Systematicity

Robert Cummins


Stable URL:
http://links.jstor.org/sici?sici=0022-362X%28199612%2993%3A12%3C591%3AS%3E2.0.CO%3B2-C

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at
http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you
have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and
you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or
printed page of such transmission.

The Journal of Philosophy is published by Journal of Philosophy, Inc.. Please contact the publisher for further
permissions regarding the use of this work. Publisher contact information may be obtained at
http://www.jstor.org/journals/jphil.html.

The Journal of Philosophy
©1996 Journal of Philosophy, Inc.

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office.
For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
SYSTEMATICITY

Jerry Fodor and various colleagues\(^1\) have argued that the human scheme of mental representation exhibits a certain sort of constituent structure—a \textit{classical} structure—on the grounds that the systematicity of thought and language, alleged to be evident and pervasive, is best explained on the hypothesis that mental representation is classical. It turns out to be more difficult than one would expect to specify either what it is for thought to be systematic or for a representational scheme to be "classical" in the required sense. The definitions in the literature are quite hopeless, and when the deficiencies are corrected, the thesis that thought is systematic is seen to be problematic at best and the argument from the systematicity of thought to the alleged classicalness of representation collapses. The systematicity of language, on the other hand, is relatively unproblematic, but it turns out that sensitivity to the systematicity of language can be explained without appeal to classical representation. Moreover, there are other systematicities—for example, in vision—that must receive some explanation or other. If the argument from linguistic systematicity to classical representation is accepted, then we should also accept an exactly parallel argument from visual systematicity to nonclassical representation. The argument from human sensitivity to systematicity to classical representation is therefore self-defeating if it is meant to establish that \textit{the} scheme of mental representation in humans is classical.

I. CLASSICAL REPRESENTATION

Let us begin with the idea of a classical representational scheme. Here is the preferred definition: a representational scheme is *classical* in the intended sense if tokening complex representation requires tokening its constituents. On the obvious reading, this is surely a tautology, since what makes a representation complex is just that it has constituents. A representation that does not have constituents will not count as complex. We can rescue this notion from triviality only by supposing that a token of a complex representation can "have" a constituent that is not tokened. Tim van Gelder\(^3\) suggests that we might broaden the notion of compositionality so that \(x\) might count as a constituent of \(y\) even if tokening \(y\) does not require tokening \(x\), provided that you can reliably get from \(x\) to \(y\) and back again. He calls this *functional compositionality*. But this appears to leave the notion of complexity in some obscurity. As P. Smolensky, G. LeGendre, and Y. Miyata\(^4\) show, we can represent arbitrarily complex structures as vectors of activations, but the complexity of the resulting vectors does not correspond to the complexity of the structure represented. Indeed, all the vectors in such a scheme are of equal complexity (since they are of the same dimension). The only sense in which one representation in this scheme is more complex than another is that some decode into more complex structures than others. This relativizes the complexity of a representation to a decoding. Since there may be more than one decoding possible, and these need not be semantically equivalent, it also makes complexity an implicitly semantic notion rather than a purely formal notion. If we reject the dubious notion of functional compositionality, the criterion of classicalness proposed by Fodor and Brian McLaughlin will make every scheme classical. Evidently, we need a different approach.

Another way of getting at the idea of classical representation is that a scheme is classical if representing a complex content requires a correspondingly complex representation, with each constituent of the representation corresponding to each constituent of the content represented. This is certainly operative in the discussion of system-

\(^2\) Fodor and McLaughlin; Fodor and Pylyshyn.


Systematicity in Fodor and Z. Pylyshyn’s classic discussion, for example. Following C. Swoyer, I shall call representational schemes that satisfy this condition structural schemes. Fodor and Pylyshyn’s discussion of systematicity suggests that classical representation is structural representation. But the idea that classical representation is structural representation faces two serious difficulties. First, this formulation requires that propositions be complex—complex, moreover, in just the way that the representations are complex. And second, it pre-

5 Fodor and Pylyshyn suggest that a Connectionist might get simplification by having a node representing \( p \& q \) activate nodes representing \( p \) and representing \( q \).

Since each node has (let us suppose) no semantically relevant internal complexity, it follows that the representation of \( p \& q \) is no more complex than the representation of \( p \), and certainly does not contain a representation of \( p \) as a constituent. The question arises, however, as to why the whole three-node net does not count as a classical representation of \( p \& q \). (Notice that it misses the point of this proposal to argue that we could cut the connections, or change the weights, allowing activation of the top node, hence tokening \( p \& q \), without tokening either \( p \) or \( q \).

The obvious response to this sort of proposal is to argue that it is unfair to build, as it were, the entailments into the representation. This response is unavailable, however, to anyone who, like Fodor himself, is an advocate of functional-role semantics for the logical connectives. To say that logical connectives are identified by their functional role is just to say that logical form is determined by what inferences are made. Functional-role semantics says that what makes something a conjunction is precisely connections to other formulas. So what makes it appropriate to label the top node \( p \& q \) is just that it is connected in the proper way to nodes labeled \( p \) and \( q \) (among other things). This is not quite the same as saying that the whole net is the representation of \( p \& q \), but it is as close as makes no difference. If you accept functional-role semantics, then nothing that does not work pretty much the way the net does will count as a representation of \( p \& q \), and that is as close as makes no difference to saying that the whole net is a classical representation of \( p \& q \) by the criterion offered in Fodor and Pylyshyn and in Fodor and McLaughlin: whenever you token a representation of \( p \& q \), you are bound to token one of \( p \), since, if you do not, you do not have the causal relations in place that make something a conjunction in the first place. Since Fodor himself accepts functional-role semantics as an account of the content of the logical connectives, it is hard to see how he could object to this construal. Those who think the meanings of logical connectives are fixed functionally are therefore committed to accepting van Gelder’s idea that functional compositionality is as good as the “real thing.” I would prefer to take this as a reductio against a functional-role approach to the connectives.

