NOTE TO READERS: THIS IS A DRAFTY VERSION OF MY COMMENTS ON BOB BATTERMAN'S *A MIDDLE WAY* FOR THE WORKSHOP AT THE U OF PITTSBURGH CENTER FOR THE PHILOSOPHY OF SCIENCE ON JANUARY 28, 2023. I'M POSTING IT HERE IN THE OFF-CHANCE THAT ANYONE MAY BE INTERESTED

BATTER (MAN)ING FUNDAMENTALISM: SOME REFLECTIONS ON BOB'S *A MIDDLE WAY*

1. Introduction. It is a great pleasure to have this opportunity to comment on Bob's book, *A Middle Way* (hereafter *MW*). I've learned a great deal from it, as well as from Bob's other work. I will leave it to others who are more knowledgeable about physics than I to comment on the details of Bob's discussions of particular physical theories. What I'm going to do instead is to use his book as an occasion to discuss some more general issues having to do with autonomy, reduction, and the role of meso-level theorizing. My very meandering discussion is organized as follows. I begin (Section 2) with some general reflections on autonomy-- on what does and does not amount to and how it might be understood in terms of a notion that I call conditional independence (Section 3). This is followed by some brief comments on the notion of level (Section 4). Section 5 emphasizes the important point that in realistic cases autonomy is a matter of degree and relative to a particular set of explananda. Section 6 connects my discussion with Bob's question **AUT** about how to explain autonomy and also emphasizes that this question is of very general methodological interest, having implications that go well beyond the particular examples that Bob discusses. Section 7 connects the discussion of autonomy to other notions, also discussed by Bob, that are in the same ballpark-- universality and multiple realizability. I also emphasize that issues about autonomy -- how it should be characterized and explained-- are to a significant degree orthogonal to those having to do with reduction, contrary to what a substantial portion of the philosophical literature supposes. Section 8 responds to the claim that multiple realizability and universality are rare. Section 9 uses some of Bob's examples to explore some fashionable ideas about the relations between micro and macro theories-- that such relations should be understood in terms of concepts like identity, realization, constitution or the like. Section 10 comments further on the notion of reduction and its connection with autonomy. Section 11 concludes with some general reflections on why the meso level is as important as it seems to be.

2. Autonomy. Let me begin with the notion of autonomy. The word is used in many different ways, but one natural meaning (and the one I will adopt) is that a theory T is autonomous to the extent that T does not need information from some other theory T* (perhaps particularly information coming from a theory at a different "level") to adequately model the range of phenomena with which T is concerned. (I'm going to understand"adequately model" to mean something like "explain" in what follows, but one can also raise parallel questions about what sort of information a theory needs for purely predictive purposes.This may be different from what is required for explanatory purposes.) To use one of Bob's examples, continuum fluid mechanics as described by the Navier-Stokes equations is autonomous from an underlying theory concerning the molecular constituents of the fluid to the extent that the former can correctly describe the behavior of fluids without information from the latter. (To foreshadow an obvious point, T may need some information from T* but not other sorts of information, so that autonomy is not an all or nothing matter.) To use some examples that are different from Bob's, a psychological theory is relatively autonomous if it does not need (or does not need much) information from neurobiology to adequately model the phenomena with which it (the psychological) theory is concerned. Micro-economics is autonomous to the extent that (contrary to claims by neuro-economists) we can adequately model the individual choice behavior and other phenomena on which economics focuses without bringing in information from neurobiology. Notice that this sort of autonomy is consistent with e.g. neurobiology being in some sense "relevant" to psychology-- relevant in the sense that, for example, if the structure of the human brain were sufficiently different, the generalizations of psychology would be false. Similarly if, say, the characteristics of the strong force in nuclei or the electromagnetic forces holding molecules together were sufficiently different, the N-S equations would not hold-- so in that sense this information about the electromagnetic force etc. is relevant to the N-S equations.

3. Conditional Irrelevance. Elsewhere (e.g Woodward, 2021) I have suggested that the features of autonomy just described can be understood in terms of a relationship that I call conditional irrelevance or conditional independence. Suppose we have an upper-level theory which formulates true or correct dependency relations (by "correct" here I will mean "correct up to some acceptable level of approximation"[1]) among a set of upper- level variables {*Ui*}-- that is, the upper-level theory contains generalizations describing e. g. how *U3* depends on *U1* and *U2* and so on. Dependence or "relevance" here means counterfactual dependence understood along broadly interventionist lines or perhaps some suitable generalization of this approach-- *U3* depends on *U1* and *U2* in accordance with some functional relationship *U3= f(U1, U2)* if *f* correctly describes how the value of *U3* changes under "wigglings" (including those produced by interventions) of the values of *U1* and *U2*. In this case I will say that *U1* and *U2* are unconditionally relevant to *U3*, with unconditional irrelevance or independence being the absence of unconditional relevance or dependence. Suppose also there is an "underlying" lower-level theory formulating true dependency relations among variables {*Lj*}. Let *E* be some candidate explanandum. Assume that some of the {*Ui*} are unconditionally relevant to *E* and similarly for some of the {*Lj*}. Finally-- this is the most important part-- let us say that a set of *L*-variables is irrelevant to (independent of) *E conditional* on the *U* variables if the *Ls* are unconditionally relevant to *E*, the *U*s are unconditionally relevant to *E, and* conditional on the values of the *U*s changes in the value of *L* produced by additional interventions consistent with these values for the *U*s are irrelevant to *E*[2]. Put differently, suppose that according to the upper level theory, *E* depends on *Uk*. Consider various values for the variables *Lj* of the lower level theory, where each of these values is consistent with the values taken by *Uk* . If, conditional on the values taken by *Uk*, further variations in the values of the *Lj* consistent with the values of *Uk* make no difference to *E*, then the *Lj* are irrelevant or independent of *E*, conditional on the values taken by the *Uk*. Put informally the upper-level variables "screen off" the lower-level variables from *E*, where the screening off relation is understood in terms of appropriately behaved (that is, interventionist or non-back tracking) counterfactuals rather than probabilistic independence. The intuition is that to the extent that this is true, all of the difference-making information in the lower-level theory that is relevant to *E* has been absorbed into the upper- level theory and it is this difference-making information that is the explanatorily relevant information. For example, the NS equations are (apart from the values of particular parameters

