 9–61.

——— (2003), *Making Things Happen.* Oxford: Oxford University Press.