

Setting Asymmetric Dependence Straight

Mark Greenberg, UCLA¹

1. Introduction

Fodor's asymmetric-dependence theory of content is probably the best known and most developed causal or informational theory of mental content. It has generated a substantial critical literature.² Much of that literature attempts to generate counterexamples to Fodor's theory. In this paper, I offer a more fundamental critique.

Fodor's theory is cast in terms of laws covering the occurrence of an individual thinker's mental symbols. Rather than attempting to identify the content-determining applications of a symbol by specifying situations in which a thinker would apply the symbol only to instances of its reference property, Fodor maintains, roughly speaking, that the content-determining laws are those on which the other laws governing the mental symbol asymmetrically depend (in a sense to be elaborated).

My central argument (section 5) is that, even if the laws that the theory requires obtain, the laws do not in fact exhibit the appropriate asymmetric-dependence relations.³ In a nutshell, I show that, in general, part of the mechanism that underwrites the crucial, supposedly content-determining law for a mental symbol is not shared by the mechanisms for the other laws covering the occurrence of the same mental symbol. As a result, the former law can be eliminated (by eliminating the non-overlapping part of the

¹ I am very grateful to Martin Davies, Gilbert Harman, Ram Neta, Paul Pietroski, Georges Rey, and Galen Strawson for extremely helpful comments and discussions.

² See, e.g., Loewer and Rey (1991).

³ Note that Fodor holds that the asymmetric-dependence conditions are sufficient, but not necessary, for content. Consequently, he insists that he need consider only cases in which it is stipulated that the conditions are satisfied. In section 4, I show that this position cannot be maintained.

mechanism) without eliminating the latter laws. The latter laws do not asymmetrically depend on the former law.

Before I can reach my main arguments, I need to do some stage setting. In the next section (2), I outline the asymmetric-dependence theory. In section 3, I discuss Fodor's views about laws, which are crucial to an understanding of his theory. Next, in section 4, I examine and dispute Fodor's claims about what work he needs to do to support his theory and about what kinds of objections are legitimate. My substantive critique is contained in section 5.

2. *The Asymmetric-Dependence Conditions*

Fodor's conditions for a linguistic or mental symbol "S" to mean X are (pp. 93, 121):⁴

- 1) It is a law that Xs cause the tokening of "S"s⁵ (information condition); and
- 2) if it is a law that any Ys (that are not Xs) cause "S"s, then that law is dependent on the law that Xs cause "S"s, but not the other way around (asymmetric-dependence condition).⁶

There are two further conditions that have an uncertain status in Fodor's account:

- 3) Some "S"s are actually caused by Xs (actual-history condition).

⁴ In this paper, all citations to page numbers alone are to Fodor 1990.

⁵ For convenience, I will use Fodor's loose way of talking according to which a letter or word in quotation marks can stand for either a linguistic or mental symbol, or can be short for an occurrence or tokening of the symbol, as in *horses cause "horse"s* (shorthand for *horses cause occurrences of the symbol "horse"*).

⁶ It should be noted that Fodor's asymmetric-dependence account assumes (at least some elements of) the language-of-thought hypothesis, for which he argues elsewhere (e.g., 1987, pp. 135-154; 1975). I will assume the language of thought hypothesis throughout this paper in order to address Fodor's theory on its own ground.

4) For some property Y (not = X), it is a law that Ys cause “S”s – *i.e.*, the asymmetric-dependence condition is not satisfied merely vacuously (robustness condition).⁷

The first two conditions form the core of the theory. They are Fodor’s way of distinguishing a symbol’s reference from other properties that reliably cause the symbol’s occurrence – in other words, they provide Fodor’s solution to the disjunction problem. Specifically, the two conditions implement Fodor’s basic notion that of all the properties that are nomically related to (the property of causing occurrences of) a symbol, the symbol’s reference is the property that figures in the nomic relation on which all the other nomic relations depend.

We will not need to be much concerned with the actual-history condition. Fodor tentatively proposed it as a way of preventing the theory from implying implausibly strong verificationist consequences (pp. 119-24),⁸ but, as described below, he has subsequently repudiated it.

The idea behind the robustness condition is that it distinguishes semantic from (merely) informational nomic relations (pp. 92-93). True symbols are “robust,” which means that they retain their meanings despite being reliably caused by properties other than those to which they refer (1990, p. 91). If, to put it loosely, it is a law that a property causes an event of a certain type,⁹ then events of that type are natural indicators and carry information about the property. (*E.g.*, it is a law that fire causes smoke, so smoke carries

⁷ See Fodor 1990, p. 118; 1991a, p. 261.

⁸ Roughly, the actual-history condition is intended to make ineligible for a symbol’s reference any property that in fact has no causal impact on the thinker. This is designed to avoid the consequence (to which the theory apparently otherwise would be committed) that in order for “X” to mean X, it must be nomologically possible for the thinker to distinguish X “from any property that *would* cause “X”s *if it were instantiated*” (p. 122).

⁹ For expository ease, I will follow Fodor in using terms such as “law,” “property,” “cause,” and “event” in a relaxed way, writing, *e.g.*, of properties causing events, or of laws that Xs cause Ys.

information about fire.) Thus, a natural indicator is not robust – it means whatever reliably causes it. In contrast, Fodor thinks, it can be a law that a property causes (occurrences of) a symbol, despite the fact that the symbol does not refer to that property. Symbols are robust – they retain their meaning, despite their occurring (both mistakenly and non-mistakenly) without being caused by their reference. (For example, “horse” means HORSE even if it is a law that cows on dark nights cause “horse” tokens and a law that thoughts of horses do.)

Fodor’s theory presents a moving target.¹⁰ The original formulation of the theory (1987, pp. 106-111) and its initial presentation in *A Theory of Content (TOC)* (1990, p. 89-100) include only the information and asymmetric-dependence conditions. Robustness enters the discussion in TOC as a feature of content that a theory must be able to explain (1990, pp. 90-93, 97-100), but in the course of the essay Fodor responds to an objection by treating the robustness condition as (a necessary) part of his sufficient conditions for a symbol’s having content (p. 118). At the end of the essay, the robustness condition is not mentioned in the canonical specification of the theory (pp. 119-124). In subsequent work, however, Fodor continues to appeal to the robustness condition (1991a, p. 261). As for the actual-history condition, Fodor introduces it tentatively, then includes it in the canonical statement of the theory and treats it as a firm part of the theory (pp. 119-124), and, as noted above, finally rejects it (1994, pp. 115-119). My discussion focuses on the two principal conditions, but I will point out where appropriate that Fodor’s account is not saved from my objections by either the robustness or actual-history condition.

¹⁰ In this paragraph, I am leaving aside significant shifts in Fodor’s position from his 1987 to his 1990, see, e.g., 1990, pp. 131-132, fn. 3, as well as further changes that are implicit in his 1998, pp. 120-145.

3. Fodor on laws

Fodor thinks that appealing to laws, as opposed to counterfactuals, will enable him to avoid problems that are fatal to other informational accounts. It will be useful to examine Fodor's views about laws in order to clarify the framework within which his theory of content is situated. I will accept this framework, at least for the sake of argument.

3.1 Preliminaries

We can begin with a few preliminary clarifications and terminological points. On Fodor's view, reliable causal relations are subsumed by laws. Laws hold of properties, not individuals. As Fodor does, I will often use "it is a law that As cause Bs" as a shorthand for "there is a nomic relation between the property of being an A and the property of causing Bs."¹¹

Suppose Qs reliably cause Rs – that it is a counterfactual-supporting generalization that Qs cause Rs. It follows that there is a law that subsumes the causal transactions between Qs and Rs, but it is a further question whether the property of being a Q figures in that law. For example, if Qs are a subkind of Ts, Qs may cause Rs in virtue of being Ts rather than in virtue of being Qs. That is, the covering law may be that Ts cause Rs. Fodor gives the example that, though small horses reliably cause occurrences of the mental symbol "horse" (in a particular thinker), the covering law may be that horses cause "horse"s, not that small horses do (pp. 101-102).