[1] It is important that it is the claimed *dependency relations* that are good approximations (or not). Lots of other features of the theory or model may not be anything like correct approximations, as when, e.g. a model with two dimensions correctly tells us about some dependency relation in a three dimensional system, even though the two dimensions in the model are not an "approximation" to the three dimensions in the real system. This reflects the idea that explanation is a matter of getting dependency relations right and not a matter of there being a correspondence or isomorphism between the model and the system in other respects.

figuring in them, such as viscosity) autonomous from the details of the underlying molecular components of fluids to the extent that given the the values for the variables figuring in those equations (fluid pressure, density, velocity etc.) further variations in the behavior of the molecular components are screened off from many explananda having to do the bulk (large scale) behavior of the fluid. Notice again that this does not say that the behavior of the fluid is completely independent of the characteristics of its molecule components-- it rather says that all or most of what is relevant at the molecular level can be absorbed into upper-level variables. To anticipate what I will say below, autonomy in this sense is a matter of degree. At one extreme a successful upper level theory may not need any lower-level information. In other, more realistic cases the upper-level theory may need only limited lower level information-- e.g., to specify the values of key parameters but not the overall form of the equations characterizing the upper-level dependnecy relations[3].

Although the above characterization does not require this, it is also worth underscoring that in realistic, interesting cases of autonomy, the autonomous theory will virtually always involve some degree of coarse- graining or loss of degrees of freedom in comparison with the theory from which it is judged (relatively) autonomous. A theory T that exactly reproduces all of the distinctions, information and dependency relations in another theory T* (as in the psychology/neurobiology example mentioned below) is trivially autonomous with respect to T* but I take it that such cases are, at best, rare and uninteresting (and in fact raise questions about whether the two theories are genuinely distinct.)

In this connection , let me also raise the following question— I would welcome input from Bob and others on this: can the conditional irrelevance idea that I have described be used to capture at least part of what is going on with the use of RG methods? That is, to the extent that the goal in RG is to find coarse-grained variables that absorb or capture the information in more fine-grained variables and representations that is is relevant to certain macro-explananda, does the conditional relevance idea above capture this notion of relevance? Of course the characterizations of conditional relevance and irrelevance given above do not provide a *method* for determining which are the relevant upper level variables and which variables are irrelevant. The RG does this, at least for certain problems. It is

---

[3] One way of trying to make this more precise is to provide a quantitative measure of the amount of information that must be supplied to the upper-level from the lower level theory for the former to be successful and /or a meaure of the information that is lost in transitoning from the lower to upper level theory. See Ay and Polani, (2008) for a possible way of approaching this.

a very interesting question whether there are other methods for separating relevant from irrelevant variables that might work for other problems.

As these examples illustrate when we ask about autonomy, we are often, although in my view not always, interested *macro/ micro relationships*-- that is, the extent to which some upper level or macro theory concerning a system S is autonomous from some lower level, micro theory that is also concerned with S, with the micro theory understood as characterizing the behavior of entities, properties etc. that make up S. However, while this is a common case it is not the only one. We might also wonder whether a theory at level *L1* is or could be autonomous with respect to information coming from what we think of as a "higher" level *L2* (as might be the case if, for example, the entities at *L1* are among the parts of those at *L2*. For example if genetic anomalies are among the causes of cancer but if (as seems to be the case) tumor growth is also causally influenced by larger scale features of the tumor environment such as facts about tissue stiffness and connectivity, then a theory of cancer that just makes reference to genes will fail to be completely autonomous. It is thus important to keep in mind that autonomy can fail in "both" directions-- a theory can fail to be autonomous from lower-level information but it can also fail to be autonomous from upper level information

4. Levels. This formulation ties the notion of autonomy to ideas about the individuation of theories or "levels" and the domains of phenomena to which theories are responsible (we have to be able to say what belongs to psychology and what to neurobiology, if we are to discuss whether the former is autonomous from the latter) and these notions are no means entirely unproblematic. On the other hand, we can often be precise enough in particular cases to capture what is meant-- we can specify, e.g a psychological theory T and a range R of phenomena which it is intended to account for and then ask whether T by itself is sufficient to account for R or whether it needs to be supplemented in some way by additional information. Bob's examples-- e.g., his discussion of the relative autonomy of continuum fluid mechanics from the "underlying" theory of the constituent atoms and molecules making up the fluid -- make similar assumptions about our ability to distinguish theories and levels and I will follow him.

Some philosophers (e.g., Potochnik, XX Eronen, XX) have recently claimed the whole notion of nature containing levels is confused. I think this is just wrong and if I have understood Bob, he does too. On my understanding a "level" is just a set of linked phenomena, related by stable generalizations, that are relatively autonomous. If you don't like some of the associations that come with the notion of "level", it will suit my purposes if you substitute "scale" instead[4].

5. Autonomy as a Matter of Degree. As noted above, this way of thinking about autonomy suggests immediately that it is a matter of degree and relative to particular class of effects or explananda that need to accounted for. For example, a theory T might require, at least for a very large range of phenomena, no or only a very limited input from some other theory or body of information T*-- e.g., T* might be required only to fix the values of certain parameters that are left undetermined in T--with the equations of T themselves needing no correction. This is the case with the Navier-Stokes equations which in their treatment of continuum level phenomena require only supplementation with lower-level parameters that fix the values of such parameters as viscosity. Alternatively it may be that the equations of T themselves breakdown or require correction for certain phenomena or effects although the equations perform just fine in connection with many other explananda. An example of this is provided by the classical statistical mechanics of dilute gases which accounts for a substantial range of phenomena but fails in connection with specific heats, where information from quantum mechanics is required.