cludes abbreviation \((r = df p \& q)\), which is surely not intended. The first consequence—that propositions themselves have a constituent structure that mirrors the constituent structure of some favored representational scheme—is controversial, and I propose to set it aside for the moment. But the second consequence defeats the proposal outright. In the sort of scheme Fodor et alia have in mind, differences in complexity need not correlate with differences in content. We want to be able to hold with Fodor that ‘Jocasta is eligible’ and ‘Œdipus’ mother is eligible’ have the same content in spite of having different constituents. It will not do, then, to suppose that a representation is classical just in case representations of complex contents require correspondingly complex representations.\(^7\)

What Fodor has in mind, of course, is a language, a language with the sort of syntax and semantics that admits of a Tarskian truth theory. So, perhaps with some de jure but no de facto loss of generality, I shall call a scheme classical if (1) it has a finite number of semantically primitive expressions individuated syntactically; (2) every expression is a concatenation of the primitives; and (3) the content of any complex expression is a function of the contents of the primitives and the syntax of the whole expression.

II. SYSTEMATICITY IN THOUGHT

Recent discussion in the literature has centered around language processing.\(^8\) A system is said to exhibit systematicity if, whenever it can process a sentence \(s\), it can process systematic variants of \(s\), where systematic variation is understood in terms of permuting constituents or (more strongly) substituting constituents of the same grammatical category. Systematicity in this sense has been amply demonstrated in connectionist systems that do not use classical representational schemes.\(^9\) L. Niklasson and van Gelder are surely right to claim that the issue is now the difficult empirical one of determining (i) what systematicity humans actually exhibit in their language processing, and (ii) what sort of architecture best models it.

---

\(^7\) See chapter six of Fodor, A Theory of Content and Other Essays (Cambridge: MIT, 1990). Notice that, for a language of thought theorist, chunking is abbreviation. Since chunking is essential to every orthodox computational architecture that has any pretensions to psychological relevance, abbreviation is an ineliminable feature of “classical” schemes.


\(^9\) See Niklasson and van Gelder; Smolensky, LeGendre, and Miyata.
I shall return shortly to systematicity in language. First, however, I want to consider the alleged systematicity of thought. The usual formulation (for example, Fodor and Pylyshyn’s) goes like this:

(T0) Anyone who can think a thought of the form \textit{Rab} can think a thought of the form \textit{Rba}.

This is hopeless as it stands because it presupposes that thoughts have forms corresponding to the forms of classical representations, and this is precisely the point at issue.\textsuperscript{10} To be non-question-begging, we need a formulation like this:

(T1) Anyone who can think a thought with a content of the form \textit{Rab} can think a thought with a content of the form \textit{Rba}.

This evidently faces a difficulty encountered above and set aside, namely, that contents may not have the relevant forms or any form at all. We can sidestep the issue by reformulating the condition:

(T2) Anyone who can think a thought with the content \textit{c} can think a thought with the content \textit{c*}, where \textit{c*} is any systematic variant of \textit{c}.

Intuitively, the idea is that \textit{c*} is a systematic variant of \textit{c} if \textit{c*} is a content you can express by permuting the constituents of your expression of \textit{c}.\textsuperscript{11} This allows (on the face of it, anyway) for the possibility that contents themselves are unstructured or not structured in a way that happens to mirror the structure of your representations. For example, it permits, but does not require, that anyone who can think the set of possible worlds picked out by ‘John loves Mary’ can also think the set of possible worlds picked out by ‘Mary loves John’.

This approach evidently relativizes the systematicity of thought to the choice of some representational scheme that allows for the permuting of constituents. That consequence is disturbing. The systematicity of thought ought to depend only on the structure of the mind, not on the structure of the representational scheme we, as theorists, choose to use in representing the contents of thoughts. We could defend a particular choice of scheme, of course, if we knew it reflected the structure of the propositions thought, but the motiva-

\textsuperscript{10} Of course, sometimes one thinks thoughts in one’s natural language. In this case, one could argue that thoughts are represented in a classical scheme since they are represented in, say, English, which is classical. In this case, the systematicity of thought simply reduces to the systematicity of language, which I shall take up below.

\textsuperscript{11} Depending on how you do your grammar, allowable grammatical permutations may not yield an expression with a content. ‘The hill went up Jack’ is, perhaps, an allowable permutation of ‘Jack went up the hill’, but may not express any proposition at all.
tion for (T2) is precisely to avoid having to make any such dubious commitment.¹² Still, I think it is pretty clear that (T2) is what underlies intuitions about the systematicity of thought, such as they are, for those intuitions clearly depend on one’s natural language. The intuitions in question are claims like the following:

(1) Anyone who can think that John loves Mary can think that Mary loves John.

The systematicity of thought that this is supposed to illustrate is clearly derived from the systematicity of the language—in this case, from the fact that ‘John loves Mary’ is a permutation of ‘Mary loves John’. The intuitive force of (1) would obviously disappear or change if we could substitute either unstructured or differently structured classical representations for the propositions, or if we could substitute nonclassical representations for the propositions. Consider:

(2) Anyone who can think Mary’s favorite proposition can think that Mary loves John.

On the obvious assumption, the italicized phrase in (2) refers to the same proposition as the italicized phrase in (1). But (2) elicits no systematicity intuitions precisely because the italicized phrase in (2) is not a permutation of ‘Mary loves John’.