The laws that figure in Fodor's theory are not paradigms of laws of nature. A law of the sort Fodor relies on is a nomic relation between a worldly property and a property

¹¹ See generally, Fodor 1987, p. 164 fn. 6; Fodor 1990, pp. 93-95.

of causing the occurrence of a particular mental symbol in an individual thinker's head. (A mental symbol here must be understood as a type, tokens of which can occur or be triggered on particular occasions.) An example is a nomic relation between the property of being a star and the property of causing the occurrence of a certain mental symbol – call it “star” – in my head (in Fodor's shorthand, the law that stars cause “star”s). Such a law is specific to an individual thinker's mental symbol; there is no suggestion in the theory that there are community- or language-wide laws. One notable way in which Fodor's laws are peculiar is their specificity to particular thinkers. A related oddity is that the laws are easily changeable. Learning, aging, and injury can create and eliminate laws.

A number of further terminological shorthands will be convenient. If there is a law covering the occurrence of some mental symbol, such that all other laws covering the occurrence of that symbol asymmetrically depend on that law, I will say that the law is *primary*. Suppose we are considering the laws covering the triggering of my mental symbol “S.” We can refer to the law that As cause “S”s as “the A law,” the law that Bs cause “S”s as “the B law,” and so on. (In this terminology, Fodor's theory of content holds, very roughly, that “S” means *X* in virtue of the *X* law's being primary.)

When the reference of “S” is A, we can also call the A law “the reference law” or “the content-determining law” and the other laws “wild laws.” Wild laws include laws covering occurrences of “S” that are not *applications*¹² of “S”, for example, occurrences of “S” that are caused by other thoughts. Such occurrences of a symbol need not be in any way mistakes. I will for the most part set aside this qualification until section 5,

¹² An “application” of a symbol to an object is a use of a symbol to characterize a presently perceived object as falling within the symbol's extension. To apply “horse” to a horse is to *call* a horse “a horse,” not merely to token “horse” in the presence of a horse (Fodor 1990, p. 97).

where laws covering non-application occurrences of symbols will become important to my argument.¹³

3.2 *Ceteris-paribus* laws as a response to Kripke's doubts about dispositions

The laws that figure in Fodor's asymmetric-dependence theory are obviously high-level laws, not the basic laws of physics. Such laws are not, in general, exceptionless;¹⁴ they have implicit *ceteris-paribus* clauses.¹⁵

Fodor explains his appeal to *ceteris-paribus* laws by considering Kripke's (1982) argument that ideal conditions will not help to solve the Kripkenstein problem. After pointing out that our actual dispositions are imperfect, Kripke considers the possibility of appealing to *ceteris-paribus* dispositions. Even if a thinker would in fact make mistakes when apparently adding, perhaps what is relevant is what he would do other things being equal. Kripke rejects this idea on the ground that we are not justified in making assumptions about what would happen under all the varying circumstances encompassed by the *ceteris-paribus* clause. For example, he argues, we "have no idea" what a thinker would do if given enough extra brain capacity to handle very large numbers; he might "go insane" (Kripke 1982, p. 27). In effect, Kripke is pointing out that the ideal conditions to which dispositional accounts appeal are impossible conditions – conditions under which we have infinite memory and time, no distractions, and so on – and that evaluating the truth of counterfactuals with impossible antecedents is problematic. In

¹³ See also note 25 below.

¹⁴ I use the term "exception" in the most obvious sense, to refer to any case in which the antecedent of a law or putative law is satisfied, but the consequent is not, even if that case is consistent with the law's obtaining (for example, because the case is accounted for by interfering factors). This contrasts with the way in which Pietroski and Rey (1995), discussed below, use the term. In their terminology, an apparent exception is not an exception if it is consistent with the law's obtaining. Thus, in their terminology, even *ceteris-paribus* laws are exceptionless (1995, pp. 87-88).

¹⁵ See Fodor 1991b.

possible-worlds terms, the closest possible worlds in which the physically or nomologically impossible conditions held might well be worlds that were wildly different from our world.

Reasonably enough, Fodor thinks that this objection is too swift. In his view, Kripke's mistake is to assume "that we can't have reason to accept that a generalization defined for ideal conditions is lawful unless we can specify the counterfactuals which would be true if the idealized conditions were to obtain" (p. 92). Fodor points out that in general it cannot be true that we have to be able to specify all that would happen if ideal conditions were to obtain in order to be justified in accepting that a generalization holds, *ceteris paribus*. In other words, we can be justified in relying on a *ceteris-paribus* generalization without being able to cash the *ceteris-paribus* clause into counterfactuals. The idea is that we can know how a system would behave in the absence of external interference, without knowing what all the possible interfering factors are – and, *a fortiori*, without being able to replace the *ceteris-paribus* clause with fleshed-out counterfactuals.¹⁶ We can have reason to accept that you would add very large numbers correctly if the limitations of memory did not interfere, without having to worry about what would actually happen if you had extra brain matter.

¹⁶ See pp. 93-95 & fn. 10. Fodor follows an unpublished manuscript of Georges Rey, which was a predecessor of Pietroski and Rey (1995), in maintaining that an ineliminable *ceteris-paribus* clause does not make a law vacuous. As Pietroski and Rey spell out the proposal, a *ceteris-paribus* law is true only if all cases in which the law apparently fails to hold can be explained by the interference of *independent* factors. In order to be independent, a factor "has to explain not only the apparent counter-example to the [*ceteris-paribus* law], but also something other than that failure" (1995, p. 90). As discussed in note 18 below, Pietroski and Rey's proposal does not address the conditions for something to be a law; rather, it concerns only a sufficient condition for non-vacuity (pp. 91-93, 98-100).

3.3 *Laws and mechanisms*

Fodor thinks we can avoid the problematic counterfactuals by appealing directly to laws. He prefaces his account with an affirmation of faith that it is ontological “bedrock that the world contains properties and their nomic relations”: nomic relations between properties are deeper than and therefore need not be analyzed in terms of counterfactual relations among individuals (p. 93). He is even more sure of the epistemological point: we can know that a particular nomic relation obtains without knowing which corresponding counterfactuals hold. We can know that the ideal gas laws hold even if we have no idea what would happen under the physically impossible conditions the laws specify (pp. 93-95). Fodor’s strategy is therefore to frame his account in terms of nomic relations, and then to resist, on these theoretical grounds, the challenge to specify all the corresponding counterfactuals. For example, he thinks he can maintain that the property of being a horse is nomically related to the property of causing “horse”s (in a particular thinker), without establishing what the thinker would do and say under various specified conditions.

Suppose we grant that *ceteris-paribus* laws need not be vacuous and are (ontologically and epistemologically) prior to counterfactuals. There remains the question of when a regularity counts as a law. We saw that Fodor’s appeal to *ceteris-paribus* laws is a response to the problem that the correlations between instantiations of properties and occurrences of symbols are imperfect. When faced with an apparent counterexample to a candidate law, Fodor need not find some way of specifying ideal conditions under which the counterexample will not occur. As long as mistakes can be attributed to other things not being equal, they do not undermine the existence of the law.

Of course, Fodor does not attribute every apparent counterexample to other things not being equal. He implicitly assumes a distinction between situations in which “other things” interfere with a law’s operation and situations in which there is no law. Fodor’s laws, as non-basic, are mediated, or underwritten, by mechanisms. To say that other things are not equal is to say that something prevents the mechanism from operating in the way that it otherwise would. Thus, the distinction Fodor needs is between exceptions to a law – situations in which a law’s mechanism is intact but conditions prevent its operation – and situations in which part or all of a law’s mechanism is lacking and thus the law does not obtain.

