Australasian Journal of Philosophy
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713659165

Lewis, Carroll, and seeing through the looking glass

John Roberts [a]
[a] University of Pittsburgh,

Online Publication Date: 01 September 1998

To cite this Article: Roberts, John (1998) 'Lewis, Carroll, and seeing through the looking glass', Australasian Journal of Philosophy, 76:3, 426 — 438

To link to this article: DOI: 10.1080/00048409812348541
URL: http://dx.doi.org/10.1080/00048409812348541

Australasian Journal of Philosophy
Vol. 76, No 3, pp. 426–438; September 1998

LEWIS, CARROLL, AND SEEING THROUGH THE LOOKING GLASS

John Roberts

Much of David Lewis's work consists of a large-scale defense of the thesis of Humean supervenience; a crucial part of this defense is his elaboration and defense of the best-system analysis of laws of nature, which makes laws out to be Humean supervenient.[1] Much of John Carroll's work consists of a campaign against reductive analyses of laws of nature; one element of this campaign is his attack on the idea that laws of nature are Humean supervenient.[2] Carroll's main argument against the Humean supervenience of laws is called *the Mirror Argument.* The aim of this paper is to argue that while the Mirror Argument does succeed in bringing out a surprising consequence of the thesis of Humean supervenience about laws, it fails as a refutation of that thesis.

The Mirror Argument aspires to be a completely general argument against Humean supervenience about laws; it is not specifically an attack on Lewis's best-system analysis. However, the availability of the best-system analysis will play a role in my objection to the Mirror Argument. My conclusion, in brief, will be that the Mirror Argument cannot successfully refute Humean supervenience about laws unless it is supplemented with an independent argument against the best-system analysis.

I. Humean Supervenience and the Best-System Analysis of Laws

For Carroll, *Humean supervenience* (henceforth, HS) is the thesis that the complete set of non-nomic features of a given possible world settles all the facts in that world. *HS about laws* is the more specific thesis that the complete set of non-nomic features of the world settles what the laws of nature are; in other words, there are no two possible worlds that agree on all non-nomic details but have different laws of nature.[3] The non-nomic features of a possible world, for Carroll, are all those that can be stated without reference to the concepts of laws, counterfactuals, causation, and other 'close cousins' of these notions that share the 'modal character' of laws.[4]

1 Lewis [7], [8], and [10].
2 Carroll [4] and [5], Chapter 3.
3 Carroll [5], p. 58. Lewis's definition of Humean supervenience differs from Carroll's in two important ways. First, in place of 'the non-nomic features of the world', Lewis has 'the spatio-temporal arrangement of local qualities'. Second, for Carroll, HS is, if true, necessarily so, whereas the thesis Lewis defends under the name 'Humean supervenience' is supposed to be merely contingently true. However, where Lewis makes this claim, it is the general thesis of HS he is talking about, rather than the specific thesis of HS about laws of nature. See, for example, [8], p. x, and [10], pp. 474–5. In this paper, I'll always use 'HS about laws' in Carroll's sense rather than Lewis's since it is Carroll's argument I am criticizing. It is worth noting that Lewis's best-system analysis of laws (about to be discussed in the text) implies HS about laws in Carroll's sense as well as his own.
4 Carroll [5], pp. 7–8, 57–8.

Lewis's best-system analysis of laws (henceforth, the BSA) runs as follows: The laws of nature are those generalizations that belong to the best deductive system of propositions all of whose members are true.[5] The *best* such system is the one that achieves the best balance of strength and simplicity, where strength, simplicity and balance are to be measured according to our present standards (or the present standards of our scientists). The true deductive systems that are to be regarded as candidates for best are required to specify only non-nomic features of the world—they are not to include any propositions about what is a law, what causes what, etc. This analysis entails HS about laws, since the non-nomic features of the world fix the true deductive systems, and then our standards of strength, simplicity, and balance (which are to be considered as fixed) determine which system is best.[6]

II. Tooley's Thought Experiment

HS about laws is a very strong negative thesis: It claims that there do not exist any two possible worlds that match with respect to the non-nomic details but have different laws of nature. So one way to argue against it is simply to try to describe a pair of possible worlds that constitute a counterexample. This strategy is employed by Michael Tooley,[7] and is also used by Carroll.[8] However, as we'll see, Tooley's argument is not conclusive. The Mirror Argument is presumably intended to be an improvement on Tooley's strategy. My complaint against the Mirror Argument will be that its apparent advantages over Tooley's argument are illusory, so that it must be just as inconclusive as the latter. For this reason, it is important to explain what is wrong with Tooley's argument.