One might object at this point that (2) does not tell us what Mary’s favorite proposition is, and hence that we are in no position to determine whether her favorite proposition is a systematic variant of the proposition that Mary loves John. This is fair enough, perhaps, but it underlines the fact that identifying the propositions in question requires the mediation of an appropriate representation. It is the structure of the mediating representation which determines whether or not we see systematicity in the thoughts. (2) seems like cheating because it does not give us canonical representations of the propositions in question. But that is just my point: (1) gets all of its

¹² One might think that propositions have classical forms on the grounds that they are expressible by classical representations. This appears to be Fodor’s line in Theory of Content. But one might equally think that propositions have pictorial forms on the grounds that they are expressible by pictorial representations. (The proposition represented by a picture is the set of possible worlds it accurately depicts. For example, the picture on your driver’s license picks out the set of possible worlds in which someone looks like you do in that picture.) Moreover, it is a notable feature of classical schemes that representations of radically different forms can express the same proposition. Indeed, since logical equivalence is not decidable, it is not even decidable in a classical scheme whether two forms express the same proposition.
appeal from the fact that it incorporates linguistic expressions for the propositions that are interderivable by permutation. You get the same dialectic if you put in Gödel numbers or activation vectors: once you lose the relevant constituent structure in the expression of the propositions, you lose the intuition for systematicity. The apparent obviousness of the systematicity of thought looks to be an illusion created by reading the structure of contents off the structure of their representations.

A comparable point can be made by noticing that thought may appear systematic relative to one scheme but not systematic relative to another. Consider:

(3) Anyone who can think that a face is smiling can think that a face is frowning.

This looks pretty implausible. But imagine a palette scheme for constructing cartoon faces:

Under the influence of this scheme, (4) looks pretty plausible:

(4) Anyone who can imagine a smiling face can imagine a frowning face.

Absent some representation-independent access to the structure of propositions, which propositions seem to be systematic variants of each other will depend on one’s preferred scheme for representing propositions. If you linguistically represent the contents to be thought, then you will want mental representation to be linguistic, since then the systematicities in thought that are visible from your perspective will be exactly the ones your mental scheme can explain. You can make things look hard for (some) connectionists by (a) covertly relativizing systematicity to a natural language, and (b) re-
minding them that they favor a nonclassical scheme of mental representation. If you can get them to accept this "challenge," they will then labor away trying to show that a user of a nonclassical scheme might still exhibit the systematicities visible from a classical perspective. Getting connectionists to accept this challenge is nice work if you can get it, and apparently you can. There may still be job openings in this area if you want employment in the confidence business.

The preceding argument depends on the claim that something like (T2) is the only way to get at the systematicity of thought. Actually, there is another way. One supposes that thoughts are relations to mental representations. According to the standard story, believing that \( p \) is harboring a representation that \( p \) in the belief box; desiring that \( p \) is harboring a representation that \( p \) in the desire box; and so on. Thoughts then inherit the forms of their associated representations, and the systematicity of thought is just a reflection of the systematicity of the scheme of mental representation. If you think that mental representation is classical, then you are entitled to:

Anyone who can think a thought of the form \( Rab \) can think a thought of the form \( Rba \).

But equally, if you think mental representations are activation vectors, then you are entitled to:

Anyone who can think a thought of the form \( <\ldots a\ldots b\ldots> \) can think a thought of the form \( <\ldots b\ldots a\ldots> \).

The whole point of the appeal to systematicity was to argue from systematicity to a conclusion about the form of mental representation. Everyone gets a free explanation of whatever systematicities are exhibited by the scheme of mental representation they hypothesize. You only get independent leverage on what sort of scheme the mind uses if you have a trick for independently identifying systematicities in thought, that is, a trick that does not depend on any assumptions about the scheme of representation the mind uses.

III. SENTENCE PROCESSING

What the systematicity argument requires is a domain \( D \) of things such that (a) some members of \( D \) are systematic variants of others, and (b) it is an empirical fact that anyone who can represent a member of \( D \) can represent its systematic variants. The problem we are having is that we have no non-question-begging way of saying which propositions are related by systematic variation. This suggests that the problem is with propositions, not with systematicity. Surely, there are other domains that loom large in human cognition and whose
members exhibit systematic relations. This is one reason why the debate has tended to focus on the issue of language processing. Sentences are systematically related in various ways, so it might seem that the issue could be joined meaningfully in that venue, where the claim at stake is this:

(L) Anyone who can understand a sentence of the form \( F \) can understand a sentence of the form \( F^* \) (where \( F \) and \( F^* \) are systematic variants\(^{19}\)).

No one, of course, is in a position to say definitively what sort of representational scheme might be required for understanding sentences. But we are in a position to ask whether various sentence processing tasks that are plausibly supposed to be necessary conditions of systematic understanding are possible for systems employing nonclassical representational schemes. So we can wonder about (L1), for example:

(L1) Anyone who can parse a sentence of the form \( F \) can parse a sentence of the form \( F^* \).

As noted above, it is now pretty clear that (L1) holds of systems that do not employ a classical representational scheme, or any scheme that structurally represents sentences or their grammatical forms.

It is interesting to reflect on what makes this possible. The answer, in a word, is complete encoding. By an encoding of a domain \( D \) (which might itself be a representational scheme), I mean a recursive mapping of the members of \( D \) onto the representations in a scheme \( S \) whose members do not preserve the internal complexity (if any) of the members of \( D \). Gödel numbering of sentences is an encoding in this sense, since it does not preserve constituent structure, while Morse code is not an encoding since it is simply an alternative alphabet, preserving the same letters in the same order, which is sufficient to preserve the syntactic structure of a (written) sentence. An encoding is complete if it provides a representation of every member of the target domain.

Unlike an encoding, a structural representation (SR) is an isomorph of the thing it represents, so a scheme of structural representations does preserve the structure of the items in the represented domain. The issue between connectionists and classicists concerning the explanation of (L1) is whether structural representation is required—that is, whether the representations of \( F \) and \( F^* \) required by the

\(^{19}\) See Niklasson and van Gelder for various levels of systematic variation that might be plugged in here to give claims of different strength.
parser must themselves be systematic variants—or whether an encoding of the target sentences is sufficient. Classical representation gets into the debate only because it happens that classical schemes can (but need not) provide structural representations of sentences.\footnote{Recall that a classical scheme can represent sentences without structurally representing them. There is abbreviation \( s = \text{'John loves Mary'} \) and there is logical equivalence ('John loves Mary' is equivalent to 'the sentence \( s \) such that \( s = \text{'John loves Mary'} \) or \( 1+1=2 \)).} It has seemed to classicists that structural representation of sentences is required to account for \( (L) \) and its corollaries (for example, \( (L1) \)).