In an article that Fodor (1990 p. 132, fn. 8) cites approvingly, Paul Pietroski and Georges Rey (1995)¹⁷ explain the role of *ceteris-paribus* laws in scientific explanation as follows: “scientists state [*ceteris-paribus* laws] in an attempt to focus on particular factors . . . and thereby ‘carve’ complex phenomena . . . in a theoretically important way” (1995, p. 92). On this picture, most phenomena in the real world are the result of the complex interaction of many different systems. Science proceeds by considering isolated systems in abstraction from other external factors. It explains the phenomena as the result of the interaction of these different systems. Thus, the question whether a generalization that is subject to apparent counterexamples is nevertheless a law is the question whether the best scientific explanation of the phenomena – of the cases in which the generalization holds and the cases in which it does not – holds that the apparent counterexamples are the consequence of interactions between an underlying system whose behavior corresponds

¹⁷ As noted above, Fodor cites what was then an unpublished manuscript of Georges Rey, which was a predecessor of Pietroski and Rey (1995).

to the generalization and factors external to that system. In other words, whether a regularity is a law depends on whether there is an intact underlying mechanism.

I don't know how to give a rigorous account of the distinction between absence of a mechanism and interfering factors, and Fodor offers no help.¹⁸ But it is clear that Fodor's appeal to *ceteris paribus* laws requires such a distinction. Also, as we will see, Fodor implicitly relies on the idea that eliminating a law's mechanism eliminates the law.

3.4 Asymmetric dependence

Central to Fodor's theory is the notion of asymmetric dependence. For law B to be asymmetrically dependent on law A is for law B to depend (ontologically) on law A but not *vice versa*. For law B ontologically to depend on law A is for the obtaining of law A to be part of what makes law B obtain.

The relevant kind of dependence is not historical or across time – “diachronic” – but “synchronic” (p. 134 fn. 18; 1987, p. 109).¹⁹ In other words, the question is not whether the existence (or elimination) of one law would, over time, lead to the existence (or elimination) of another, but whether, at a particular point in time, one law

¹⁸ Fodor (p. 132, fn. 8) quotes a passage from Rey that, taken out of context, might sound as if it offers a way of spelling out sufficient conditions for a generalization to be a *ceteris-paribus* law: “The viability of a *ceteris-paribus* clause depends not upon the actual specification or realizability of the idealization, but rather upon whether the apparent exceptions to the law can be explained as due to independently specifiable interference.” Pietroski and Rey emphasize, however, that they do not purport to offer sufficient conditions for a generalization's counting as a *ceteris-paribus* law, but only to offer sufficient conditions for the non-vacuity of a *ceteris-paribus* clause, on a narrow reading of vacuity. See, *e.g.*, 1995, pp. 98-99. Accordingly, they define an “independent” factor as one that explains something other than the failure of the generalization (in a sense that they specify more precisely) (pp. 89-90). Given this definition, the condition that any exception to a generalization be explained by an independent factor, cannot be (and does not purport to be) a sufficient condition for the generalization to count as a law. As Pietroski and Rey write, as far as their proposal goes, “every singular causal claim of the form A caused B is a candidate for an interferable [*ceteris-paribus* law]” (1995, p. 99). They make no attempt to distinguish situations in which a putative law lacks a mechanism (and thus the law does not obtain) from situations in which the law's mechanism exists but is interfered with.

¹⁹ See text accompanying note 39 below.

ontologically depends on another – in other words, whether the latter law’s obtaining is part of what now makes the former law obtain.

Just as laws are more basic than counterfactuals, and thus are not analyzable in terms of them, the same is true of the relations between laws (pp. 93-95). Thus, strictly speaking, counterfactual formulations do not capture what dependence (or asymmetric dependence) is – rather, the dependence is what makes the counterfactuals true.²⁰ In practice, however, Fodor often treats the claim that law B is asymmetrically dependent on law A as equivalent to the counterfactual (indeed, counterlegal) claim that, *ceteris paribus*, if law A did not exist, law B would not exist (or, as he often puts it, that, *ceteris paribus*, eliminating law A would eliminate law B), but not the converse.²¹

Subsequently, in response to an objection from Paul Boghossian, Fodor offers a different formulation according to which law B’s asymmetric dependence on law A amounts to the following: other things being equal, you would have to change more to eliminate law A without eliminating law B than to eliminate law B without eliminating law A (1991a , p. 276). (Also, although he resists the idea that possible-worlds translations are needed (p. 95; 1991a , p. 271), Fodor offers such translations and often relies on them (113-116; 1991a , p. 276-277).²²)

²⁰ In his 1987 (pp.108-109), Fodor offered the counterfactual formulation as an analysis of asymmetric dependence. In his 1990, he resists the idea that it is an analysis, but, as explained in the text, he nevertheless works with it in practice.

²¹ See, e.g., pp. 91, 93, 107, 112-113, 115-116. (“*but that “cow” tokens carry information about cows, they wouldn’t carry information about anything;*” “No noncow-caused “cow”s without cow-caused “cow”s;” p. 91 (italics in original).)

²² In earlier discussions, Fodor often works with possible-worlds formulations of asymmetric dependence – there is some variation among them – that hold, roughly, that law B asymmetrically depends on law A if and only if there are (nearby) possible worlds in which A obtains and B does not, and they are nearer to our world than any world in which B obtains and A does not (1987, p. 109; 1990, pp. 113; 1991a , p. 276). When Fodor introduces the later understanding of asymmetric dependence described in the text, he explains that, in possible worlds terms, the claim that law B asymmetrically depends on law A should be understood as the claim that, other things being equal, worlds in which A holds but B does not “are closer to us than *corresponding* worlds” in which B holds but A does not (1991a , p. 276, emphasis added). As we will see,

It may be harmless for most purposes to follow Fodor in treating the dependence of law B on law A as tantamount to the counterfactual truth that, *ceteris paribus*, law B would not obtain were it not for law A's obtaining. Nevertheless, I want to pause to point out that my argument does not depend on a conflation of asymmetric dependence with the corresponding counterfactuals. First, as Fodor makes clear, the claim that one law is asymmetrically dependent on another *entails* the corresponding counterfactual claim (p. 95). Since I will be arguing that the asymmetric dependences to which Fodor's theory is committed do not hold, it would be sufficient for my purposes to show that the corresponding counterfactuals do not hold. Second, it is difficult to evaluate the truth of the relevant counterfactuals in the abstract. Thus, in order to evaluate the claim that one law depends on another (and also to evaluate the corresponding counterfactual claims), I will examine the mechanisms that underwrite the laws, and the relations between those mechanisms – which, as we will see, is the method that Fodor generally employs. In other words, rather than determining dependences between laws by evaluating counterfactuals, I will determine dependences (and indeed evaluate the corresponding counterfactuals) by considering the relations between the mechanisms that sustain the laws. Thus, the counterfactual formulations will not play an essential role in the argument.

3.5 *The lawfulness of systematic mistakes*

Since Fodor's theory addresses the disjunction problem by holding that wild laws asymmetrically depend on reference laws, Fodor holds that thinkers' dispositions to make

there may be diverse ways of changing the world so as to eliminate law A or law B. The point is that we are to compare each way of eliminating law A with the corresponding way of eliminating B – the way of eliminating B that is most similar to it. As I make clear in section 5, my arguments do not depend on the precise possible-worlds treatment of asymmetric dependence.

mistakes can be lawful. According to his view, if a thinker is systematically disposed to mistake cows on a dark night for horses, there will be a law under which cows on a dark night cause “horse”s.²³ (Of course, the relevant law might not be that cows on a dark night cause “horse”s; the kind that figures in the law might be, for example, horsey-appearing objects. See notes 34-35 below.) This is just an application of his more general position that “it is necessary and sufficient for such reliable causation that there be a nomological – lawful – relation between certain (higher-order) properties of events” (1987, p. 99).²⁴

In fact, it is easy to see that Fodor’s asymmetric-dependence theory would be superfluous (and, because of the robustness condition, its conditions for content would never be instantiated) if properties other than a symbol’s reference were not nomically related to the causing of the symbol. The problem at which the theory is directed is that there will be multiple laws governing the causation of a given symbol. Fodor’s solution is that the law on which the other laws asymmetrically depend is the content-determining law. “Horse” means HORSE because all laws covering the causing of “horse”s by non-horses are asymmetrically dependent on the law that horses cause “horse”s.