Tooley's argument asks us to consider a description of a possible world that specifies ten types of fundamental particles. Thus, there are 55 pairs of kinds of particles, and so 55 types of possible two-particle collisions. The history is specified so as to contain many examples of 54 of these 55 kinds of two-particle collisions, but no collisions of the 55th kind, which consists of collisions between X-particles and Y-particles. Finally, the possible world is described as having a set of laws of nature including laws of free particle motion as well as laws governing two-particle collisions of the 54 types that actually occur. This world-description is not to include any law of nature that rules out X–Y collisions; it's just that the initial conditions, together with the specified laws, preclude any X–Y collisions, 'by accident' as it were. Evidently, this is a coherent world-description, but it seems that the description is not complete, because there appears

[5] Strictly speaking, this gives only Lewis's analysis of deterministic laws of nature; in order to extend the analysis to cover statistical laws, Lewis must modify it a bit (see [10], pp. 480–2). In this paper I'll ignore this complication.

[6] One very common objection raised against the BSA maintains that in making our present standards a fixed element in the analysis of laws of nature, the BSA makes laws out to be either subjective or arbitrary. A discussion of this objection would reach beyond the scope of this paper; Lewis gives one reply to such objections in his [10] (on p. 479), and I give a different reply in [11].

[7] Tooley [13], pp. 669–72. Tooley uses this example to argue for the claim that the truth-makers of law statements must not be facts about particulars, a conclusion that could reasonably be construed as a way of denying HS.

[8] Carroll [4], pp. 202–6. Here, Carroll adopts Tooley's thought experiment in an argument aimed directly at the BSA (pp. 202–3), but Carroll's argument, if successful, proves (more generally) that HS is false.

to be room for a fundamental law governing X–Y collisions, and by hypothesis nothing in the description settles what this law should be. We could supplement the description by specifying that it is a law of nature that whenever an X–Y collision occurs, an event of type P occurs, or we could supplement the description by specifying that when such a collision occurs, an event of type Q occurs, where Q and P are inconsistent event-types. These two supplementations would result in descriptions of two possible worlds that agree on the non-nomic details but differ on the laws of nature. So, it may be argued, the two possible worlds described constitute a counterexample to HS.[9] .

But there are serious questions about the legitimacy of this kind of argument. How do we know that the worlds described in cases like Tooley's really are genuine possible worlds? Mightn't it be the case that not every way of supplementing Tooley's world-description with a particular X–Y collision law answers to a genuine possible world? (In particular, why couldn't it be the case that in any world answering to Tooley's original description, there just is no objective matter of fact about what would happen if an X–Y collision occurred, and so no law of nature that governs such collisions?)

Someone sympathetic to Tooley's argument would insist that we can be confident that each of Tooley's two world-descriptions, complete with their specified X–Y collision laws, describe genuine possible worlds. On what grounds could this claim be defended? Evidently, there is nothing logically inconsistent about either world-description. But it isn't clear that every logically consistent world-description picks out a genuine possible world; for example, Kripke argues that, although it is logically consistent that heat is not the kinetic energy of molecules, there is no possible world in which this is true.[10] It might be urged that Tooley's world-descriptions seem intuitively unproblematic, and that *prima facie* there is no good reason to suppose that they do not represent genuine possibilities. This could be questioned for many different reasons, though. First of all, strong and unambiguous intuitions about what the laws would or could be like in other possible worlds are not always universally shared, and it would be dangerous to treat them as if they were conclusive. After all, many of us share a very strong pre-philosophical intuition that it is possible for heat to be distinct from molecular kinetic energy; if Kripke's views on necessity and identity are correct, then they serve as a stern warning against putting too much trust in our pre-theoretical modal intuitions. Tooley's argument appeals directly to intuitions that concern not only modality, but also the perplexing concept of nomicity, so it would seem that we have even more reason to be wary of these intuitions.

Even if we grant the intuitive plausibility of Tooley's two putative worlds, there are a number of reasons that might be put forward for doubting the veridicality of these intuitions. Philosophers attracted to HS about laws are aware this view has some counterintuitive consequences; presumably, they maintain that HS has theoretical advantages that trump the intuitions it offends against. For example, they might hold that it is desirable to maintain that the laws are determined by the objective facts about the world, and that it is objectionable to admit the notion of an objective, non-logical necessary connection in nature, and that the only way to satisfy both of these considerations (which, taken jointly, may be thought to trump many of our pre-theoretical

[9] For other arguments along similar lines, see Armstrong [1], p. 71, and Carroll [4], pp. 211–7.
[10] Kripke [6], p. 99.

intuitions) is to affirm HS.[11] Nor is this the only source of doubt about Tooley's worlds: Some philosophers hold that the identity conditions of properties and kinds include the laws of nature in which they figure.[12] A consequence of this is that two possible worlds with different laws of nature could not contain instances of the very same kinds or instantiations of the very same properties. For such philosophers, Tooley's two world-descriptions provide only an apparent counterexample to HS: A kind of particle that exists in the first world, and a kind of particle that exists in the second world, have both been given the label 'X-particles', even though considerations about the identity of properties show that the two kinds could not be identical, and hence that the non-nomic histories of the two worlds could not be exactly the same.