I noted above that one obvious sets of context in which issues about autonomy arise have to do with cases in which we have an upper level or macro-theory concerning a set of systems S and a lower level micro-theory, also concerned with S, with the micro-theory understood as characterizing the behavior of elements that make up S and where we are willing to treat the lower -level theory as "correct" or a ground truth (again in the approximate sense described above). If the upper-level theory is relatively autonomous from the lower-level theory in the sense described above, this means that some of the information in the lower-level theory does not matter for the behavior (that is, the correct characterization of the behavior) that the upper-level theory provides. This implies that some of the distinctions and degrees of freedom that are recognized in the lower-level theory are not (because they do not need to be) recognized in the upper-level theory. In other words, the lower-level theory will treat various microstates as heterogenous, while distinctions among these states are not recognized in the upper level theory, so that some kind of coarse-graining of the

---

[4] For more on levels, see Woodward, XX. The autonomy or scale-based notion of level that I favor should be distinguished from accounts that tie the notion of level to the subject matters of particular sciences (the "biological" vs the "chemical" level) or to part/whole relations-- individual cells are at a lower level from multi-celled organism because the former are "parts" of the latter. Woodward, XX argues that these notions of level have little to recomend them.

variables of the lower-level theory is present in the upper-level theory. Of course, as Bob illustrates, this coarse graining can take many different forms and need not involve just simple versions of averaging.

6. BOB on AUT. All of this leads up to the issue that Bob is concerned with when he formulates the question that he calls AUT:

> How can systems that are heterogeneous at some (typically) micro-scale exhibit the same pattern of behavior at the macro-scale?

This is formulated as a question but it assumes a characterization of autonomy that I think is very like that described above: autonomy has to do with the extent to which systems that are heterogenous at some (typically) microscale exhibit the same pattern of behavior at a more macroscale and this implies that the heterogeneous details at the micro- level do not matter to some aspects of the upper level behavior. The world presents us with examples of macroscale behavior that are autonomous in this sense and Bob wants an explanation (or explanations) for this.

Bob formulates his AUT as a question about explanation but it is worth noting that there is a closely associated issue that is more related to methodology and research strategy: In many areas of science we have lots of low-level information about the behavior of constituents of systems-- in the case of the brain, information about the biochemistry of neurotransmitters, the factors affecting synaptic growth, the generation of action potentials in neurons, wiring diagrams and other information about connectivity at various scales for parts of human and non-human brains and so on. But for many reasons we want to be able to formulate higher level theories about the large or macro-scale behavior of brains-- in cognition, emotional processing, choice and so on. Such higher-level theories are going to have to make use of some of this lower-level information but, in part just for reasons of tractability, if such higher-level theories can be formulated at all, they are going to have to abstract away from much of this more micro information, ignoring lots of detail which we hope can be treated as efffectively irrelevant.. Put differently there had better be such irrelevant detail for some of the explananda in which we are interested since if all of the lower-level information is relevant, the result will be far too complicated for incorporation in a coherent upper-level theory. Similarly for many other complex systems -- the influence of genetics on many behavioral traits and diseases, the human heart and perhaps the macro economy.

It would be naive to suppose that answers to the question of how the stable upper-level generalizations of hydrodynamics or solid materials are possible will automatically transfer directly to these other areas. Nonetheless, I think there may be something more general to learned from Bob's examples -- roughly that a key may be to look for middle level or meso-structures that mediate between micro and macro theorizing, where these structures that are often in some broad sense "topological", in the sense that they concern larger scale correlations, patterns of connectivity, scaling relations and similar matters. I acknowledge that this is speculative and will try to say a little more about it below. In any case, I want to emphasize that I see Bob's question about explanation in his AUT and my more methodological concerns as intertwined-- if we have an understanding of how autonomy can arise in some cases where the relevant science is well understood, as it is in Bob's examples, this may help us in figuring out where we might expect to find it in other scientific areas or whether instead the search for autonomous theories may be hopeless in some areas.

7. Autonomy in Relation to Multiple Realizability, Universality, and Reduction. But turning to some of these issues related to methodologies for understanding complex systems, however, I want to say a bit more about what autonomy implies when understood as above and its relations to some other notions that Bob discusses that are in the same ball park. These include universality, multiple realization, and how these relate to so-called reduction.

To begin with, autonomy as I understand it (and I think also as Bob understands it) is very different from what some other writers have understood by this notion. For example, according to Jerry Fodor, psychology is autonomous if it is not reducible to neurobiology, where by reduction Fodor means Nagelian reduction with type- type identities-- that is every property or variable in one's psychological theory is type-identical with properties or variables in the underlying neurobiological theory. Note, however, that if this were the case, the generalizations of the psychological theory would exactly mirror or track the generalizations of the neurobiological theory. If the latter are exceptionless-- the ground truth-- the same must be true for the generalizations of the psychological theory. So in doing psychology there would be no need to advert to information from the neurobiological theory; thus the psychological theory would be trivially autonomous in my sense but not in Fodor's. Without wishing to get involved in a semantic dispute about how best to understand autonomy, I think it is pretty clear that Fodor has departed from some of our usual understandings of this notion. (Of course, the suggestion that there might be such a Nagelian reduction of psychology to neuroscience is completely fanciful-- I mention it just for purposes of illustration.)

We can also see from this example, that as a number of other writers have recently claimed, trying to understand autonomy as the "opposite" of "reduction", conceived along Nagelian lines, does not seem fruitful. Indeed, given the characterization above, it is entirely possible, given a sufficiently permissive notion of reduction (see below), that in realistic, non-fanciful cases, T can be autonomous with respect to T* and yet reducible to T*. As Bob's examples illustrate, autonomy of T with respect to T* does not require that there be no connection between T and T* or that we be unable to understand from the perspective of T* why T is successful. (By contrast, Fodor seems to think that the autonomy of the upper level theory requires that its success be completely ununderstandable-- a kind of miracle-- from the perspective of the lower level theory.) Rather what matters is *how* the subject matter of T depends on T* and whether that dependence is of such a character as to support conditional independence relations of the sort described above. In typical cases in which autonomy of *T* fails in an extensive way, *T* cannot even be formulated (that is, in terms of correct dependency relations) without including lots of input from *T** . In such cases, within a Nagelian framework there is no self-contained *T* which is even a candidate for reduction to *T**. In other words, we have a failure of autonomy *and* a failure of reduction.