²³ Many philosophers seem to have assumed that mistakes are not supposed to be lawful on Fodor’s account. A careful reading leaves no doubt that this is incorrect (though perhaps Fodor is not quite as explicit about this point as he should have been) (*e.g.*, pp. 93, 121-122, 126). Indeed, as explained in the text, the asymmetric-dependence theory would not be needed to address mistaken applications if they were not lawful. In *A Theory of Content*, Fodor manages to confuse even himself with respect to an immediate implication of the nomic status of mistakes. He maintains there that in order for a mistake to be a mistake and not content-determining, it must be nomologically possible for the thinker to avoid the mistake: “An organism can’t have a kind of symbol which it necessarily misapplies, *i.e.*, which it misapplies in every world consonant with its psychology” (p. 108; see also p. 132 fn. 10). For, Fodor argues, if the putatively mistaken application were necessary given the organism’s psychology, the application would be correct (p. 108). Fodor later (1991a, p. 263) recognizes that since the mistaken connections are nomic, it is nomologically (but not physically) impossible for the thinker to avoid them.

²⁴ Fodor thus eschews the implausible approaches 1) of denying the possibility of systematic mistake (see, *e.g.*, pp. 90-91) and 2) of maintaining that a symbol’s connections to its reference are nomic, but that its systematic connections to other properties are not. If laws are the kind of exception-ridden generalizations that Fodor needs them to be, systematic mistakes can be nomic.

Fodor goes further than merely holding that systematic error is a problem that a theory of content must address. As we saw, he conjectures that robustness is what distinguishes a symbol from a mere natural sign. Thus, on his view, the existence of multiple laws governing a symbol's causation is partially constitutive of content (pp. 90-93, 99-100, 117-119, 128-130). Accordingly, Fodor imposes the robustness condition, which requires that, in order for a symbol to have content, there be at least one nomic relation between a symbol and a property other than its reference.²⁵

3.6 High-level laws and unorthodox "natural" kinds

As noted, the laws that figure in Fodor's theory are high-level – "special science" – laws. Fodor cannot and does not take the view that only (natural) kinds of basic, or even relatively low-level, sciences can figure in laws. On his view of content, after all, any property that cannot figure in a law cannot be the content of a (primitive) symbol.²⁶ In order for "red," "messy," "fuzzy," "vase," "rage," and so on to mean what they do, it has to be the case that it is a law that redness causes occurrences of "red," that messiness causes occurrences of "messy," and so on.²⁷ Thus, Fodor's account requires the possibility that properties that are gerrymandered from the point of view of relatively low-level sciences can figure in the relevant laws. For it would be absurd to rule out the

²⁵ The reliable causal relations between properties other than a symbol's reference and the causing of a symbol's occurrence include not only those relations that are intuitively mistaken, but also those that involve non-mistaken associations or chains of thought (pp. 80-81; 90-91). As noted above, for convenience, I often omit this qualification in the present discussion.

²⁶ The qualification "primitive" does not make much difference to the point. Fodor holds that most mental representations for lexical concepts are primitive. And, anyway, it is not plausible that most ordinary concepts are definable in terms of concepts of relatively low-level sciences.

²⁷ As he must, Fodor has explicitly accepted that properties like that of being a shirt or being crumpled can figure in laws (1991a, pp. 256-257). He thus rejects his earlier position (1986) that such properties were anomic. More precisely, according to his 1986, to be an intentional system is to be able to respond selectively to a property "even though no law connects the property it responds to and the selective property of its response" (1986, p. 14). See Antony and Levine (1991).

possibility of having concepts with references that are not physical-, chemical-, or biological-kind properties. On Fodor's account, then, highly non-natural, interest-relative properties will qualify as "natural kinds" – kinds that figure in scientific laws.²⁸

4 Mere sufficient conditions and who gets to stipulate the counterfactuals

Fodor emphasizes that he is making only the limited claim that he provides *sufficient* conditions for a state to have particular content. Although his discussion often hints at larger aspirations, Fodor at least officially does not claim that any actual mental states ever satisfy his conditions. He explicitly leaves it open that all or some of our mental states have their content because they satisfy different conditions (p. 131).

Fodor maintains that since his conditions purport only to be sufficient for content, he can restrict consideration to hypothetical cases in which he stipulates that his conditions are satisfied: "Don't forget, this stuff is supposed to be philosophy. . . . I get to *stipulate* the counterfactuals. It's enough if I can make good the claim that 'X' *would mean* such and such if so and so *were to be* the case" (p. 96). Accordingly, in replying to objections, he protests repeatedly that attempts to show that his conditions do not hold in actual cases are irrelevant (*e.g.*, p. 102; 1991a, pp. 258-59, 263). I have several grounds for rejecting Fodor's view of the dialectical position.

1) It is false that one who offers (mere) sufficient conditions for content need only defend the view in hypothetical cases in which a symbol is stipulated to satisfy the conditions. Fodor seems to fail to recognize that his sufficient conditions for content imply necessary conditions. In order for a thinker to have an attitude with a certain

²⁸ In his 1998, Fodor takes the view that such unorthodox laws (and kinds) are laws about our minds, not genuinely laws about, for example, doorknobs, coat-hangers, or messiness.

content, it is *necessary* that the thinker *not* satisfy Fodor's sufficient conditions for having a different content. If actual thinkers' mental symbols do satisfy his conditions for meaning something other than what they intuitively mean, it would threaten even his limited official claim that he has provided conditions such that, if a symbol ever satisfied them, the symbol would have the indicated content. (For another thing, Fodor's specified conditions for an attitude's having a given content are meant to be sufficient conditions for the attitude's having only that content.) Hence it is not enough for Fodor to show that hypothetical thinkers who satisfied his conditions would have the appropriate contents; he also needs to show that his conditions do not assign intuitively incorrect contents to actual thinkers.

2) We have seen that Fodor takes the position that all he needs to do is to show that a symbol that satisfied his conditions for meaning *X* would mean *X*. He combines this position with a specific view about what is required in order to show that a symbol means *X*. The view is that it is necessary to show only that the symbol would carry information about *X*s and that occurrences of the symbol can be (systematically) caused by non-*X*s. I will now argue that this view – given Fodor's position that he need show only that a symbol that is stipulated to satisfy his conditions would have the appropriate meaning – illegitimately skews the dialectical position in Fodor's favor.

Fodor assumes that if he can give an informational account that allows for symbol tokens that are reliably caused by objects not within the symbol's extension, he has succeeded in naturalizing content (p. 89). He is explicit on this point: "I've tacitly assumed throughout this paper that if you can get a theory of content that squares the intuition that "*X*" means *X* only if "*X*" tokens carry information about *X*-instantiation

with the intuition that “X” means *X* only if you can have *X*-tokens that aren’t caused by *X*s, then you’ve done all that a solution to Brentano’s problem is required to do” (p. 128).²⁹ To put it another way, he assumes that to have both information and robustness is to have content (p. 128).³⁰ Accordingly, the bulk of his argument is concerned with showing that the conditions would allow for error (pp. 89-92, 122, 127-128).³¹ (He also offers some intuitive considerations in support of his theory, which I discuss immediately below.)

We can now see why Fodor’s restricting consideration to cases in which he stipulates the counterfactuals has a peculiar effect when combined with his assumption that showing that his conditions allow for error is all that he needs to do. If we restrict attention to hypothetical cases in which we stipulate the laws (or dispositions or causal relations), allowing for error is too easy. The problem is that we can simply stipulate counterfactual situations in which the person’s erroneous applications differ in some crucial way from the non-erroneous applications. For example, Fodor stipulates counterfactuals in which the person’s erroneous applications asymmetrically depend on his non-erroneous ones.