The issue concerning Tooley's thought experiment thus seems to come down to the question of which is more plausible: the intuitions that favor Tooley's putative possible worlds, or the theoretical considerations that motivate HS. The task of resolving this issue directly looks terribly difficult. Intuitive plausibility and theoretical advantage are both valuable features in a philosophical view, and the trade-offs between them that are sometimes necessary are difficult to negotiate; we may have here a case of incommensurable values. One way to begin to resolve issue would be to produce a compelling argument against the skepticism concerning non-logical necessary connections in nature that often motivates advocates of HS. Carroll attempts to do just this in his [5], but his argument includes the Mirror Argument as a crucial component, and as I'll argue in section IV, the Mirror Argument is incomplete as it stands.

A more workable strategy might be to try to show that if a defender of HS rejects a Tooley-esque thought experiment on the grounds that it doesn't represent a real possibility, then we can rationally lead (or push) her into a position that is manifestly implausible, even by her own lights, without making any illicit appeal to any modal or nomological intuitions that we can be sure she will not respect. I take it that this is what the Mirror Argument aims to do. Beginning with some very intuitive premises about a couple of possible worlds that seem to be acceptable, even by the lights of someone inclined to favor HS, Carroll argues, relying on standard ideas about counterfactuals and possible worlds, for the existence of a distinct pair of possible worlds that constitute a counterexample to HS. If a defender of HS is going to dismiss the latter pair of putative possible worlds as spurious, then she is going to have to reject one or more of the seemingly uncontroversial premises from which the Mirror Argument begins. This would be a dubious move that (Carroll argues) should lead us to reject her dismissal.

This is a promising strategy; however, as I'll argue below, it doesn't quite work. In order to get us to accept all the premises of the Mirror Argument, Carroll has to get us to believe a contentious proposition that no defender of the BSA can be reasonably expected to accept. This contentious proposition will turn out to be extremely similar to the proposition that Tooley's two world-descriptions represent genuine possible worlds. Carroll might insist that we should accept this premise on the grounds of its intuitive

[11] To be fair, I should note that the intuitions that HS and the BSA are widely thought to offend against are not limited to those involved in Tooley's argument. In particular, it has been argued that irreducibly statistical laws raise a great difficulty for defenders of HS. Lewis discusses and responds to this difficulty in [10].

[12] Swoyer [12]; Bigelow, Ellis and Lierse [3].

plausibility, but if this move is made, then we find ourselves back at the very same stand-off between intuitions and theoretical considerations described above. For this reason, the Mirror Argument turns out to be no more convincing than Tooley's argument.

III. The Mirror Argument

The Mirror Argument begins by asking us to consider two possible worlds, U_1 and U_2. Each world contains X-particles and Y-fields. The proposition that all X-particles in Y-fields have spin up is baptized L_1; L_1 is a law in U_1 but not in U_2.[13] In each of these worlds, each X-particle's trajectory takes it into and then out of one Y-field, and none of these particles will ever enter another Y-field. In U_1, since L_1 is a law, it follows that each of the X-particles has spin up while inside a Y-field. In U_2, however, L_1 is not a law and it is not even true, for one of the X-particles (called 'particle b' in each world[14]) has spin down while traversing its Y-field. The two worlds are assumed to be identical in all other non-nomic details. In particular, in each world, particle b passes by a mirror just before it enters its Y-field. The mirror's position is called 'position c'. The mirror is mounted on a well-oiled hinge which it makes it possible to rotate it into a position called 'position d' in which it blocks the path of particle b. Further, in each world, the laws governing particle motion are the same, and in particular these laws imply that a particle that approaches a mirror in position d with the trajectory of particle b will be deflected in a certain direction, and in U_1 and U_2, nothing lies in that direction. In short, we would find it intuitive to say that, in each world, if the mirror were in position d when particle b came by, then b would be deflected in such a way that it would zoom off without ever interacting with anything else again.

Are U_1 and U_2 plausible enough to count as 'real' possible worlds? Carroll argues persuasively that they are. It is possible to fill in the details as richly as we want, so long as we keep U_1 and U_2 qualitatively identical in all non-nomic details except for the differences already noted, and so long as we don't add anything otherwise inconsistent with Carroll's descriptions.[15] It may seem strange that in each of these worlds, a particular type of particle never encounters a particular type of field more than once. But it isn't clear just how strange this is; though it may seem unlikely, it isn't clear why we should even think that this sort of situation isn't to be found in our own world. Is it reasonable to

[13] Carroll writes as if a field were a localized item that a particle can either be inside or outside of, rather than a function defined over all the points in a space. This is harmless, however; we can take 'Y-field', in Carroll's usage, to mean a region in which the value of the Y-field is greater than some value V (or within some range R).