Next, universality and multiple realizability. Universality is of course a physicist's notion, meaning roughly the independence of aspects of the behavior of set of systems from the lower-level details of their components-- thus sameness of behavior in systems that are non- trivially different at some lower level of analysis. MR is a philosopher's notion, being present, roughly, when variables that are different from the point of view of the underlying theory, or perhaps different values of the same variable in the underlying theory are mapped into the same variable or value in the upper-level theory, presumably with the added requirement that this mapping leads to true formulations of dependency relations in the upper-level theory. As it is usually understood, for MR to be present not just any many-one function will do-- the mapping must also involve the lower level variables or values bearing some relation to the upper level variables or values that is more intimate than ordinary causation or correlation-- this relation is variously described in the philosophical literature as "realization", "constitution", "token identity" and the like. I will put these claims about the nature of the upper/ lower relation aside for the moment but will return to them below.

When interpreted literally as characterized above, MR seems weaker than universality and does not seem capture the kind of autonomy we are interested in. Consider the following bit of science fiction, which I help myself to because everyone else does. The generalizations of some upper-level psychological theory (say, folk psychology-- FP) can be realized in a human brain but only when that brain is in a very special condition-- the wiring patterns, levels of neurotransmitters and so on have to be just so. If these conditions are different in any way the folk psychological generalizations break down. There also Martian brains based on (what else?) silicon and a similar thing is true of them-- the generalizations of folk psychology (e.g., relating Martian-type beliefs and desires to behavior) are realized in Martian brains when these are in a very special condition (different of course from the realizing condition for humans) but not otherwise. Here we have a case of the MR of FP in the literal sense but the case is not one in which FP is autonomous from the underlying carbon and silicon-based neurobiological theories that characterize humans and Martians. On the contrary, under our imagined scenario FP breaks down under most underlying neurobiological conditions and we need both human and Martian neurobiology both to understand when this break-down occurs and also to formulate more accurate generalizations. Without presuming to speak for physicists I also doubt very much that they would regard this as an example of universality[5]. (In addition to the instability that is present, in the story as I have told there is nothing that suggests any interesting "intrinsic" characterization of the universality class in the example -- e.g. there is no interpretable smooth transformation that takes the human case to the Martian case while including other cases in between.)[6]

---

[5] In thinking about universality it may be worthwhile to distinguish two different kinds of cases. In the first, we have a stable upper level generalization characterizing a single kind of system S that continues to hold across changes in lower level states that are physically possible transitions that in many cases actually occur in S . For example, the ideal gas laws continue to hold for an ideal gas as the velocities and momenta of its component molecules change. In the second kind of case, a stable upper level generalization holds for systems that differ more radically in kind-- e.g., gases and ferromagnets. Here, in contrast to the first case, there is no physically possible way of transforming or transitioning from a state of one of these systems to a statemof the other-- the transformations relating the systems are instead abstract and mathematical. My impression is that physicists tend to use "universality" to characterize the second kind of case but that it is often assumed that behavor corresponding to first sort of case is also present in the systems of interest.

[6] As Bob notes, Fodor and others who hold similar views have claimed that the postulated instantiations of FP in humans and other organisms need not have anything at all in common (other than that they are instantiations of FP). So from this perspective, the absence of any such intrinsic characterization is not a bug at all, since it is part of the view that there is no further explanation of why these different organisms behave in accord with FP. Instead this is just an arbitrary (and massive) coincidence. Of course one of Bob's messages is that this is not what one finds in realistic cases of universal behavior.

For what it is worth my diagnosis of Fodor is that he basically wants to be a dualist and his claim that psychological generalizations can hold in very different organisms with no underlying theory of a more physicalist or neural sort that explains why this is so reflects this impulse toward dualism-- it reflects the characteristically dualist commitment to the

In any case, to the extent that FP is autonomous with respect to neurobiology this will be because it behaves very differently from what we see in the above example. At the very least, some non-trivial degree of autonomy in FP requires that the generalizations of FP be stable across many of the lower-level changes that "normally" occur in individual human brains as well as normal variations across human brains, where presumably the latter are something like the variations that occur in the former case. Indeed we know on very general grounds that to the extent that there is any regularity in behavior and cognition at the macroscopic or whole person level, there *must* be some substantial degree of independence of such regularities from realizing lower level detail. The behavior of individual neurons is highly stochastic, neurons die and in some cases are replaced by others, new synapses are continually forming and so on and all of this is consistent with some significant degree of stability in human cognition and behavior. If organism level behavior and cognition was highly sensitive to such variation, creatures like ourselves would not exist.

One conclusion we can draw from this is that the mere presence of MR is not remotely enough to secure autonomy or to give those who think that upper-level sciences are sometimes at least somewhat autonomous from lower level realizers what they want. Of course we can draw the same conclusions from Bob's examples-- the universal behavior that he talks about in connection with continuum level fluid mechanics and even more so the similar behavior of ferromagnets and fluids near their critical points involves something much stronger than MR as characterized above.

8. **Objections to MR and Universality**. In MW Bob notes in passing that the claim that MR is common has been criticized by a number of philosophers, including Shapiro and Polger (SP) (whom he cites) . It is worth considering their argument since it seems to me rest on a misunderstanding of what MR and the stronger versions associated with universality and autonomy require-- a misunderstanding that when corrected, helps to illuminate their structure. At an abstract level SP's argument goes as follows. Consider a putative example of MR in which upper -level property *U* is realized by two lower- level properties *L1* and *L2*. (SP put things in terms of properties rather than variables and I will follow this, even though I think that there is much that is misleading about property talk in this context). One possibility is that the two properties *L1* and *L2* are not different

---

inexplicability of the mental in non-mental terms. At the same time he recognizes that respectable philosophers are physicalists, so he tries to have things both ways by combining his views about the inexplicability of the mental from the point of view of any underlying physicalistic theory with an endorsement of token physicalism. The upshot is a kind of "effective dualism" (dualism in all but name) with genuflection in the direction of physicalism.

enough to count as distinct. If so, this will not really be a case of MR since there are not two distinct realizers for *U*. Alternatively, suppose *L1* and *L2 are* distinct. Then, SP argue, it is extremely likely that *L1* and *L2* will have *some* different effects and they take this to show that the case will not really involve MR. In other words, their effective criterion for whether MR is present is that lower-level realizers of *U* must be both distinct and have exactly the same effects at every level of analysis. For example, in their discussion of well-known experiments in which rewired ferrets use portions of their auditory cortex to "see"[7], they note that the rewired ferrets differ from normals on some fine-grained visual discrimination tasks (although not on other visual discrimination tasks), and infer that because of this difference, this is not a case in which the same property is differently realized. (Somewhat oddly, they also suggest that the normal and rewired ferrets may not be relevantly different after all, since auditory cortex in the rewired ferrets exhibits a columnar organization similar to that of the visual cortex in normals and thus that the case may fail to illustrate MR because the realizers are not different. Obviously this is a sort of "heads I win, tails you lose" strategy, since both similarities and difference in behavior are counted as objections to MR.)