Suppose, by way of comparison, that my theory is that possessing property *Z* is sufficient for laws (or dispositions or causal relations) to be content determining. By

²⁹ Fodor uses the term “disjunction problem” for the problem of providing such an account, but, as the quotation in the text makes clear, he is using that term to mean only the problem of avoiding the consequence that systematic error is not possible – that is, of finding a theory that makes it possible to have occurrences of a symbol that are caused by *Y*s without it following that *Y*s are instances of the symbol’s reference.

³⁰ Fodor goes on to toy inconclusively and very briefly with the idea that perhaps we need “to throw in” some consciousness or normativity. He also suggests that the psychological might be a superset of the intentional and thus that conditions for intentional content might not be conditions for being in any psychological state (pp. 128-130).

³¹ He also tries to show that he can handle some philosophical problems for dispositional accounts of meaning. For example, in addition to addressing variations on the disjunction problem (pp. 101-110), he tries to show that his account can account for contents that refer to uninstantiated properties (pp. 100-101, 123-124) and that it does not have undue verificationist implications (pp. 119-123).

Fodor's rules, I get to stipulate the counterfactuals. So I will of course stipulate counterfactuals in which all the wild laws lack property Z and all the (intuitively) content-determining ones possess property Z. Consequently, by Fodor's lights, my theory solves the problem of error. Thus, Fodor's rules make it trivial to solve – or more accurately, to avoid facing – the problem of error. To put it another way, the problem of accounting for error is a problem that arises in real cases – it is the problem of finding in actual cases a property that both distinguishes the wild laws from the (intuitively) content-determining ones and is a plausible candidate for a constitutive element of content. If we limit ourselves to stipulated hypothetical cases, we do not address the real problem of accounting for error, but stipulate it away. In sum, in order to evaluate whether Fodor has solved the problem of allowing for error, we need to ask whether wild laws asymmetrically depend on reference laws in actual cases. (Or else we need to recognize that allowing for error in stipulated cases does not show that one's proposed conditions are sufficient for content.)

3) It is revealing to consider Fodor's attempt to explain the intuitions that lie behind his theory. Fodor's discussion has the following structure. First, Fodor considers our actual linguistic practices: "Some of our linguistic practices presuppose some of our others, and it's plausible that practices of applying terms ... are at the bottom of the pile" (p. 97). He goes on to suggest that non-information-carrying symbol tokens asymmetrically depend on the information-carrying ones (p. 98). Next, Fodor points out that the mechanisms that sustain these asymmetric dependences are our linguistic intentions and policies (pp. 98-99). Finally, Fodor speculates that the asymmetric-

dependence relations, rather than the intentions or other mechanisms that mediate them, might constitute content (pp. 98-100).

My point here is not to criticize this argument. Rather, I want to note that Fodor is attempting to support his theory by showing that, in actual cases, asymmetric dependences plausibly fall the way his theory suggests. With some reason, he thinks that this claim lends intuitive support to his speculation that the asymmetric dependences are sufficient for content. (In other places, Fodor similarly tries to show that the asymmetric dependences would fall in the way his theory specifies in actual cases (*e.g.*, p. 115).³² But if he can support his theory by arguing that actual cases exhibit a suggestive pattern of asymmetric dependences, it is surely fair to respond that, in fact, actual cases do not exhibit the appropriate asymmetric dependences. The point is not merely *ad hominem*. There is surely some plausibility to Fodor's idea that the existence of asymmetric dependences in actual cases bears on whether asymmetric dependence is constitutive of content. Why should we think asymmetric dependence has anything to do with content, if we find no asymmetric dependence in the only cases of content with which we are familiar? (My arguments suggest that in the actual world, it is in general not the case that, among the laws covering the occurrence of a symbol, one law is primary. If this conclusion is correct, it at least weakens the case that asymmetric dependence is a plausible candidate for a constitutive account of content (or even for a sufficient constitutive condition for content.)

In light of the foregoing discussion, I therefore consider it fair game, in evaluating Fodor's theory, to ask whether he has plausibly accounted for actual cases of content. Given Fodor's insistence that he purports to provide only sufficient conditions for

³² See also section 5 below.

content, I want to emphasize this point. Thus, once again: *I will not limit the discussion to cases in which the appropriate nomic relations (and dependences between nomic relations) obtain, but will examine whether it is plausible that they do obtain (at least typically) in actual cases in which a certain symbol has a particular content.*

5 *Why the star law is not the primary law*

In this section, I argue that the asymmetric dependences do not fall the way Fodor needs them to. For expository purposes, I will use the example of my (or any person's) linguistic or mental symbol "star." As we have seen, on Fodor's theory, "star" has the content STAR because the following conditions are satisfied.³³ First, it is a law that stars cause "star" to be triggered. Second, there are other, wild laws covering the occurrence of "star." Third, all wild laws are asymmetrically dependent on the star law.

For the purpose of challenging Fodor's assumptions about which laws asymmetrically depend on which, we can grant Fodor the *existence* of the laws. Thus, let us assume that it is a law that stars cause "star"s and that it is a law that, say, relatively slow-moving comets cause "star"s. (Suppose I systematically mistake such slow comets for stars.)^{34,35}

³³ Once again, as argued in section 4, although Fodor's official position is not that his theory explains the contents of any actual person's words or thoughts, it is warranted to explore his theory's plausibility as an account of actual cases.

³⁴ As we've seen (sections 3.5-3.6), Fodor's account depends on there being wild laws for every meaningful symbol. Nothing will depend on the exact details of the example; we could equally use a law that electrical storms, catastrophic cosmic collisions, planets, or moons of certain kinds (or under certain conditions) cause "star"s. My examples, like Fodor's, are somewhat unrealistic for purposes of expository simplicity. In light of two qualifications, however, the examples are less unrealistic than may at first appear. First, as noted, Fodor's theory concerns the contents of symbols of individual thinkers, so our assumption is only that a single thinker systematically makes the relevant mistakes. Second, to say that a connection, whether involving a mistake by the thinker or not, is nomic is not to say that it occurs in an exceptionless way. So there may be a slow-comet law even though slow comets do not cause "star"s when they are close to the horizon or seen through a telescope (say). In fact, it is consistent that it is a law that slow comets cause "star"s and that it is not a law that slow comets close to the horizon cause "star"s. The

As we saw, Fodor's rough test for whether one law asymmetrically depends on another is whether, other things being equal, eliminating the latter law would eliminate the former, but not *vice versa*. Since he thinks that the laws and their dependence relations are the consequence of mechanisms (pp. 99-100; 1991a, p. 313 fn. 8), he answers the test question by examining the mechanisms that he assumes underwrite the relevant laws (pp. 113, 115-116, 117; 1991a, pp. 276-277). The example that Fodor develops most fully concerns why the law that XYZ causes occurrences of "water" asymmetrically depends on the law that water causes occurrences of "water."

My point is that the intention to use "water" only of stuff of the same kind as the local samples has the effect of making its applications to XYZ asymmetrically dependent on its applications to H₂O *ceteris paribus*. Given that people are disposed to treat "water" as a kind term . . . it follows that – all else equal – they would apply it to XYZ only when they would apply it to H₂O; specifically, they would apply it to XYZ only when they *mistake* XYZ for H₂O; only when (and only because) they can't tell XYZ and H₂O apart. Whereas, given a world in which they *can* tell XYZ and H₂O apart (and in which their intentions with respect to "water" are the same as they are in *this* world), they will continue to apply "water" to H₂O and refrain from applying it to XYZ (p. 115).

Fodor uses the same basic form of argument in other kinds of cases (*e.g.*, pp. 112-114, 116). When Fodor introduces his later formulation of asymmetric dependence,³⁶ he gives a useful clarification of the argument: "to get to the [kind of world in which only H₂O causes 'water' tokens], you'd only have to make XYZ and H₂O distinguishable; to get to

argument of this section will offer something toward a more precise understanding of what counts as a (*ceteris paribus*) law and an interfering factor.