The work of Smolensky, LeGendre, and Miyata demonstrates that a complete encoding is sufficient, for they prove an equivalence between a parser written in TPPL, a LISP-like language that uses classical representations, and a network using fully distributed representations (\textit{op. cit.}). The network devised by Niklasson and van Gelder appears to demonstrate the same point (\textit{op. cit.}). Structural representation of sentences is not required for parsing tasks, and since the only reason for preferring classical schemes in this case is that they can provide structural representations of sentences, the argument to classical representation from language processing appears to collapse.

\textbf{IV. SYSTEMATICITY}

Friends of the systematicity argument need not throw in the towel just yet, however. They can legitimately complain (though they have not) that \( (L) \) and \( (L1) \) are incautiously formulated. The explananda in this debate have the form:

\((S)\) Anyone who can represent \( c \) can represent \( c^* \).

where \( c \) and \( c^* \) are systematic variants in some domain \( D \).\footnote{(L1) does not actually have this form, but it entails something of the right sort, since it says, in effect, that anyone who can represent a parse of \( F \) can represent a parse of \( F^* \), and these will be systematic variants of each other on the assumption that \( F \) and \( F^* \) are systematic variants.} The problem is that \((S)\) will be satisfied by any system employing a representational scheme that is complete for \( D \). Being able to represent systematic variants is a trivial consequence of being able to represent everything. So, for example, if you can represent every sentence of \( L \), it will follow that if you can represent \( s \in L \) you can represent its systematic variants. This is evidently not what is wanted for the systematicity argument, since the whole idea is that a system handles
systematic variants in some way that is special. What is needed is the idea that, in a system sensitive to a systematicity in $D$, generating a representation of $c^*$ from $c$ is easy or principled in a way that moving from one representation to another generally is not. This is what makes structural representation a natural candidate for explaining sensitivity to systematicity: to get from a representation of $c$ to a representation of $c^*$, all you have to do is to permute constituents of the representation of $c$, you do not have to start from scratch.\(^{16}\)

A system is sensitive to a systematicity in $D$, then, if there is something about its processing of representations of systematic variants in $D$ which is special. There must be, in short, some effect $E$ that a system exhibits when it processes representations of systematic variants and does not exhibit otherwise. A system might, for example, get from a representation of $c$ to a representation of $c^*$ faster than to a representation of $d$. Or perhaps the time required to move from a representation of $c$ to a representation of $c^*$ is constant or independent of the complexity of $c^*$. Moving between representations of systematic variants might be less error prone or prone to some characteristic sort of error. All of these would count as exhibiting sensitivity to systematicity in $D$, and they are the sorts of effects one would expect if the system employs a structural representation, since permuting elements of $c$ should be simpler than constructing a new representation.

With this rudimentary understanding of systematicity, we can see that the success of connectionist parsers does not really settle the question. On the other hand, to my knowledge, no definite question has been posed. That is, no one has specified a systematicity effect which human parsers exhibit which does not simply reduce to the ability to parse any grammatical sentence that does not overload short term memory. Perhaps there are such effects—they may even be well known—but they have not made their way into the systematicity debate as yet.

The relation between systematicity and classical representation is only a bone of contention between classicists and connectionists because connectionists (some of them) believe in nonclassical representation. The deeper issue is whether any nonclassical system—any system using nonclassical representation—can achieve sensitivity to systematicity. As we have just seen, we have a substantive systematicity

---

\(^{16}\) I am reading ‘permute’ liberally here to include turning over the “smile” in a cartoon face, as well as, say, exchanging it with an eye to give a surprised person winking.
effect if a system exhibits some characteristic effect when moving from a representation of \( c \) to a representation of \( c^* \), a systematic variant of \( c \). Within a broadly computationalist framework, there are two ways in which effects can arise: (I) they are computed; (II) they are incidental, that is, side-effects. An effect is computed just in case exhibiting the effect is a matter of computing the characteristic argument-to-value pairs; it is incidental just in case exhibiting it is, as it were, a by-product of computation: not a matter of which values are computed from which arguments, but a matter of some effect of doing the computation, such as how long it takes, how much the system heats up, or how much sugar is metabolized. An example will make the distinction clear and, at the same time, introduce a methodological problem that arises in drawing it.

Consider two multipliers, M1 and M2. M1 uses the standard partial products algorithm we all learned in school. M2 uses successive addition. Both systems exhibit the multiplication effect: given two numerals, they return a numeral representing the product of the numbers represented by the inputs. M2 exhibits the linearity effect: computation is, roughly, a linear function of the size of the multiplier. It takes twice as long to compute \( 24 \times N \) as it does to compute \( 12 \times N \). M1 does not exhibit the linearity effect. Its complexity profile is, roughly, a step function of the number of digits in the multiplier. The linearity effect is incidental; the multiplication effect is computed.