[14] Carroll might be using 'b' as a rigid designator, i.e. a referring term that picks out the same particle in each of the possible worlds under consideration. But if so, then this can be dropped without significantly changing the argument. All that is necessary is that there be an X-particle with the specified features in each world under consideration, and that in each world this particle is picked out by 'b'. So critics of the notions of rigid designation and trans-world identity needn't be worried by this feature of Carroll's argument.

[15] Actually, Carroll specifies that U_1 and U_2 are empty except for the X-particles, the Y-fields, and the mirror. But he claims that details of U_1 and U_2 could be filled in without affecting the argument. 'It could all take place in an isolated portion of an ordinary universe, one much like ours, that includes people, tables, and all sorts of things.' (p. 65) It seems to me that Carroll can go even further: The X-particles and Y-fields need not be in an isolated portion of a universe, so long as the facts mentioned above are all exactly as specified.

suppose that, given Carroll's descriptions of the non-nomic details of U_1 and U_2, that the laws of nature in these worlds could be as he describes them? Carroll notes that someone might object that U_1 and U_2 could not agree on all of their laws except for L_1, on the grounds that laws of nature are necessarily deeply interdependent (p. 66). But even if we grant the interdependency of laws, we don't necessarily have an objection here:

> We surely want to acknowledge that there are worlds that partially overlap on their laws. Also, nothing in my argument requires that the only difference in the nomic structure of U_1 and U_2 be that L_1 is a law of U_1 and not of U_2. In these two worlds, there may be two entirely different networks of laws governing particle/field interactions. So, to challenge my argument, some very specific connection needs to be established between laws governing the motions of X-particles and particle/field interaction laws like L_1. ([5], p. 66)

So long as U_1 and U_2 don't differ on any non-nomic details except for the spin of particle b as it traverses its Y-field, Carroll is willing to let the laws of these worlds differ as much as is necessary to secure that L_1 is a law in U_1 but not in U_2. To sum up, if there is an objection to the admissibility of U_1 and U_2 as possible worlds, then it doesn't appear to be an obvious one. There may well be a damning not-so-obvious objection, but I want to move on. I'll assume that worlds U_1 and U_2 are in fact possible.

Carroll now appeals to the highly intuitive claims that in U_1 and in U_2, it is physically possible that the mirror be in position d when particle b comes by, and that in each of these worlds, it is true that if the mirror were in position d when particle b came by, then b would sail off on a certain trajectory T, never again to interact with anything else.[16] Thus, there is a possible world that would result from taking U_1 and modifying it so that the mirror is in position d when b comes by; Carroll dubs this world U_1*.[17] In U_1*, then, b hits the mirror and is deflected into trajectory T, and never again interacts with anything else. Similarly, there is a possible world, called U_2*, that results from taking U_2 and modifying it so that the mirror is in position d when b comes by; in U_2*, b hits the mirror and is likewise deflected into trajectory T. U_1* and U_2* are evidently identical in their non-nomic features.

But, according to Carroll, U_1*and U_2* differ in their laws of nature. In arguing for this claim, Carroll relies on the following pair of principles:

> (SC*) If P is physically possible and Q is a law, then Q would (still) be a law if P were the case.
>
> (SC′) If P is physically possible and Q is not a law, then Q would (still) not be a law if P were the case.[18]

[16] I'm introducing 'T' as a convenience; it isn't used by Carroll.

[17] Of course, strictly speaking, there are many possible worlds that would result from modifying U_1 in this way, even though Carroll speaks of U_1* as '*the* world that would result if the mirror were in position d in U_1' ([5], p. 63, my italics). Carroll clearly intends U_1* as a label for the closest world to U_1 in which the mirror is in position d (or (presumably), in the case of a tie, an arbitrary member of the set of closest such worlds—although Carroll might think that there can be only one). Similarly for world U_2*, which is about to be described in the text.

[18] (SC*) and (SC′) are defined at p. 59 of Carroll [5].

From (SC*) it follows that in U_1, the following counterfactual is true:

If the mirror were in position d, then L_1 would (still) be a law

and from (SC′), it follows that in U_2, the following counterfactual is true:

If the mirror were in position d, then L_1 would (still) not be a law.

Now, given the standard possible-worlds semantics for counterfactuals, it follows from this that, since U_1* is the closest possible world to U_1 in which the mirror is in position d, L_1 must be a law in U_1*. By parallel reasoning, L_1 must not be a law in U_2*. So U_1* and U_2*, while agreeing on all of their non-nomic details, have different laws of nature. Therefore, Carroll concludes, HS about laws is false.