³⁵ In my view, it is more plausible that the relevant wild law is not that slow comets cause "star"s but that, very roughly, things with a starry appearance – the appearance that stars have in the *actual* world – do. I don't have space to argue for this point here, but it would only strengthen my argument; the law that starry-appearing things cause "star"s, unlike the slow-comet law, can be eliminated only by making changes in the thinker's semantic dispositions or intentions. (See the discussion in text below.) Moreover, in other work, I argue that stars themselves cause "star"s not in virtue of their being stars but in virtue of their having a starry appearance. If that is right, Fodor's theory does not get off the ground since it is not a law that stars cause "star"s. Thus, in order to grant Fodor the strongest case, I begin with the assumptions in the text.

³⁶ See section 3.4 above.

the [kind of world in which only XYZ causes 'water' tokens] you'd have to both do that *and* alter the mechanisms that underlie our intentions to use 'water' only of stuff that's of the same kind as our local samples." (1991a , p. 277).

We can apply Fodor's argument to the slow-comet/star example. In order for one law to hold, but not the other, I (the thinker) have to be able to tell slow comets and stars apart. (As Fodor does, I'll consider the case of linguistic symbols for simplicity.) If we change things so as to make slow comets and stars distinguishable, but change nothing else, my intentions with respect to the use of "star" (or the mechanisms that underlie those intentions) will ensure that I continue to apply "star"s to stars. But I will no longer apply "star"s to slow comets. In contrast, Fodor thinks, in order to have the slow-comet law but not the star law, we would need both to change things so as to make slow comets and stars distinguishable and to change my intentions with respect to the use of "star."

Fodor's reasoning assumes that my intentions about how to use "star" (and the mechanisms that implement those intentions) are the crucial mechanism underwriting the star law. His view is that the only way to eliminate the star/"star" connection is to change my intentions with respect to the use of "star;" in other words, to change the meaning of "star." Now it is uncontroversial that the various uses of a term depend on its having the meaning that it does. That is, if I changed my intentions with respect to the use of "star," thus changing its meaning, the wild laws, including the slow-comet law, would, as a rule (*i.e.*, barring coincidence), be eliminated along with the star law.

But – and this is the crucial point – there is another way to eliminate the star law, one that does not involve changing our intentions, or anything about our psychology. We can change the stars themselves so that, even when there are no interfering factors – even

when conditions are ideal – the stars do not have the effects that they dependably have in the actual world. For example, we can cool them down or change their composition or surround them with opaque clouds of dust. (I discuss shortly the objection that these changes count as interfering factors.) Since I (the thinker) am unchanged – it is not legitimate to take into account the changes in me that would, over time, come about as I took into account the changes in the stars – such a change in the stars will prevent me from applying “star”s to stars, even when there are no interfering factors. But such a change need not alter the way in which slow comets (and other starry-appearing non-stars) affect us. All the talk of “changes” may be misleading. The point is that there must be worlds in which a non-psychological part of the (actual world) mechanism for the star law does not exist.

Let us take the argument step by step. The crucial first step is that part of the mechanism underwriting the star law (or any nomic relation between a worldly property and the property of causing occurrences of a symbol) must be outside the thinker. Here is the argument for this first step. As we have already seen, (non-basic) laws require mechanisms. The thinker’s intentions cannot be the entire mechanism for the star law, for in order for there to be a real regularity connecting a worldly property to a symbol, the world must do its part. The star and the rest of the world around the thinker have to have certain dependable effects. And the thinker must have certain sensory equipment, beliefs (for example, about how stars look and how they affect various instruments), and intentions about how to use the word “star.”

Fodor recognizes that the world must do its part in order for “star” to mean STAR. First, he emphasizes that nothing purely mental is sufficient for content – there

must be real patterns of causal dependences (pp. 98-99).³⁷ Second, he insists that in order for there to be a law that Xs cause “X”s, there must be a world consistent with the thinker’s psychology in which the thinker can discriminate Xs from all non-Xs that cause “X”s (at least all non-Xs that do so in the actual world) (pp. 122-123). That is, it must be consistent with the thinker’s psychology that she can distinguish Xs from all nonXs (that cause “X”s in the actual world). To say that the thinker must be able to discriminate stars from all other causes of “star”s is to say something about the world as well as about the thinker. The world must be such that stars and non-stars have effects on thinkers that are different enough for the thinker to be able reliably to distinguish them. If intentions were enough for there to be a star law, it would not be necessary for the thinker to be able to discriminate stars from other things.

We have established that part of the mechanism that underwrites the star law must be located outside the thinker. The conclusion that the star law can be eliminated without changing our intentions or anything about our psychology follows immediately. If laws depend on mechanisms, and a (non-redundant) part of the mechanism for the star law is out in the world, then the world can be changed in such a way as to eliminate the mechanism, without changing the thinker (supposing – safely enough – that the external world can be different without a difference in the thinker’s mind).

It might be objected that whatever “change” we make will just be the addition of an interfering factor. The objection cannot be correct. We can see this without having an account of how to draw the distinction between the presence of an interfering factor and

³⁷ Fodor is explicit on this point: “words can’t have their meanings *just* because their users undertake to pursue some or other linguistic policies; or indeed; just because of any purely *mental* phenomenon, anything that happens purely ‘in your head’.... For there to be a relation between “John” and John, something has to happen in the *world*” (pp. 98-99). See also Fodor 1991a, pp. 289-290.

the absence of the mechanism. There must be such a distinction, and that is enough for our purposes. Suppose, for *reductio*, that every change counts as an interfering factor. In that case, removing the mechanism for a law counts as introducing an interfering factor. It follows that in a world in which there is no mechanism for a given law, the law still holds. In other words, if every change counts as an interfering factor, laws do not require mechanisms. Every law obtains in every world; the absence of a mechanism is just an interfering factor. But this is absurd. Thus, for every part of a mechanism, there must be changes that do not introduce interfering factors but eliminate the part of the mechanism.

Since part of the mechanism for the star law is in the world outside the thinker, there must be changes in the world outside the thinker that constitute the elimination of part of the star law's mechanism. Once again, the basic point is just this: since part of the mechanism underwriting the star law must be located outside the thinker, it must be possible to eliminate the law without changing the thinker.

Now we can spell out the conclusion of the argument. If the world were different in such a way that stars did not have the dependable effects that they actually have, it would not be a law that stars cause "stars." Suppose, for example, that, while holding everything else constant, including everything internal to the thinker, we drastically lowered the temperature of the stars or changed the facts of chemistry or optics in a way that prevented light from escaping from the stars.³⁸ As noted, Fodor is clear that such counterfactuals must be understood *synchronically* rather than diachronically (p. 134 fn.

³⁸ Compare Kripke's (1980, p. 118) example in which it turns out that gold is actually blue, and that an optical illusion has caused us to think it is yellow.

18; 1987, p. 109).³⁹ That is, the point is not that if stars had looked different, we would not have applied “star” to them – presumably we would have done so because our beliefs and theories would have developed differently. (Similarly, it is beside the point that if stars changed their appearance we would eventually develop new beliefs and theories that would enable us to apply “star” to stars.) Rather, the point is that we, as we actually are, would not apply “star” to stars if stars were changed in such a way that, even with no interfering factors, they would not have any of their characteristic effects on us or on our instruments.⁴⁰ The same goes for scientific experts. Given the theories and methods they use in the actual world to detect stars, they would no longer apply “star” to stars if stars were changed in the way I have described. (It would of course beg the question to say that we would apply “star” to (the changed) stars if we were told that they were stars. The informational theorist cannot simply grant himself a reliable star detector and the knowledge that it is one. For in that case, his position would amount to the view that to mean STAR by “star” is to intend to apply “star” to whatever a known-to-be-reliable star detector tells us is a “star” – not an informational theory at all.)

If we make the minimal change in the world that is necessary for stars not to have starry appearances, slow comets will still have starry appearances. (Or, if this claim is false because of some nomological connection between the appearances of stars and slow

³⁹ See also Loewer and Rey (1991, p. xxxvi fn. 51). Admitting diachronic effects would undermine Fodor’s theory since, as Fodor points out, diachronic dependences cannot be expected to be the right way around (1987, p. 109).