Of course, the linearity effect might be computed. We could design a system M3 that not only computes products, but computes reaction times as well, timing its outputs to mimic a successive addition machine. M3 might be quite difficult to distinguish from M1 on behavioral grounds. It need not be impossible. The timing function might be disabled somehow without disabling the multiplier. More subtly, computation of the relevant output times might itself be nonlinear, in which case M3 will not be able to fool us on very large inputs (assuming it can process them at all). Or it might be that the linearity effect in M3 is cognitively penetrable.\(^{17}\)

This last possibility is important, because it reminds us that incidental effects, being architecture dependent, cannot be cognitively penetrated. Since a difference in representational scheme is precisely a difference in architecture in the relevant sense, it follows that two systems using different representational schemes cannot exhibit all the same incidental effects. It follows further that classical systems

and distributed connectionist systems cannot exhibit all the same incidental effects. It seems clear that proponents of the systematicity argument are thinking that systematicity effects are incidental.\textsuperscript{18} They are thinking, recall, that you can easily get from a representation of the sentence ‘John loves Mary’ to a representation of the sentence ‘Mary loves John’ because (I) you represent these structurally—classical schemes providing structural representations of sentences—and (II) given a structural representation of one sentence you can get a structural representation of the other simply by permuting constituents, a much simpler operation, one might suppose, than constructing a new representation, as a system that encodes sentences must do. Indeed, a system that encodes sentences will have to compute the relevant effect. As we shall see shortly, this may make an encoding explanation of the effect ad hoc in a certain sense, but it need not invalidate it. When it comes to systematicity,

\textsuperscript{18} Suppose the effects in question are \textit{computed} by a system that uses a scheme of structural representation. Then they can be computed by a system that uses an encoding scheme. Let $D$ be some domain exhibiting systematicity, and let $f$ be a computational procedure that manipulates structural representations of members of $D$ in such a way as to exhibit sensitivity to the relevant systematicity. Now let $e$ be an encoding of the structural representations manipulated by $f$, and define $f^*$ as follows:

\[ f^*(x) = e(f(e^{-1}(x)) \]

The function $f^*$ is evidently computable, since $f$, $e$, and $e$'s inverse are computable. Any relation $\tilde{f}$ imposes on the things its structural representations represent will be imposed on those things by $f^*$ as well. If you have a classical explanation of computed systematicity effects, you can have a nonclassical explanation, too.

I have heard the following objection: "If the only way to compute $f^*$ is to decode into structural representations, then you do need structural representations to do the job." But there is no reason to assume that the only way to compute $f^*$ involves computing $e$. Indeed, there are bound to be other ways of doing the job. Without loss of generality, we can suppose $e$ to be a numerical encoding. Since $f^*$ is a recursive function that takes numerical arguments and values, it follows that it can be defined in terms of the standard arithmetical functions plus minimalization and composition. There is therefore an arithmetical algorithm $A_r$ for computing $f^*$ which traffics only in numerals, never having anything to do with their decodings. We can think of $A_r$ as manipulating numerical encodings of $D$ directly, since an encoding of a scheme of structural representations of $D$ is easily exchanged for an encoding of the relevant structures of $D$ themselves. Let $h$ be a function taking elements $d$ of $D$ onto their structural representations. Then $h^*$ encodes $D$ directly, where $h^*(d) = e(h(d))$. Evidently, $f^*(h^*(d)) = h^*(d')$ if and only if $f(h(d)) = f(h(d'))$. So we can think of $A_r$ as operating on a direct encoding of $D$.

Connectionists, of course, must do more than merely provide an encoding of a classical grammar if they are to explain (L1). They must provide an encoding into activation vectors, and they must demonstrate that a network can compute the appropriate star function. (This is exactly the approach of Smolensky, LeGendre, and Miyata.) These are, of course, nontrivial tasks. But the possibility of $A_r$ is enough to demonstrate that structural representation of a systematic domain is not required to explain sensitivity to that systematicity.
some encoding explanations are inevitable, as I shall argue in the next section.

V. NONLINGUISTIC SYSTEMATICITY

If we focus exclusively on systematicities in language, classical representation occupies a kind of favored position, because it affords structural representations of sentences. Structural representation is not the only way to go, as we have just seen, but it is a natural way to go. If you want to explain sensitivity to structure in \( D \), it is natural to hypothesize a representational scheme whose members preserve (that is, actually have) the relevant structure.

Sensitivity to nonlinguistic systematicities appears to abound in human psychology, however. Consider, for example, the perception of objects in space: anyone who can see (imagine) a scene involving objects \( o_1 \) and \( o_2 \) can see (imagine) a scene in which their locations are switched:

\[ (SP) \text{ Anyone who can see (imagine) a scene involving objects } o_1 \text{ and } o_2 \text{ can see (imagine) a scene in which their locations are switched.} \]

Again, any system employing a complete scheme will satisfy (SP), but it is surely the case that spatial representation underlies a number of substantive systematicities as well. One might well take the extensive literature on imagery to provide examples. Indeed, the imagery literature appears to put the shoe on the other foot, the friends of imagery arguing that imagery effects are best explained as incidental effects of employing a structural representation of space, while opponents argue that they are computed by a system employing an encoding in a classical scheme.

Once we see that there is sensitivity to systematicity in nonlinguistic domains, an argument parallel to that given for the classical representation of language would be available for the nonclassical representation of these other domains. The perception of objects in space appears to be a case in point. A map in three dimensions that has object representations as constituents would provide a structural representation of the domain, and hence allow for a "simple" explanation of the systematicity effects suggested by (SP). A classical scheme of representation will have its work cut out for it, however, since three-dimensional scenes are not isomorphic to Tarskian structures, that is, to the kind of structure a representation must have to be subject to Tarskian combinatorics—the kind of structure we are calling classical. No scheme can structurally represent both sentence structure and spatial structure, since these are not isomorphic to each other. Friends of the Fodorian argument that moves from sensi-
tivity to linguistic systematicity to classical representation ought to be friendly to an analogous argument that moves from sensitivity to spatial systematicity to nonclassical representation, for the Fodorian argument is really just an argument from sensitivity to systematicity in \( D \) to structural representation of \( D \). If you like this kind of argument, you had better be prepared for as many schemes of mental representation as there are structurally distinct domains with systematicities to which we are sensitive. People can transpose music and color schemes as well as spatial scenes and sentences. One flavor of representation will not fit all.

Of course, there is no exclusive choice that has to be made here. Maybe there are three schemes of structural representation (sight, touch, smell-taste, hearing) and a lot of encoding to boot. Maybe there are two structural schemes, or twenty-five. My point here is just that you cannot regard the need to explain a substantive version of (L) as decisive for classical representation unless you also regard the need to explain a substantive version of (SP) as decisive for nonclassical representation. Seeing that much is enough to show you that you cannot get from human sensitivity to systematicity to the conclusion that human mental representation is exclusively classical.