Of course, Carroll could have just started by describing worlds U1* and U2*. But in that case, he would have had to ask us just to accept two possible worlds that agree on the spatio-temporal arrangement of local qualities, but in one of which L_1 is a law, whereas in the other L_1, while true, is not a law. This would be to rely on the kind of appeal to intuitions about what kinds of possible worlds there are that, as I argued above, is unsatisfying. In order to make the argument effective, Carroll has to prove that the two possible worlds at issue are in fact acceptable, even to a reasonable defender of HS and the BSA who would not be inclined to accept a pair of possible worlds like U_1* and U_2* on the force of modal and nomological intuitions alone. That is the point of introducing U_1 and U_2.[19] The acceptability of U_1 and U_2, together with (SC*), (SC′) and the consequences drawn from them above, is supposed to entail the acceptability of U_1* and U_2*. And finally, (SC*) and (SC′), as well as the acceptability of U_1 and U_2 are supposed to be things that can be reasonably taken to be true independently of any appeals to intuition that are not likely to impress a defender of HS and the BSA. In this way, the Mirror Argument appears to attain a rational authority over defenders of HS and the BSA that Tooley's ten-particle argument does not.

IV. The Trouble with the Mirror Argument

The Mirror Argument depends crucially on two premises that Carroll does not explicitly recognize. Let 'H' denote the total world-history that results from taking the (non-nomically characterized) history of U_1, altering the position of the mirror from c to d,[20] keeping the trajectories of all X-particles except b and the locations of all Y-fields exactly the same, modifying the motion of particle b so that after it collides with the mirror it goes into trajectory T, and leaving all other non-nomic details exactly the same. In other words, H is the world-history, completely but non-nomically characterized, that

[19] Carroll [5], pp. 64–5.

[20] This modification in the mirror's position should be thought of as holding throughout all time, or as coming to be at some time prior to particle b's encounter with the mirror, whichever is deemed less disruptive to the history of U_1. If there is more than one way to specify the way in which the mirror comes to be in position d that are equally non-disruptive, then we can let H contain an arbitrarily selected such way.

Carroll assumes to obtain in both U_1* and U_2*. Now we can state the following premises, each of which is crucial to the Mirror Argument:

1. In U_1, it is physically possible that the total history is H.
2. In U_2, it is physically possible that the total history is H.

Premises 1 and 2 are not explicitly affirmed by Carroll—Carroll does argue forcefully that in U_1, it is physically possible for the mirror to be in position d, but this is a weaker claim than premise 1—so why do I call 1 and 2 crucial premises of the Mirror Argument? Carroll claims (and he must claim, in order to run the Mirror Argument) that the world U_1* is physically possible relative to U_1, and that U_2* is physically possible relative to U_2. Furthermore, in arguing for his claim (also crucial to the Mirror Argument) that U_1* and U_2* agree on all non-nomic details, he walks through a construction of the non-nomic details of these worlds, allowing as little divergence from worlds U_1 and U_2 as is necessary in order to allow the mirror to be in position d; in walking through this construction, he assigns to each of U_1* and U_2* the non-nomic history I've labeled H.[21] Hence, the Mirror Argument depends on premises 1 and 2. For, if 1 were false, then U_1*, as constructed by Carroll, would not be physically possible relative to U_1, and similarly for premise 2.

In fact, the Mirror Argument requires more than the conjunction of premises 1 and 2. It requires also that the notion of physical possibility used in these premises be interpreted in a certain way. We can distinguish between three interpretations of physical possibility as follows. First, on what may called the *intuitive interpretation* of physical possibility, proposition P is physically possible in world W iff we can imagine or conceive of P's being the case in W without offending any of our physical intuitions. Second, on what I'll call the *weak interpretation*, P is physically possible in world W iff P is logically consistent with the truth[22] of all of the laws of nature in W. Finally, on what I'll call the *strong interpretation*, P is physically possible in W iff there exists a possible world, having exactly the same laws of nature as W, in which P is true. Carroll understands physical possibility in terms of the strong interpretation, whereas Lewis understands it in terms of the weak interpretation.[23]

The Mirror Argument is sound only if premises 1 and 2 are both true in the strong sense of physical possibility. For unless this is the case, (SC*) and (SC′) may not both be true in the cases in which Carroll applies them. If 1 and 2 are true only in the weak sense of physical possibility, there might be only one possible world with total history H: The proposition that the total history of the world is H would then be physically possible (in the weak sense) in both U_1 and U_2, since H is consistent with the truth of the laws of both worlds. However, it would not then be the case that (SC*) is true in U_1 and (SC′) is true in U_2, because that would entail that in the one possible world with total history H, L_1 is both a law and a non-law. (SC*) and (SC′) are both true, then, only if the strong interpretation of physical possibility is presumed.

[21] [5], pp. 63–4.

[22] But not necessarily the nomicity!