⁴⁰ It might be objected that if either stars or slow comets were different in a way that affected their appearance, thus making them look different to us, then, given our linguistic intentions or dispositions, we would apply the term only to the stars. Thus, the objection continues, other things being equal, it is easier to eliminate the slow-comet law. The problem with this line of reasoning is that it evaluates the dependence diachronically rather than synchronically. It is true that if stars and comets had always looked different, whether because stars were different or because comets were different, we would have applied the term “star” only to stars. But given how we now apply “star,” if stars were made to look different, we would now – before we had time to adjust – apply the term to the things that looked the way stars used to look.

comets, there will be many other examples that can be used to make the same point). Since the thinker is unchanged, the thinker will still apply “star”s to objects with starry appearances. So the law that slow comets cause “star”s still obtains. Thus, *prima facie*, eliminating the star law does not eliminate the slow-comet law, so the latter law does not depend on the former. What about the other way around? Just as we can eliminate the star law by changing the appearance of stars, we can eliminate the slow-comet law by changing the appearance of slow comets (suppose for example, that they are surrounded by clouds of dust). And that change will leave the star law intact. At least with respect to this kind of case, it seems that the two laws’ situations are symmetrical; neither law depends on the other. In sum, it is false that, *ceteris paribus*, eliminating the star law eliminates the slow-comet law. Therefore, the slow-comet law is not asymmetrically dependent on the star law.

So far I have been arguing that, by changing the stars, the star law can be eliminated without eliminating the slow-comet law (and that, in a parallel way, the slow-comet law can be eliminated without eliminating the star law). It might be objected that there are other ways of eliminating the star law that would eliminate the slow-comet law. As we have seen, Fodor’s preferred way of eliminating the star law is to change my intentions with respect to the word “star”. Eliminating the star law *in this way* will in general eliminate the slow-comet law as well.

Now the mere fact that *some* ways of eliminating the star law will eliminate the slow-comet law is irrelevant. An asymmetry is required. But the objector’s basic idea is to question whether the counterfactual situations that I have considered are the relevant

ones. Thus, the objection raises the issue of the proper way of understanding asymmetric dependence.

As Fodor emphasizes, the fundamental issue is not about counterfactuals (or possible worlds interpretations thereof), but about the ontological dependence of one law on another. I will now argue that we can resolve the issue of whether the slow-comet law asymmetrically depends on the star law by directly examining the relations between the laws' mechanisms. We need not enter into the subsidiary, technical, and tricky, issues about which counterfactuals are relevant.⁴¹

The fundamental issue is whether the slow-comet law's existence depends on the star law's existence, but not *vice versa*. The existence of the laws in question, as we have seen, depends on mechanisms. Thus, in general, one law's existence depends on another law's existence if, and only if, the former law's existence depends on the existence of the mechanisms that sustain the latter law. It is this kind of thought that drives Fodor's intuitions about which laws depend on each other (as well as his method of arguing for the dependence of one law on another). For example, as we have seen, Fodor's very natural idea is that an XYZ law obtains because 1) the mechanism for the water law is in place, and 2) the thinker can't distinguish XYZ from water. Thus, the reason it seems plausible, at first blush, that the XYZ law asymmetrically depends on the water law is that the XYZ law's existence seems to depend on the existence of the mechanisms that sustain the water law. Notice also that questions about laws' mechanisms are more fundamental than questions about whether a law can be eliminated without eliminating

⁴¹ In other work, I show that both Fodor's original, looser counterfactual formulation and his later one support the conclusion that the star law is not primary. Since the issues about counterfactuals are secondary I omit the details here in order to save space.

another law, or about whether eliminating one law requires more change than eliminating another law.

Precisely what relations between the mechanisms of laws A and B are necessary in order for law B to depend asymmetrically on law A? If no part of law A's mechanism is part of law B's mechanism, law B does not depend on law A.⁴² If only part of law A's mechanism is part of law B's mechanism, law B's existence does not depend on law A's existence, but only on the existence of part of law A's mechanism. (That part of law A's mechanism could exist and law B could obtain without law A's obtaining.) Thus, in order for law A's obtaining to be part of what makes law B obtain, all of law A's mechanism has to be part of law B's mechanism. Moreover, in order for the dependence to be asymmetric, at least part of law B's mechanism cannot be part of law A's mechanism. Therefore, in order for law B to depend asymmetrically on law A, law A's mechanism must be a proper part of law B's mechanism.

I have showed, however, that the slow-comet law's mechanism and the star law's mechanism will have non-overlapping parts – *i.e.*, neither is a proper part of the other. Thus, the slow-comet law's obtaining does not depend on the star law's obtaining, so the slow-comet law does not asymmetrically depend on the star law. It is only by ignoring

⁴² It might be objected that even if the mechanisms are distinct, one mechanism could exist only because the other one does. If the objection is that the former mechanism would not have come into being had the first one not existed, then it covertly appeals to diachronic rather than synchronic dependence. See text accompanying note 39 above. Another version of the objection instead relies on synchronic counterfactual claims. But, as discussed in the text, the present discussion is premised on the idea that relations of ontological dependence are more explanatorily basic than such counterfactuals. Hence it is not legitimate to object, in a situation in which there is no mechanism that could mediate ontological dependence, that there is a counterfactual dependence. Nevertheless, it may legitimately be objected that even if law B's mechanism is not a proper part of law A's mechanism, there could be a more basic mechanism that ensures an ontological dependence of law B's mechanism on law A's mechanism. The answer to this objection is that, although such ontological dependences are possible in particular cases, it would require a massive cosmic coincidence to ensure that the asymmetric dependences fall the way Fodor's theory entails. The argument in the text is therefore a harmless oversimplification. I'm grateful to Ram Neta for pressing me on this point.

the part of the star law's mechanism that is out in the world that it can seem plausible that all of the star law's mechanism is a proper part of the slow-comet law's mechanism. That we can eliminate either law without eliminating the other is just a consequence of the more basic fact that neither law's mechanism is a proper part of the other. And it is because of this more basic fact that the slow-comet law does not depend on the star law.

Where did things go wrong? Fodor's argument assumes that if we make stars and slow comets distinguishable, we will eliminate the slow-comet law but not the star law. In order to eliminate the star law, he thinks, it is necessary to change our intentions about the meaning of "star," which will eliminate the slow-comet law as well. But we should have expected something was wrong at this point. Even if it is kosher for Fodor to appeal to our intentions to use "star" with a certain meaning, those intentions plainly cannot constitute the entire mechanism of the star law. There has to be a component of the mechanism that is in the world outside the thinker. It follows that the law can be eliminated not only by changing the meaning of "star" but by interfering with the part of the mechanism that is outside the thinker.⁴³

The argument so far has concerned wild laws, such as the slow-comet law, that cover systematically erroneous applications of symbols. It is worth noting that the argument is even stronger for a different kind of wild law. Fodor emphasizes that robustness includes not just erroneous occurrences of symbols but, for example, occurrences of symbols caused by other thoughts (1990, pp. 80-81; 1991a, p. 261). It

⁴³ How could Fodor think that my intentions ensure that the slow comet law asymmetrically depends on the star law? One possibility is that he simply overlooked the possibility of eliminating the star law by making a change in the world. He invariably characterizes the change as making As and Bs distinguishable, without analyzing what that would involve, thus making it possible to overlook the possibility and implications of changing the stars. Another is that he unwittingly slipped into a diachronic interpretation of the relevant changes. That is, if we changed the appearance of stars, our intentions would, over time, produce changed linguistic dispositions so that stars but not slow comets would cause "star"s. See note 40 above.

might be that *thoughts* about nuclear fusion or about Van Gogh (because of his famous painting of a night sky) reliably cause “star”s. Now if you assumed, as Fodor does, that eliminating the star law requires changing the meaning of “star,” then it would seem to follow that eliminating the star law would eliminate the law that, say, thoughts of Van Gogh cause “star”s. As we have seen, however, changing the stars can eliminate the star law. Changing the stars would not in general (synchronically) affect the various nomic relations between thoughts and the causing of “star” tokens, however. Even if stars were different – even if they had outer crusts that blocked all light – it would still be a law that *thoughts* of nuclear fusion would cause “star”s; the mechanisms that (now) mediate that law do not depend on how the stars are. The explanation is that the thoughts-of-fusion law is mediated by beliefs, for example, the belief that fusion occurs at the core of stars. And, synchronically, those beliefs do not depend on the nature of the stars. Similarly, Van Gogh thoughts would still cause “star”s because (synchronically) my association of Van Gogh with stars is not mediated by actual stars, but by beliefs, memories, and other mental states.