**VI. AN ARGUMENT FOR STRUCTURAL REPRESENTATION**

There is another Fodorian argument that goes like this.

You can get sensitivity to systematicities in a domain without structurally representing its members. But an explanation in terms of structural representation is better, other things equal, because postulating structural representations constrains which systematicities you are going to get, namely, the ones mirrored in your representational scheme. Roughly speaking, given structural representations, you are going to get sensitivity to the mirrored systematicities whether you like it or not. The encoding hypothesis, on the other hand, leaves it open which systematicities you are going to get. With encoding, it all depends on the processing: you program in just the systematicities you want and leave the others out. So the encoding approach is ad hoc, while the structural representation approach is principled.\(^{19}\)

This argument was intended to favor classical schemes over (distributed) connectionist schemes. But, given the fact that we have nonisomorphic domains to deal with, not all systematicity can be explained by appeal to the structural properties of any single scheme. If you favor an explanation of sensitivity to systematicity in terms of structural

\(^{19}\) This is generalized version of an argument found in Fodor and McLaughlin, systematically translated into my terminology, of course.
representation, then you are going to have to postulate at least one nonclassical scheme. If you want to stick with a classical scheme, then some systematicity effects will have to be explained via a nonstructural encoding of the domain. As an argument for classical representation and only classical representation, therefore, the proposed tactic is doomed by the fact that we are sensitive to quite distinct systematicities. Still, it is worth investigating the idea that explanations of sensitivity to systematicity are ad hoc unless they appeal to structural representation. There are two reasons why we need to get straight about this. First, there is the fact that connectionist explanations of sensitivity to systematicities in language do not appeal to structural representations of language. And second, there is the diversity of systematicities to which we are sensitive: if we do not postulate a different representational scheme for each case, then some systematicity is going to need an encoding explanation, and those are going to be ad hoc by the standard currently on the table.

In what sense, then, are encoding explanations supposed to be ad hoc? To get at what is behind this idea, we need to return to the fact that structural representation schemes are themselves systematic. It is a defining characteristic of structural representation that systematicity in the represented domain is mirrored by a corresponding systematicity in the representational scheme. Whenever two elements of the domain are systematic variants of one another, their representations are systematic variants of one another as well. It follows from this that whenever the scheme can represent c, it can represent its systematic variants, and this seems to be just what is wanted for explaining systematicity effects.

As we have seen, however, the point is not just that whenever you have a representation of c you are guaranteed to have a representation of its systematic variants, for any complete scheme of representation can make this claim. The point is rather that generating a representation of c* from a representation of c is easy and principled in a scheme of structural representation but not in an encoding scheme. The idea is that you can state a formal rule that captures the systematic variation, and that rule can, as it were, be made a principle of cognition. For example, you can easily imagine a scene with

---

59 There might be some worry about how the scheme is known to be complete. In a scheme of structural representation, we are guaranteed representations of all the relevant variants. What is the corresponding guarantee in an encoding scheme? Well, think of Gödel numbers: you have a representation for every sentence, so you have a representation for every systematic variant of s for variable s. This example shows that there are ways of guaranteeing the completeness of an encoding scheme.
the positions of two objects swapped because it is a simple matter to

generate the required representation: you just swap the object repre-
sentations in the scene representation. You cannot do this in an en-
coding scheme, because encodings are arbitrary: there is no formal

relation an encoding of \( c \) bears to an encoding of \( c^* \) if and only if \( c \)

and \( c^* \) are systematic variants.

\textit{Objection}: there is the following relation. If \( c \) and \( c^* \) are systematic vari-
ants, then the system using the encoding scheme will simply compute

the relevant effects rather than have them as incidental behavior. As-

suming, for example, that the encoding is numerical, there will be some

arithmetic relation \( A \) that holds of \( r \) and \( r^* \) just in case these are en-
codings of systematic variants. A system that can compute \( A \) can exhibit

the relevant effects. Any systematicity effects exhibited by a system using

structural representations can be mimicked by a system using a non-

structural encoding.

The objection is overstated, since not all incidental effects can be

mimicked by a system that must compute them. But the objection

misses the point of the argument in any case. The relation \( A \) is not
determined by the representations, but must be programed inde-

dependently in your computational model. Since you can program in

whatever \( A \) you want, the resulting explanation is ad hoc. The only

constraint on which \( A \) you program in is data coverage.

Of course, the permutation rules that drive the explanation of sen-
sitivity to systematicity in a system using structural representation
have to be programed in as well; they do not come for free. But the

theorist is constrained by the form of the representations: you can

only write permutation rules when there are permutable con-

stituents. Models using encoding schemes have no way of enforcing

a comparably principled constraint. To get a feel for this, consider a

vector encoding. Vectors can, of course, be systematic variants of

each other, but since we are assuming an encoding scheme, we are

assuming that systematic variation in a vector does not mirror system-

atic variation in what it represents.\footnote{The use of vector encoding in connectionist systems is therefore different than the use of vector representation in mechanics, in which systematic variation in direction and magnitude does mirror systematic variation in the quantity represented. Vector schemes are structural schemes in mechanics, whereas they are typically encoding schemes in connectionist systems.} But vectors do stand in spatial

relations to one another, and this might make it possible to state

rules of the form: given a representation of \( c \), construct a representa-
tion of a systematic variant by performing spatial transformation \( T \).