[23] [5], p. 18, note 8 (here Carroll defines physical necessity as the dual of physical possibility in the strong sense); Lewis's endorsement of the weak interpretation is entailed by his remarks on p. 20 of his [9]. Concerning the weak interpretation, see also van Fraassen [15], pp. 43–5.

An advocate of the BSA has a very straightforward reason to deny that 1 and 2 could both be true given the strong interpretation of physical possibility. To see why this is so, consider *any* possible world W in which the complete non-nomic history is H. Among the deductive systems that are true in this world, either there is, or there is not, a unique one that achieves the best balance of strength and simplicity.[24] If there is such a system, then for Lewis, the generalizations belonging to it are the laws of nature in the world W, and L_1 either does or does not belong to this system. Since W was chosen as an *arbitrary* world with total history H, it follows that L_1 either is a law or fails to be a law in *all* worlds having this total history, including U_1* and U_2*. If L_1 is a law in all such worlds, then it is a law in U_2*, so that it is not, after all, true that U_2* has exactly the same laws as U_2. Thus, in U_2, it is not physically possible, in the strong sense intended by Carroll, for the total world-history to be H, because any world with history H will not have exactly the same laws as U_2.[25] In other words, premise 2 is false in this case. Similarly, if L_1 fails to be a law in all worlds with total history H, then it will not be a law in world U_1*, and premise 1 will be false. On the other hand, if there is no unique deductive system that achieves the best balance of strength and simplicity, then on Lewis's analysis, there will simply be no fact of the matter about what the laws of nature are in world W. In this case, both premise 1 and premise 2 will be false: There will be no possible world with exactly the same laws as U_1 (or U_2) that has the right non-nomic facts, because any world with the right non-nomic facts will have no determinate set of laws of nature at all. In any case, then, if we assume the strong reading of physical possibility and the BSA, then at least one of premises 1 and 2 will be false.[26] (Which one(s) is (are) false will depend on the details of our standards of strength, simplicity, and balance.)[27]

[24] In his original version of the BSA [7], Lewis requires only that there be a definite class of deductive systems that are tied for best, and he defines the laws of nature as those generalizations in the intersection of all of these equally-best systems. In more recent work [10], Lewis revises this, and requires that there be a uniquely best system if the concept of a law of nature is to have an unambiguous extension. The remarks in the present paragraph can easily be modified in obvious ways to accommodate the earlier version of the BSA.

[25] However, in any such world, the laws of U_2 will still be *true*.

[26] A similar argument could be given substituting for Lewis's view any other theory of laws that entails HS. For example, anyone who finds the naive regularity account intuitively appealing will surely find premises 1 and 2 intuitively offensive on the strong reading of physical possibility, since it is obvious that they are jointly inconsistent with this view of laws. Similarly, an advocate of an epistemological account of laws (cf. Ayer [2] and Urbach [14]), according to which laws of nature are those regularities for which we can have evidence of a certain kind, will object to the conjunction of 1 and 2 (given the strong reading of physical possibility) on the grounds that in any H-world (that is, any possible world with total history H), the available evidence for the 'law' L_1 would be exactly the same, so that its nomic status in the two worlds would have to be the same.

[27] Defenders of the BSA may disagree as to which of premises 1 and 2 are false, since they may disagree on how our standards of strength, simplicity, and balance would apply in the case of a world with history H. It seems to me, however, that the best case can probably be made for the view that L_1 *is* a law in any such world. Recall that we are presupposing that L_1 is a law in U_1, so that the total non-nomic history of U_1 is such that including L_1 in a deductive system greatly increases its strength while not sacrificing too much simplicity. Now, the strength–advantage of L_1 in a world in which H is true is evidently virtually as great in as it is in U_1; moreover, it is hard to see how any disadvantage with regard to similarity that it suffers in an H-world would not be equally present in U_1. Given these facts, it is tempting to say (if somewhat facetiously) that the only reason why L_1 fails to be a law in U_2 is that it has a single counterexample there. Since this reason is lacking in any H-world, it seems that L_1 would have to be a law in any such world. Thus, on my view, it is premise 2 that a defender of the BSA ought to reject.

Thus, the conjunction of 1, 2 and the strong interpretation of physical possibility can be true only if the BSA is false. The root reason for this is that on the BSA, it is not in general the case that smooth and intuitively non-jarring distortions to the non-nomic history of a world will leave the laws of nature exactly the same, and this in turn follows from the fact that on the BSA (as on any regularity analysis of laws), the laws are determined by the total pattern of events in the world, so that a strategically placed change in the latter can result in a change in the former. Hence, the issue seems to come down to the question of which is more plausible: the premises of the Mirror Argument (including the hidden premises 1 and 2, and including the strong interpretation of physical possibility), or the philosophical motivations behind the BSA.