This line of argument seems to me to neglect the observation above that universality and autonomy must be understood as relative notions--a theory is autonomous or not with respect to certain explananda, effects or information but not others. The Navier-Stokes equations are universal or autonomous in the sense that the general form of those equations holds for fluids of very different molecular compositions but different fluids will differ in some respects-- for example, in their viscosity, which enters as a parameter in the NS equations. It would be obviously misguided to argue that the NS equations are not illustrations of universal behavior on the grounds that, say, fluids like water and oil have *some* different effects-- for example, differences in effects that are related to viscosity differences, not to speak of other effects, such as their consequences for ingestion. A similar moral seems to me to apply to the ferret case.

One unfortunate consequence of the requirements on MR and related notions imposed by SP is that attention is diverted from phenomena that are genuinely puzzling and stand in need of explanation. It is prima-facie puzzling -- something that stands in need of explanation-- that ferrets can use their auditory cortices to successfully perform some visual tasks even though the structure and anatomy of auditory cortex is in some respects rather different from their visual cortices. Similarly it is puzzling that fluids of very different material composition can

[7] In these experiments the auditory pathway leading to the thalamus is destroyed with the result that optical pathways connect to auditory cortex as well as to visual cortex . The optical pathways then make use of auditory cortex.

exhibit continuum level behavior that is the same in important respects. In MW Bob describes how this latter fact can be explained. An explanation in the case of the ferrets also seems available: As noted above, in the case of the ferrets it turns out, unsurprisingly, that there are higher level structural similarities between the organization of ordinary visual cortex and the organization of rewired auditory cortex (the latter acquires a columnar organization similar to that involved in ordinary visual cortex) as well as similarities in the computations performed and this provides an explanation of how both normal and rewired ferrets can succeed on certain visual discrimination tasks.

(An aside: I've also sometimes encountered the following objection to talk of MR-- presumably if the objection is cogent, it would also apply to universality and autonomy. The objection is that whenever a theory fails to cite some factor as a difference-maker, it is trivially true that it is autonomous with respect to that factor-- hence autonomy and universality are so ubiquitous as to be uninteresting. For example, Newtonian gravitational theory treats masses of different colors in exactly the same way-- differences in color can be ignored. Is this a case of autonomy? We might consider biting the bullet and saying "yes," and that if this judgment seems odd this just reflects the fact the case is obvious and uninteresting. However, I think a better response is to observe that there is no underlying physical theory or body of information that suggests that color might be relevant to the behavior of fundamental physical forces. Thus in this context, masses of different colors do not count as heterogenous in any relevant sense and there is no basis for an expectation that they will behave differently which then needs to be addressed by a universality argument. Put in terms of my idea about conditional independence, there is no underlying fine grained theory that makes discriminations on the basis of color that is then screened off by a more macro theory. Matters are quite different with respect to the behavior of fluids and ferromagnets which are regarded as in important respects quite different from the perspective of a number of well-founded theories.

If we think of universality and autonomy in the ways described above it seems obvious that (1) there is quite a lot of it, (2) if there were not quite a lot of it, the world would be unrecognizably different and it seems unlikely that we would be around (since our existence seems to require a substantial amount of upper-level stabiity in nature's behavior and corresponding insensitivity to many variations in lower level details-- a point also made by Bob) and (3) to a substantial extent any understanding we are going to be able to achieve of macroscopic systems, especially those with any degree of complexity is going to depend on the presence of the right forms of universality and autonomy, so that strategies for identifying and exploiting these, which I assume will be largely mathematical, are crucial. Arguments that MR (and thus universality) are rare or impossible lead us to miss all of this. They also of course lead to a failure to recognize that universality and autonomy stand in need of explanation.

With respect to (1), we have as Bob has emphasized a substantial amount of autonomy or independence of realizing low level detail in connection with many familiar materials -- fluids, metals, and so on. The construction of tools and machines depends on this. In living organisms and structures like brains, there is a great deal of lower-level noise and lower level variation both within single organisms and structures and across these which is somehow consistent with a good deal of stability in upper-level behavior—as noted above, the behavior of individual neurons is highly stochastic the operation of our brains do not change radically when a single neuron dies, and so on. Of course, the way in which this stability is achieved in organisms is somewhat different from the way in which it is achieved in non-living things since in the former case there is *selection*[8] for structures that actively maintain higher-level order and (at least typically) this is not the case for non-living things. However it also appears there is considerable similarity in the ways in which upper level stability or relative autonomy is achieved in both living and non-living things— in both cases this is often achieved via averaging, exploitation of facts about correlations and so on. Natural and other forms of selection find and exploit relatively stable upper-level casual relations and these will relations that are also be present in non-living things. So the last thing that we want is a treatment of MR, universality and autonomy that denies the existence of such stable, relatively autonomous relations (or defines them in such a way that it is hard to see how they could possibly exist)-- instead we need characterizations of these notions that allow us to understand when and why they occur.

9. Identity, Realization, Constitution, Parts and Wholes

I noted above that these are terms of art that are widely used in the metaphysics literature and to a significant extent in the philosophy of biology and psychology literatures to characterize relations between entities and magnitudes in lower-level theories and those in upper-level theories. I see these notions as motivated at least in part by expectations and assumptions associated with classic Nagelian models of reduction and subsequent attempts to modify such models. Nagel's model focused on issues having to do with the deducibility of the reduced theory from the reducing theory but, as subsequent discussion made clear, deducibility, even if it is to be had, is not sufficient to capture what reductionist-minded philosophers were generally after. One might imagine, for example, an

[8] Selection of various sorts-- not just natural selection but various forms of learning.

account according to which mental properties and entities are completely distinct from physical properties or entities but linked to them via bridge laws that permit the deduction of claims about the mental from claims about the physical. This amounts to dualism, rather what is usually intended by a successful reduction of the mental to the physical. Instead, the usual understanding is that for a successful reduction we need some much more intimate relation between the entities and properties of the reduced and reducing theory. Some form of identity, either type or token, is the most obvious candidate and the one originally proposed. However since this leads to familiar problems, more recently philosophers have been focused on other candidate relations like constitution, composition, realization and the like. The basic idea is that these relations are different from straightforward identity but at the same time also different from from relations like causation or correlation which we think of as obtaining between completely distinct entities. One of the many interesting things about the examples that Bob discusses from continuum fluid mechanics and materials science is that they provide explicit scientifically fruitful accounts of relations between theories at different levels but accounts that do not seem to make any very heavy duty of the notions (identity, composition, constitution) described above.