It is easy to see, however, that this sort of rule is not comparable to the
permutation rules available to users of structural representations. Suppose we have: given a scene containing \( o_1 \) and \( o_2 \), generate a representation of the same scene with \( o_1 \) and \( o_2 \) swapped by performing transformation \( T \). But now suppose we want a representation in which \( o_2 \) and \( o_3 \) are swapped. We need a new and independent transformation. Because representation of the objects is not separable from the representation of the scene, we cannot state the rule in terms of the objects whose representations are to be swapped.\(^{22}\)

Being principled is a virtue in explanation. But the virtue is purely methodological. Principled explanations are easier to test, but they are no more likely to be true. If there are independent reasons for preferring a connectionist architecture (as there clearly are), then the methodological weakness of the ensuing explanations of our sensitivity to systematicity in language must simply be swallowed. If you are convinced that the mind is a network, you should not be dismayed by the fact that your explanation of our sensitivity to linguistic systematicity is not as easily tested as the other guy's. To repeat, there is no relation between being easily tested and being true. And anyway: no single representationnal scheme can structurally represent every domain exhibiting systematicities to which we are sensitive. If the other guys have not got a different scheme for each case, then some of their explanations are going to be exactly as ad hoc as yours.

VII. PRODUCTIVITY

It is generally supposed that structural representation buys you a lot if the systematicity in the target domain is unbounded. Finite systematicity can be handled by a look-up table, and encodings are as good

\(^{22}\) Permutation rules can be defeated by context sensitivity. Constituents of cartoon faces are thoroughly context sensitive, since whether something counts as a nose, an eye, or a mouth depends solely on its relative position in the whole face. If I try to swap the mouth and the left eye, I get a winking face with a tongue sticking out, not a face with a mouth where the eye should be.

If I make the features sufficiently realistic, however, the effect of context vanishes and the swap works. This, presumably, is because a "sufficiently realistic" feature is just a feature with enough internal complexity to constitute a complete picture in itself.
as structural representations in a look-up table. Structural representation begins to pay off as the amount of targeted systematicity grows large, because structural representation allows you to represent variations in a structure by corresponding variations in the elements of its representation. Permuting a small number of elements in the representation yields great representational power at low cost. Given enough time or memory, you get unbounded expressive power with finite means.

Reflecting on this familiar fact, it might seem that encoding strategies are bound to run into trouble in dealing with unbounded capacities to mirror systematicities. For the point about nonstructural representation is that you cannot rely on intrinsic properties of a representation \( r \) to construct a representation \( r^* \) that represents something systematically related to the thing represented by \( r \). So you cannot get sensitivity to systematicity via a rule that applies to an unbounded set of representations in virtue of their shared form. Because nonstructural representations are arbitrary, it is tempting to conclude that getting \( r \) to generate \( r^* \) whenever the content of it is systematically related to the content or \( r^* \) is going to require something like a look-up table, and hence be essentially bounded.\(^{23}\) If each element of \( D \) is represented by an arbitrary number, how in the world could you effect, with finite means, an unbounded set of pairings among those numbers which mirrors the unbounded systematicities in \( D \)?

We have already seen the answer to this question: at least sometimes, you can compute the effect directly. There is a more interesting answer, however, that is relevant to the dispute between connectionists and their detractors. Connectionists, of course, represent with vectors, and vectors stand in spatial relations to one another. Hence, if \( c \) and \( c^* \) are systematically related, the possibility arises that vector representations of \( c \) and \( c^* \) might stand in a characteristic spatial relation. For some arbitrary \( c \) in \( D \), then, it might be possible to find a systematic variant of \( c \) by effecting a characteristic spatial transformation on its representation. In this way, the structure of the vector space itself can be used to mirror the desired sys-

\(^{23}\) Notice that the problem is not the lack of unbounded representational power. Suitable encodings do provide unbounded expressive power with finite means. The example of Gödel numbering shows this nicely: you get a unique decimal numeral for each sentence, and you get it with finite means. The game is still permuting elements, of course, but without any correspondence between those permutations and corresponding permutations in the target domain. The trouble with encoding is not that you run out of representations. The trouble, such as it is, is that it is structurally arbitrary.
tematic relations, even though the systematically related items have no structural representations in the system.

No actual physical computing system can compute an unbounded function, of course. When we say that our adding machines compute addition, we must be understood to mean that they would compute addition given freedom from resource constraints and from physical wear and tear. The point is that the algorithm is perfectly general; the boundedness of performance arises solely from resource constraints and physical disintegration. Talk of unbounded computational capacity (or productivity, as it is sometimes called) thus makes sense only when it makes sense to distinguish algorithm from resources. G. Schwarz\textsuperscript{24} has recently argued, however, that this distinction is not available to connectionists. The argument is simple. Imagine a connectionist device $\Sigma$ that computes some finite restriction of an infinite function $f$. To add memory to $\Sigma$, you have to add either nodes or precision to already existing nodes. In either case, you must reconfigure the weights to keep performance consistent with $f$. But changing the weights amounts to building a new device; for it is only the weight matrix that distinguishes $\Sigma$ from the infinite class of nets having the same number of nodes and the same activation functions. Think, for example, of training a net to compute a certain function. After training, it computes $f$; before training, it does not. But the only thing that has happened is that the weights have changed. But if we can change an adding machine into a multiplier by changing weights, then weight change is surely change of algorithm. The evident underlying principle is: different function computed implies different algorithm executed.

Tempting as this argument is, I think it has to be rejected. The flaw is in thinking of the function computed as a function from input to output—addends to sum, factors to product—as we typically do when thinking about calculators,\textsuperscript{25} rather than as a function from input and initial state to output and final state as we must when we are thinking about cognitive systems. Consider any AI system. The input-to-output relation it exhibits is a function of its stored knowledge. Change what the system knows and the same input will yield a different output. An airline reservation system cannot price a plan involving a flight it does not know about. But fortunately, given the fact that flights are added and deleted every day, you do not have to reprogram the system every time a new flight becomes available or an old one is dropped. More se-


\textsuperscript{25} We can get away with this because the internal state is generally the same. Forbid use of the clear button, however, and the illusion quickly vanishes.
riously, you cannot think of learning algorithms as pairing inputs and outputs, for the whole point of a learning algorithm is to replace ineffective input-output pairings by better ones. Leaning algorithms, by definition, remain constant over changes in input-to-output relations.