There are at least two ways in which the argument against the star law’s primacy is stronger for wild laws such as the thoughts-of-fusion law and the thoughts-of-Van-Gogh law than for wild laws covering erroneous tokenings of “star.” First, wild laws of the first kind cannot be eliminated except by making changes in the thinker’s mind. The reason is that the mechanism connecting thoughts with tokenings of “star” is entirely mental. A consequence is that there can be no question whether it requires less change in the world to eliminate the star law than to eliminate the relevant wild laws (since no amount of change in the external world will eliminate the relevant wild laws).

Second, and more importantly, there are psychological changes in the thinker of just the sort that Fodor focuses on, that will eliminate the star law without eliminating the thoughts-of-fusion and thoughts-of-Van-Gogh laws. Suppose we change the thinker's beliefs about how stars look or we change his visual acuity so that he no longer applies "star" to dots of light in the sky. The thinker continues to believe that nuclear fusion takes place in the core of stars, so the thoughts-of-fusion law still obtains. And assuming the thinker associates Van Gogh with stars directly (rather than via mental images of Van Gogh's painting of the night sky), the thoughts-of-Van-Gogh law still obtains. (Once again, it is irrelevant to synchronic dependence that, as a historical matter, the thinker's disposition to apply "star" to dots of light in the sky played a role in the thinker's acquiring the Van Gogh-star association.) Thus, the argument against asymmetric dependence is even stronger for these laws than for wild laws covering erroneous tokenings.⁴⁴

In this section, we have seen that it is false that, in general, laws covering the causing of "star"s depend – let alone asymmetrically depend – on the star law. And, as we have noted, there is nothing special about "star." In general, it is not plausible that the reference law – supposing there is such a law – for a mental or linguistic symbol will be the primary law. Given that a thinker means something by a symbol, she will develop dispositions, beliefs, mental associations, and so on that mediate between objects (including thoughts) and occurrences of the symbol. Once those mediating mechanisms are in place, they are not synchronically dependent on whatever mechanisms mediate the

⁴⁴ I noted above the possibility, for which I argue elsewhere, that slow comets and stars cause "star"s in virtue of being, very roughly, starry-appearing objects (*i.e.*, there is no slow-comet law or star law). That might lead one to think that, on Fodor's theory, "star" would mean starry-appearing object. For essentially the reasons discussed in the text, however, the thoughts-of-fusion law and the thoughts-of-Van-Gogh law do not asymmetrically depend on the starry-appearing-objects law (nor the converse).

connection between the symbol and objects in its extension. Thus, wild laws do not in general depend on the reference law.

6 Conclusion

As discussed, Fodor's strategy is not to argue that his theory accounts for any actual cases of mental content, but rather to claim that *if* a symbol satisfied his conditions, it would have the meaning specified by his theory. He begins by trying to make that claim intuitively plausible. He then defends the theory largely by showing that 1) it solves the disjunction problem and 2) fighting off putative counterexamples. Given this strategy, the intuitive motivation for Fodor's theory is particularly important. I have already argued that Fodor's approach does not solve the disjunction problem so much as stipulate it away.⁴⁵ Hence without an intuitive motivation, he has given us very little positive reason for believing the theory.⁴⁶

I will therefore conclude by pointing out a problem with Fodor's attempt to bring out the intuitive plausibility of his theory. He appeals to the notion that applications of symbols to objects that fall under them (as opposed to mistaken applications and non-application occurrences of symbols) are the most basic uses of symbols, the uses on which all others depend (pp. 96-99). If we find this notion plausible, he suggests, we will find it plausible that at least roughly the right kind of asymmetric dependences will tend to be established. What makes the notion plausible in the first place is the thought that, for example, the explanation of why we apply "horse" to cows on dark nights is that "horse" applies to horses (and cows on dark nights look like horses). In contrast, the

⁴⁵ See section 4 above.

⁴⁶ Note 31 above describes a few other minor ways in which Fodor supports the theory.

explanation of our application of “horse” to horses does not involve cows on dark nights. Thus, there is at least an initial temptation to think that the distinction between primary laws and other laws might mirror the relevant structure of reasons for uses of symbols. In particular, the uses of symbols that are content-determining might be part of the explanation of the uses of symbols that are not content-determining, but not *vice versa*. Accordingly, there is some plausibility to the idea that reference laws will be primary.

There is a serious worry about this intuitive motivation for the asymmetric-dependence theory. In assessing the suggestion that non-application uses of symbols depend on (correct) applications, we must be careful not to confuse it with the proposition that non-application uses of symbols depend on the symbol’s having the reference that it does. It is uncontroversial that all (relevant) uses of a symbol depend on its having the reference that it does. (In fact, since we are granting for purposes of argument the informationalist’s assumption that meaning is reference, the proposition that (meaningful) uses of symbols depend on the symbol’s having the reference that it does is just the platitude that (meaningful) uses of symbols depend on the symbols’ meanings.) This is not equivalent to the claim that all other uses of symbols depend on applications; in fact, an application of a symbol to an object depends on the symbol’s reference as much as any other use: we would not apply the symbol “slab” to a slab if “slab” did not refer to slabs.

In sum, it is uncontroversial that we apply, say, “star” to thoughts of Van Gogh only because “star” means *star* or refers to stars. It is highly controversial – and, in effect, the basic claim of Fodor’s theory – that we apply “star” to thoughts of Van Gogh only because “star” has a nomic connection with stars. It is crucial not to slip from the

plausibility of the uncontroversial claim to the plausibility of Fodor's theory. Fodor in fact seems to equivocate in just this way (1987, p. 108; 1990, p. 97).

References

- Antony, L. & Levine, J. "The Nomic and the Robust," in Loewer and Rey 1991, pp. 1-16.
- Boghossian, P. 1991: "Naturalizing Content," in Loewer and Rey 1991, pp. 65-86.
- Dretske, F. 1981: *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.
- 1988: *Explaining Behavior*. Cambridge, MA: MIT Press.
- Fodor, J. 1975: *The Language of Thought*. Cambridge, MA: Harvard University Press.
- 1986: "Why Paramecia Don't Have Mental Representations," in *Midwest Studies in Philosophy*, 10, pp. 3-24.
- 1987: *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- 1988: Reply to Boghossian (unpublished), APA Eastern Division Meeting, December 1988, quoted in Boghossian, 1991, pp. 81-82
- 1990: *A Theory of Content and Other Essays*. Cambridge, MA.: MIT Press.
- 1991a: "Replies," in Loewer and Rey 1991, pp. 255-319.
- 1991b: "You Can Fool Some of the People All of the Time, Everything Else Being Equal; Hedged Laws and Psychological Explanation." *Mind*, 100, pp.19-34.
- 1994: *The Elm and the Expert: Mentalese and Its Semantics*. Cambridge, MA: MIT Press.
- 1998: *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Kripke, S. 1982: *Wittgenstein on Rules and Private Language*. Oxford: Blackwell.
- Loewer, B. and Rey, G., eds. 1991: *Meaning in Mind: Fodor and His Critics*. Cambridge, MA: Blackwell.
- Pietroski, P. and Rey, G. 1995: "When Other Things Aren't Equal: Saving *Ceteris Paribus* Laws from Vacuity," *British Journal for the Philosophy of Science*, 46, pp. 81-110.