Consider for starters claims about identity (of either the type or token variety). Compare these with the relationships between meso- level structures in materials and macro-level parameters like young's modulus, that are established, as Bob explains, through the use of representative volume elements and homogenization techniques. Does it illuminate anything to say that these upper-level parameters or their values are identical with some collection of meso-level features? Even if we could make sense of such identity claims, they don't seem to capture what is going on. The upper-level parameters reflect the extraction of certain limited amounts of information from the meso-level features-- the information that is relevant to the upper-level behavior of interest with irrelevant information being discarded. Claims about identity don't seem to help us understand how this works. Indeed, they seem to point us in exactly the wrong direction: prima-facie, a necessary condition for an upper level property U to be identical with a lower level property L is that U and L have the same dimensionality or figure in theories with the same degrees of freedom, but the whole point about the cases Bob considers (and many of the others mentioned above) is that the upper-level theory has lower dimensionality or fewer degrees of freedom than the lower-level theory. By contrast, identity seems to imply that everything about the lower-level entity or property is relevant to the upper-level identity or property with which it is identified or at least an identity claim does not tell us what is not relevant. (Saying that the relation in question is one of token identity between "events" or other sorts of particulars does not seem to help either, since if individual event u is token identical with individual event l, that still implies that everything about l is relevant to u— indeed the usual understanding of identity is that identical events -- or whatevers-- enter into exactly the same causal relationships.) Instead what is going on is that much of the lower level information is irrelevant to the upper level behavior and what we want to know is how this comes about. Again, claims about identity don't seem to answer such questions-- at best they are going to be justifiable only after we have answered questions about what is relevant or irrelevant.

I think that a similar conclusion holds for notions like realization and constitution, which are popular in the literature on "mechanisms" and elsewhere. Of course one can say things like "a fluid is constituted by the molecules that are its components" or "a fluid is realized by its constituent molecules". But even if we assume that there are interpretations of "constitution" and "realization" for which these claims are true, by themselves they gives us no insight into what is relevant and what is irrelevant to the continuum level behavior of the fluid. Similarly for saying the relationship between the molecules and the fluid is one of parts to a whole.

I'm inclined to think that the focus on identity, constitution and so on in relations between theories is largely an artifact of Bixin Guo XX has called ontology- first models of theories and what David Wallace XX calls a language-oriented rather than a mathematics-oriented understanding of theories. If the ontological commitments of a theory are central to understanding it, then of course in understanding the relationship between theories at different levels we need to provide an account of how their ontologies are related, which naturally leads to questions like "what among the entities and properties postulated by the lower level theory can be identified with (are constituted by etc.) the entities and properties in the upper level theory". Suppose, however, that inter-theory relations should instead be understood in terms of mathematical relationships involving limits, approximations, topological relationships like self-similarity at different scales, patterns or networks of connectedness among heterogeneous units and the like. These may have no natural interpretation in terms of notions like identity and constitution which, to repeat, are notions introduced to relate ontologies.

My suggestion is thus that we drop notions like identity, constitution, realization and so on in characterizing relations between theories or at least that we stop pretending that use of these notions gives us much insight into what is going on when we attempt to understand these relations. As Bob's examples illustrate, what matters for the relations between theories is something more like the extraction of relevant information that can be passed from one to the other.

How this works is what philosophers of science should be focusing on, not the metaphysics of constitution or realization. I hasten to add that (of course) this does not mean that upper-level properties and entities are unreal or fictitious -- it is rather that they needn't (and often don't) line up with properties and entities in lower level theories in a way that supports talk of identity, constitution and the like. It is just a mistake to suppose that if we can't relate elements in the ontology of upper and lower level theories via some relation like identity or constitution, we must be eliminativists about the upper-level objects.

10. More on Reduction. What about the other component of Nagelian "reduction"-- the deducibility part? Nagel's account is of course formulated in terms of first-order logic, where "deduction" has a clear meaning. But what if anything does "deduction" (and still less "deducibility in principle") mean when we consider mathematical relationships that are not well captured in first order logic: various sorts of limiting relationships, including "singular" limits, relations involving approximations of various sorts, including arguments based on empirical considerations that certain terms can be neglected because they are "small", cases in which we can extract information from the first few terms of a divergent series and so on. Notions like these are needed to characterize relationships between theories including those discussed by Bob and so the tendency in the recent literature on reduction has been to expand the notion of deduction to include such relations. In a relatedly permissive vein, recent writers (e.g. Palacios XX) have contended that "derivations" of upper-level behavior from lower level theories that make use of some upper-level information (e.g. information about how lower-level entities behave when there are lots of them rather than just a few, upper-level information encoded in constraints and boundary conditions) fall within the ambit of an extended notion of reduction. Two observations about this:

1) There is not necessarily anything wrong with such maneuvers but they do involve reformulating our understanding of what counts as deduction in much more permissive directions, under pressure from examples like those discussed by Bob. Moreover, if the goal is to show how the upper level behavior can be explained from the resources of the lower level theory alone, it is not clear that this goal is met if one helps oneself to upper level information in the deduction.

2) Merely focusing on whether there is deducibility in this expanded sense elides various distinctions we may wish to make-- for example, is the reduced theory one we can obtain from the reducing theory merely by an averaging procedure or is something more complicated, making use of correlational or broadly topological information required? Put differently, it is not as though the only interesting question is whether the upper-level theory can be connected in some way or other with the lower level, underlying theory. It is also interesting

whether the connection is of such a character that the upper level theory has substantial autonomy or not with respect to he lower level theory. Whether reduction holds in some broad sense does not settle this second question.