What might a friend of the Mirror Argument say on behalf of the conjunction of 1 and 2 given the strong interpretation of physical possibility? The evidence on behalf of 1 and 2 seems to consist solely in an appeal to widely shared intuitions about what would be physically possible in worlds U_1 and U_2. These intuitions are admittedly very appealing, and it is doubtful that an advocate of HS or the BSA could plausibly deny their force. But these intuitions only concern *what is physically possible*; they do not seem to militate in favor of a particular interpretation of the notion of physical possibility (such as the strong interpretation). The point can perhaps best be put in the following way. The evidence in favor of 1 and 2 consists of highly plausible statements of physical possibility, given the *intuitive interpretation* as defined above. But the premises themselves can support the Mirror Argument only if physical possibility is construed according to the strong interpretation. So the Mirror Argument needs to rely on two inferences of the form 'P is physically possible in the intuitive sense at world W; therefore, P is physically possible in the strong sense at world W'. Whether such inferences are valid or not obviously depends on which interpretation of physical possibility best represents the modal status that our physical intuitions can be assumed to be sensitive to. A defender of the BSA needn't reject the physical intuitions that support premises 1 and 2; all she needs to do is say that whatever notion of physical possibility these intuitions reliably indicate is weaker than the strong interpretation of physical possibility. The Mirror Argument, then, can be convincing only if we can be given some reason for thinking that our physical intuitions are reliable guides to physical possibility *in the strong sense*. This is a very strong claim: It implies that our pre-philosophical modal and nomological intuitions give us access to information about strong existence claims concerning possible worlds specified in both nomic and non-nomic terms. (For example, it implies that our untutored intuitions are sensitive to the existence or non-existence of a possible world in which H is the total history of the world and the laws of nature are exactly are same as those in U_2.) Carroll gives no argument in favor of this claim.

Consider what would be implied by an insistence that our intuitions underwrite premises 1 and 2, together with the strong interpretation of physical possibility. Such an insistence would entail commitment to 3:

> (3) There exist two possible worlds, each of which has total (non-nomic) history H, such that L_1 is a law of nature in one of them but not in the other.

This claim is very similar in form to the crucial premise of Tooley's ten-particle argument, namely that there are two possible worlds having an identical total non-nomic history, but

which differ in their fundamental laws. So, since the Mirror Argument is sound only if this claim is true, it is difficult to see what advantage this argument has over Tooley's argument, which was criticized above.

One more tack a defender of the Mirror Argument might try is the following:

> U_1 and U_1* are stipulated so as to differ only on their initial conditions. We generated U_1* by taking its stipulated initial conditions and evolving them according to the laws of U_1. Now, since U_1* is stipulated precisely as the world you get by evolving these initial conditions according to the laws of U_1, isn't it unnatural to suppose that U_1* has any other laws than those of U_1? Similar remarks apply for U_2 and U_2*. So U_1* and U_2* must be distinct worlds with different laws of nature.

The problem with this argument is that it presupposes a certain picture of the way the laws of nature figure in the world, a picture which any Humean is bound to reject. Namely, it presupposes that there are two basic genetic factors that determine the history of the world: the laws of nature and the initial conditions. The initial conditions are thought to be arbitrary, brute facts, which determine a starting point for a dynamic evolution that will be governed by the laws. This idea is going to be found unacceptable by anyone sympathetic to a neo-Humean account of laws of nature such as the BSA. Such a philosopher is likely to reply that when we 'generate' U_1* by taking its initial conditions and evolving them according to the laws of U_1, the 'generation' in question is not the ontological generation of a world's history within that world itself, but rather the 'generation in thought' of a world-history, a process of mental construction. There may be more than one way to 'generate in thought' a single world, so the fact that we generate two possible worlds by mentally evolving a stipulated set of initial conditions according to two distinct sets of stipulated 'laws' does not automatically guarantee that the worlds we thus generate are really distinct—for the two distinct sets of 'laws' are not features of the worlds but rather features of the ways that we generated them in thought. Analogously, two different constructions of a geometrical figure do not necessarily result in two distinct figures. So this way of bolstering the Mirror Argument cannot be reasonably expected to cut any philosophical ice for someone attracted to the theoretical advantages of the BSA.

In conclusion, then, assuming Carroll's strong interpretation of physical possibility, premises 1 and 2 of the Mirror Argument are both acceptable only if we can be sure that Lewis's BSA can be ruled out. Again, this is ultimately because according to the BSA (or any sophisticated regularity analysis of laws), the total pattern of non-nomic events determines what the laws are (rather than the other way around), so that an imagined modification in the non-nomic history of a possible world that strikes us as smooth and does not bother our physical intuitions can nonetheless, if it is strategically placed, result in a possible world with different laws of nature from those of the world we started with. This is, perhaps, a consequence that doesn't sit well with all of our common pre-theoretical intuitions about laws of nature. The Mirror Argument, then, serves to bring out vividly one of the ways in which the BSA is intuitively surprising. But, here again, the intuition that is being offended against is a mixed modal-nomological intuition, and a defender of HS and the BSA is already going to be willing to accept the consequence that some of these intuitions will have to be revised. And here again, a bare appeal to such an intuition does not constitute a compelling argument against HS (or against the BSA).