There is nothing wrong with the principle that a different function computed implies a different algorithm executed. The problem is in thinking of the function computed by a connectionist system as a function from inputs to outputs rather than as a function from an activation vector and point in weight space to another activation vector-weight space pair. We should think of a point in weight space as a point in stored knowledge space. From this point of view, we do not build a new network when we change weights any more than we build a new rule-based system when we change its stored knowledge.

The productivity issue, as we have seen, turns on whether it makes sense to idealize away from memory constraints, and, for connectionism, this depends on how connectionist networks are individuated. Schwarz is quite right in supposing that, if we make identity of computational route from input to output a necessary condition of system identity, then we cannot coherently idealize away from memory constraints in connectionist systems. But the argument proves too much, for if we make identity of computational route from input to output a necessary condition of system identity, we cannot coherently describe learning in either classical or connectionist systems, nor, more generally, can we coherently describe any system whose behavior depends on stored information.

**VIII. INFERENTIAL COHERENCE**

Fodor and Pylyshyn illustrate what they take to be the advantage of the classical view with examples of inference.

Because classical mental representations have combinatorial structure, it is possible for classical mental operations to apply to them by reference to their form. The result is that a paradigmatic classical mental process operates upon any mental representation that satisfies a given structural description, and transforms it into a mental representation that satisfies another structural description.\(^{26}\) (So, for example, in a model of inference one might recognize an operation that applies to any representation of the form \(P \& Q\) and transforms it into a representation of the form \(P\).) Notice that since formal properties can be defined at a variety of levels of abstraction, such an operation can apply equally to representations that differ widely in their structural com-

\(^{26}\) Their statement here is a bit sloppy. Vector transformations apply to vectors in virtue of their (spatial) form as well. The point is rather that when you have a correspondence between form and meaning, as you do in any scheme of structural representation, similar formal transformations produce similar semantic transformations. This, as emphasized above, is what structural representation buys you.
plexity. The operation that applies to representations of the form $P \& Q$ to produce $P$ is satisfied by, for example, an expression like \((AvBvC) \& (DvEvF)\), from which it derives the expression \((AvBvC)\) (op. cit., p. 13).

We can reconstruct such truth preserving inferences as if Rover bites then something bites on the assumption that (a) the sentence 'Rover bites' is of the syntactic type $Fa$, (b) the sentence 'something bites' is of the syntactic type $\exists x(Fx)$, and (c) every formula of the first type entails a corresponding formula of the second type... (op. cit., p. 29).

You can see this as a point about systematicity if you take the domain in question to be propositions (and maybe properties or propositional functions), together with their semantic relations. The claim is then that classical representation provides a structural representation of that domain, and thereby facilitates an explanation of our sensitivity to the systematicity it exhibits. The underlying idea is just that logical notation provides a structural representation of interpropositional entailment relations, and hence looks to be a good theory of representation on which to ground a story about human inference.

Two troubles loom. First, there is an undefended assumption here that standard logical notation does, in fact, provide a structural representation of propositions and other inferentially related contents. This assumption is bound to remain undefended pending an independent account of the structure of propositions and other inferentially related contents. You cannot take the soundness and completeness of standard logic as an argument that it provides a structural representation of the domain, since encodings of the standard notations will do as well. Since you do not need a scheme isomorphic to standard logical notation to capture all and only the valid inferences—you can do it with Venn Diagrams, for example—there can be no argument from standard logic and our sensitivity to inferential relations to classical representation.

The second trouble looming is that there is ample empirical evidence that our sensitivity to the inferential relations captured by standard logic is extremely limited. Humans are a good deal better at

---


modus ponens than they are with modus tollens, a fact that classical schemes will have to explain in some relatively ad hoc way. For similar reasons, it is an embarrassment to classical schemes that structure sensitive processing, unless constrained in some way, will treat \((AvBvC) \& (DvEvF)\) in the same way it treats \((P \& Q)\). I am not sure about this particular example, but no one who is up on the human reasoning literature could think that humans are very good at detecting sameness of logical form. Classical schemes based on standard logical notation appear to predict far more sensitivity to entailment relations than humans actually exhibit. Nonclassical schemes might do much better in this regard for all anyone knows at this time. The unwanted predictions generated by classical schemes have to be blocked in some relatively ad hoc way, a fact that appears to level the playing field on which classicists and their opponents have to play."

IX. CONCLUSION

So-called classical representational schemes get into the debate about systematicity only because they provide structural representations of linguistic expressions. Moving away from an exclusive focus on language (sometimes misleadingly disguised as a focus on thought) allows us to see that the real issue has to do with whether sensitivity to systematicity is better explained by appeal to structural representation than by appeal to encoding schemes.

While the alleged systematicity of thought is problematic because there is no noncontroversial way of specifying the structure of propositions, there are many domains that do exhibit systematicities to which we are sensitive. Since not all of these are isomorphic to each other, it follows that either some sensitivity to systematicity cannot be

---

9 See, for example, Rips, *The Psychology of Proof*.

* If there is a classical language of thought, then it might be that unconscious automatic processing is supremely sensitive to identities of logical form. But this is beside the present point, which has to do with whether humans are observed to be sensitive to systematicities in some independently identified domain.

9 Here is an example of how to block a prediction of too much sensitivity: the effect disappears when the representations get complex, because you run out of short-term memory. This would be perfectly legitimate if logical breakdowns occurred when classical representation requires more than seven chunks. Unfortunately, things break down much earlier. No one thinks the problem with modus tollens is a resource problem.
explained by appeal to structural representation, or we must postulate a separate scheme of structural representation for every distinct systematicity to which we are sensitive. It is certainly possible to account for sensitivity to systematicities by appeal to nonstructural encodings of the domain. Although the resulting models will be methodologically ad hoc in a certain sense, they are no less likely to be true for all that.

ROBERT CUMMINS

University of Arizona