11. Conclusion: Why Meso? I conclude by returning to some questions that I raised at the beginning of these remarks. First, consider again the role of the middle or meso -level. Bob emphasizes the importance of this in connection with the physics examples he discusses, with the meso-level serving as a crucial bridge or intermediary between more micro and macro theorizing. As he remarks in passing other scientists have had similar ideas-- perhaps most notably Denis Noble in his work on modeling the heart. Noble describes a strategy he calls "middle out" which (very roughly) consists in beginning at a middle level, which in his case is the modeling of the circuits controlling the beating of the heart and then working "down" from there to a more molecular level-- e.g. to an understanding the ion channels underlying these circuits and also "up" to levels encompassing the behavior of the whole heart— the mechanics of its overall pumping activity and so on. He suggests that this strategy is more likely to be effective in understanding the behavior of the heart than either one that starts at a molecular level and attempts to work up, without any guidance from upper-level information or one that is exclusively top-down.

One can imagine a certain kind of philosopher of science who is impatient with the suggestion that there is anything special or indispensable "in principle" about middle levels. The argument might go as follows: Suppose there is a mapping M1 from the lower level to the middle level and a mapping M2 from middle to upper. Then (it might be claimed) taking the composition of these two mappings will take us directly from the lower to the upper, without any need for a stop along the way at the middle. So Bob's middle way is dispensable in principle. Needless to say, this is not a line of argument that I find convincing. When presented with some salient feature of scientific practice I think that one of the first duties of philosophers of science is to try to understand why that feature is present-- what the rationale if any is for the feature or what strategy underlies its presence. This contrasts with the impulse to construct arguments showing that the feature is dispensable in principle-- even if there is such an argument, it won't show why it is that investigators do not dispense with the feature in practice. (This tendency to write as though the only interesting question to ask concerning some feature of scientific practice is whether it is dispensable in principle rather than instead asking why the feature is present in the first place in the actual practice of science is one of many curious aspects of contemporary philosophy of science.)

Thus as I see it the question we should ask about the middle way is why focusing on it is sometimes an effective research strategy when we are trying to relate models or theories at different levels. Why do we need the middle rather than attempting to go directly from bottom to top, micro to macro? We should also generalize this question since in many cases there is no single middle level but instead a hierarchy of levels relating lower/ micro to upper/ macro but for ease of exposition I will often write in what follows as though there is a single middle level[9].

What I am going to say next is suggested by some of Bob's observations but a lot of it is sheer speculation. I hope you (the audience) may help to make it better. My suggestion is that middle levels are important (or one kind of case in which they are important is) when the large-scale organization or connections of the micro units-- how they relate to each other in terms of patterns of correlations, spatial/geometrical relations including connections that can transmit various quantities (currents) and so on-- have important influences on a system's behavior. This also includes facts about such "topological" features as basins of attraction and fixed points in dynamical landscapes -- biochemical, genetic, neural and so on. The units themselves may be largely homogeneous (e.g. the components of a fluid) but they also may not, as with Bob's examples from materials science. In cases of this sort because the large- scale relations between units matter, simple strategies of averaging, aggregation, use of mean field models etc will not work because they do not incorporate such relational or structural information. Moreover, just considering the behavior of the individual units themselves taken in isolation or the behavior of some small number of these is likely to be uninformative about the collective behavior of many units because it is only at the latter scale that the role of what I have been calling organizational or topological features becomes important. Of course this is not to say that the laws governing the individual units are violated or rendered inoperative when large collections of such units are present. At least sometimes we can combine upper-level information about the structure or organization of the collectivity with lower level information about the behavior of the individual units to explain the upper level behavior, as Bob's

[9] See, e.g Herz et al. who describe XX distinct levels of modeling of neuronal behavior. One issue which I do not discuss in what follows but is surely relevant is that different forms of mathematical representation may be most appropriate or natural at different levels and this can create non-trivial problems in extracting information from one level which is relevant from another. For example, Herz et al describe a lower level multicompartment level of analysis of neuron behavior where the appropriate mathematics is partial differential equations, a whole neuron level where ordinary differential equations are approriate, a black box input out level where Bayesian modeling may be most appropriate and so on. One reason why intermediate levels are needed is that it may be particularly difficult to go directly from the mathematics employed at a lower level to the mathematics employed at a much higher level without going through intermediate levels,

examples and many others illustrate. The point, however, is that in order to do this, we need input about structure and organization of the sort captured at the meso level-- information about the behavior of the units, either taken individually or in some very small collection does not suffice. The upshot is what Philip Anderson describes in his classic paper, "More is Different". Once we observe the upper-level behavior and get some insight into the structures with which it is associated, we may be able to use this information to trace the behavior back to lower-level laws but this is very different from being able to start just with information about the behavior of individual units and derive the upper-level behavior. (Once again I urge you to resist the temptation to claim that the latter "must be possible in principle" and focus instead on why it does not happen in practice.)

To what extent can we generalize any of this beyond the physics examples Bob discusses? I don't know but will make the following observations. In many areas of science-- physics but also biology, neurobiology, and social science-- "networks" are all the rage. There are models of gene networks, networks of protein interactions, network models of the brain at many different levels of analysis, network models of mental illness and so on. A goal of such models is to understand how large-scale features having to do with the connectivity or topology of such networks influences various sorts of upper-level behavior-- of entire genomes, brains etc. A practical motivation for such network models is that starting entirely from the bottom up-- e.g with the behavior of individual neurons or the biochemistry of individual proteins-- and somehow aggregating to get the upper level behavior of a large collective is unlikely to be a successful strategy. There are just too many units, the low level details of their interactions are too complicated and moreover the individual contributions of the units may be very small . The hope is that some important features of the upper-level behavior of these systems-- not all of their behavior but some of it --- is driven to a substantial degree by large scale facts about connectivity, patterns of correlation and so on of a sort that might be captured in a network analysis[10]. Thus we get the suggestion that various mental illnesses involve abnormal patterns of connection between various

[10] Collin Rice points out to me that straightforward homgenization techniques like those described for material science are also sometimes appropriate in dealing with biological complexity. But I wonder whether we can't think of many network analyses themselves as employing a kind of homogenzation. Such analyses commonly abstract away from what is distinctive about the behavior of individual units , the details of their causal connections or even, when undirected graphs are used, assumptions about the direction of the causal connections among units. The network representation homogenizes lots of things that are different in an overall representation that one hopes contains useful information about the behavior we observe.

brain areas-- sometimes connections that are too strong rather than connections that are too weak-- a reaction to the pretty much complete failure so far to find common variant genes that have any substantial effect on the incidence of such illnesses, the failure to find chemical imbalances (too little serotonin, too much dopamine) that are single "causes" mental illnesses and so on.