So, the argument is convincing as a refutation of HS only if it can be supplemented with an independent argument against the BSA.[28] The Mirror Argument itself, if sound, would refute Lewis's view, since it would rule out any theory of laws according to which the nomic is Humean supervenient, but this argument cannot be used to establish one of its own premises on pain of circularity. Hence, it appears that the Mirror Argument is either unconvincing or incomplete as it stands.

V. Conclusion

There is a moral to be learned here. Philosophical discussions of laws of nature typically take the central problem to be that of the notion of a law, and pay much less attention to the notion of physical possibility. It is frequently assumed that the notion of physical modality is rather straightforward, once the notion of a law has been clarified.[29] This is a dangerous assumption to make. As we have seen, there are at least two quite distinct conceptions of physical modality that have been advocated in the literature—Carroll's conception, which I have been calling the strong interpretation of physical possibility, and Lewis's conception, which I have been calling the weak interpretation of physical possibility. Hence, even if the concept of a law is fully clarified, this still leaves some room for disagreement concerning physical modality. It may seem that this disagreement would be merely terminological. But this isn't the case, simply because many important arguments that have been advanced concerning laws of nature turn on appeals to our intuitions about what is physically possible and what is not. Before we can adequately evaluate such arguments, we need to know exactly how much is warranted by these intuitions. For example, if there is widespread intuitive agreement that a certain state of affairs (e.g., the mirror's being in position d) is physically possible, then does this warrant us in presuming that the state of affairs is in fact physically possible, where physical possibility is understood according to some particular technical rendering? If so, then which of the competing technical renderings is the right one? Does the intuition tell us that the thing is possible in *some particular* sense, or does it just tell us that it is possible (in *some* sense, which we must now try to explicate)? Sound methodology would seem to require that we test our analyses of both the notion of a law and the notion of physical possibility against such intuitions at once, rather than simply assuming an analysis of physical possibility in order to test an analysis of laws. More generally, all of our analyses must face the tribunal of our intuitions as a whole, rather than one by one.[30]

University of Pittsburgh

Received: June 1997.
Revised: February 1998.

[28] In fairness to Carroll, he does offer an independent argument against the BSA, in both [4] (pp. 202–6) and [5] (see especially Chapter 2). This argument is based on the objection to the BSA mentioned in note 6. I argue that this argument fails in [11].

[29] This assumption is suggested by remarks of van Fraassen in his [15], p. 45. Notably, however, van Fraassen takes for granted Lewis's weak interpretation of physical possibility, rather than the strong interpretation taken for granted by Carroll.

[30] I would like to thank Joe Camp, John Earman, and two anonymous referees for many helpful comments on earlier drafts.

REFERENCES

1. David M. Armstrong, *What is a Law of Nature?* (Cambridge, Cambridge University Press, 1983).
2. Alfred J. Ayer, 'What is a Law of Nature?' *Revue Internationale de Philosophie* 10 (2) (1956), pp. 144–65.
3. John Bigelow, Brian D. Ellis, and C. Lierse, 'The World as One of a Kind: Natural Necessity and Laws of Nature', *British Journal for the Philosophy of Science* 43 (1992), pp. 371–88.
4. John Carroll, 'The Humean Tradition', *Philosophical Review* 99 (1990), pp. 185–219.
5. ——, *Laws of Nature* (Cambridge: Cambridge University Press, 1994).
6. Saul Kripke, *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1972).
7. David Lewis, *Counterfactuals* (Oxford: Blackwell, 1973).
8. ——, *Philosophical Papers*, Volume 2 (Oxford: Oxford University Press, 1986).
9. ——, *On the Plurality of Worlds* (Oxford: Blackwell, 1986).
10. ——, 'Humean Supervenience Debugged', *Mind* 103 (1994), pp. 473–90.
11. John Roberts, 'On David Lewis's Alleged Chauvinism', unpublished manuscript, 1998.
12. Chris Swoyer, 'The Nature of Natural Laws', *Australasian Journal of Philosophy* 60 (1982), pp. 203–23.
13. Michael Tooley, 'The Nature of Laws', *Canadian Journal of Philosophy* 7 (1977), pp. 667–698.
14. Peter Urbach, 'What is a Law of Nature? A Humean Answer', *British Journal for the Philosophy of Science* 39 (1988), pp. 193–210.
15. Bas van Fraassen, *Laws and Symmetry* (Oxford: Clarendon Press, 1989).