As another illustration, consider contemporary macro-economics, in the form of dynamic stochastic general equilibrium theories of the economy. These attempt to go directly from alleged micro-foundations having to do with individual choice behavior (this is actually more than a bit of a con) to the overall behavior of the entire economy-- that is, the meso level is entirely neglected. The result has been some spectacular failures-- failures to predict (and even to explain) the collapse of financial institutions, recessions, inflation and so on. It is of course possible that there is no approach that will work better but a not implausible conjecture is that more middle level modeling may be helpful. Such modeling might focus, for example, on particular institutions and their interactions-- financial institutions and their interrelations, market sectors like housing and so on. These are much larger than individual people or households but smaller than the whole economy. Here network level analysis and patterns of connections and correlations are likely to be crucial-- connections among financial institutions that can contribute to financial contagion and cascades of failures, patterns of correlation among housing prices in different part of a country and so on. (The naive expectation that housing prices and mortgage defaults would move independently of one another in different parts of the country contributed substantially to the 2007 recession-- a striking example of neglect of an important meso-level correlation.)

12. Concluding Unscientific Postscript/Other Things to Talk About

There are several further issues raised by *MW*. There is a prominent philosophical view about the status of upper-level generalizations in the so-called special sciences that, prima-facie at least, seems quite different than what is suggested by Bob's explorations. This is the so-called "Mentaculus Vision" developed by Loewer and Albert. (e.g. XX) It consists of the following three components: (i) the fundamental dynamical laws, (ii) the Past Hypothesis, according to which the macro state M(O) of the universe shortly after the Big Bang was one of very low entropy, and (iii) a Statistical Postulate, according to which "there is a uniform probability distribution specified by the standard Lebesgue measure over the physically possible microstates that realize M(0)". Loewer and Albert claim that all of the generalizations of the special sciences follow (that is, are derivable, "in principle") from these assumptions. They do not, as far as I know, explicitly discuss the examples that figure in *MW* but presumably they think that e.g, the Navier-Stokes equations, the universal features of critical point behavior for various substances and so on, all follow from the Mentaculus Vision assumptions.

My question is how, if at all, such claims relate to Bob's discussion in *MW*. Very roughly, I take Bob to be claiming that there are stable upper-level generalizations because (or to the extent that) various lower-level micro-details don't matte , where this presumably has to do (at least in large part) with the way in which the micro-level laws relate to macro-level generalizations. Loewer and Albert don't talk about autonomy and related matters but to the extent that there are stable upper-level generalizations, I interpret them as claiming that this has a lot to do with the fact that certain initial conditions obtained in the early universe—a low entropy macro-state and a uniform probability distribution over the micro states that realize this initial low entropy state[11]. In *MW* Bob, by contrast, does not talk about the role of special initial conditions in generating stable upper level behavior. Needless to say, Loewer and Albert don't actually exhibit (and there is no reason to think it is possible to exhibit) the derivations they associate with the Mentactulus while (to his great credit) Bob shows in detail how stable upper-level behavior can emerge from lower level laws. But putting that consideration aside, is there anything more that can be said about how the two approaches relate?

A natural interpretation of Loewer and Albert is that they hold, not just that the various upper level generalizations follow from their assumptions but that those assumptions are *needed* for those generalizations (and others-- e.g, concerning entropy) to be true— in other words, if the Past Hypothesis and the Statistical Hypothesis did not hold, those upper-level generalizations would not hold. What should we make of such a claim? Presumably if the universe was initially in a high entropy state it would have remained in such a state, with the consequence that various macro structures with which we are familiar would not have existed— no fluids, brains, ice cubes etc. But even if this is so, it seems to give us no insight into why the upper-level generalizations governing the various macro structures take the particular form that they do. Moreover, we should distinguish two different claims: i) if initial conditions in the early universe or later on were to take a very special or "unusual" form, then given the fundamental laws, the macro-level generalizations would be very different (or maybe there would be no macro-level structures for macro-level generalizations to be about) vs ii) the macro-level generalizations and structures we see around us are very sensitive to the precise initial conditions that obtained in the early universe and perhaps subsequently in

[11] Maybe this is a misinterpretation. I'd be happy for any corrections

the sense that if those conditions had been even slightly different, everything macro would behave differently. Unlike ii), i) is consistent with (iii) it's being the case that for "most" initial micro level initial conditions meeting some broad set of (perhaps macro) constraints, we see macro-level behavior similar to what we actually see. To the extent that iii) is true, we have a kind of independence or autonomy of upper level behavior from the exact details of the distribution of micro- level initial conditions (although not of course independence from relevant macro-level initial conditions.) iii) does seem to characterize our world to some substantial degree and is apparently required for the relative autonomy of upper level generalizations. It is unclear to me what Loewer and Albert think about ii) vs iii)— the derivability of upper level generalizations from PH and SP does not by itself tell us anything about how sensitive those generalizations are to the precise details of PH and SP. For example do Loewer and Albert think that *any* non-uniform distribution over the micro-states consistent with PH would lead to different upper level generalizations? If not, perhaps they should replace SP with the claim that the distribution over micro-states is not highly "special" or "atypical".

References

Ay and Polani (2008) "Information Flows in Causal Networks" *Advances in Complex Systems* 11, 17-41.

Gao "Two Approaches to Reduction: A Case Study from Statistical Mechanics"

Lin, Tegmark, Rolnick "Why does deep and cheap learning work so well?

Polger and Shapiro. *The Multiple Realization Book.*

Wallace "Stating Structural Realism: Mathematics First Approaches to Physics and Metaphysics"

Woodward, "Explanatory Autonomy: The Role of Proportionality, Stability and Conditional Irrelevance

Eronen

Potochnik

Shapiro and Polger

Wallace XX