There is a certain view abroad in the land concerning the philosophical problems raised by Tarskian semantics. This view has it that a Tarskian theory of truth in a language accomplishes nothing of interest beyond the definition of truth in terms of satisfaction, and, further, that what is missing — the only thing that would yield a solution to the philosophical problem of truth when added to Tarskian semantics — is a reduction of satisfaction to a non-semantic (and ultimately physical) relation. It seems to me that this view either misidentifies the philosophical problem altogether, or encourages a seriously misleading picture of the nature of the problem.

The view I have in mind is nowhere more persuasively at work than in a recent paper by Hartry Field. In this paper Field argues that a Tarskian theory of truth for a natural language is impossible if we insist on Tarski's case-by-case elimination of 'satisfies'. More fundamentally, however, he argues that a Tarskian theory could provide nothing of philosophical interest beyond the admittedly interesting reduction of truth to satisfaction, and his ground for this claim is, roughly, that a Tarskian theory does not reduce its primitive semantic relation — satisfaction — to a non-semantic relation. In what follows, I will concentrate on Field's paper, but my aims are not even primarily critical. I hope to show that a Tarskian theory, where it can be formulated at all, does provide something of interest beyond the reduction of truth to satisfaction (section II), and that the philosophical problem left unresolved by Tarskian semantics is not the problem of reducing satisfac-

R. Cummins

tion to a non-semantic relation (sections V and VI). In the process (sections III and IV) I hope also to clarify somewhat the nature of the problem Field is concerned with, viz., the reduction of the semantic to the non-semantic.

II

What exactly does Tarskian semantics do beyond reduce truth to satisfaction? Certainly the most striking thing is that it provides a recursive definition of satisfaction in non-semantic terms. Let’s quickly review how this is accomplished. Suppose we have a language L whose syntax is that of the first order predicate calculus. For simplicity, I will assume that there are no singular terms. Such is the simple language of classes Tarski originally used to illustrate his method.\(^2\) We begin by defining truth as satisfaction by every sequence of objects. We then explain satisfaction recursively as follows. We say that a sequence satisfies a conjunction just in case it satisfies each conjunct, that a sequence satisfies a negation just in case it fails to satisfy the expression negated, and that a sequence satisfies a universal quantification binding the i-th variable just in case the result of deleting the quantifier is satisfied by each sequence differing from the given one in at most the i-th place. This yields truth conditions for each sentence of L given satisfaction conditions for the constituent atomic expressions of that sentence. To complete matters we must provide a specification of satisfaction conditions for the atomic expressions. If we follow Tarski, we will handle this matter as follows. For each primitive predicate, e.g., \(\subseteq\), we include a biconditional like this:

(1) A sequence s satisfies an expression consisting of the i-th variable followed by \(\subseteq\) followed by the k-th variable iff \(s_i \subseteq s_j\).

(Here \(s_i\) represents the i-th member of s.)\(^3\)

The trick here is to use on the right-hand side of the biconditional the predicate which is mentioned on the left-hand side. Evidently, if we handle matters in this way there will be a separate biconditional for each primitive predicate, so there must be finite number of primitive predicates, and every sentence of L must be a first-order combination of these.

Intuitively, the idea is that anyone who understands the primitive predicates of L better than the truth predicate applied to sentences of


\(^3\) I assume here that the object language L is included in the metalanguage (“mathematical English” in this case). If not, then the use of \(\subseteq\) on the right of the biconditional must be replaced by its translation in the metalanguage.
The Problem of Truth-of

L is in a position to make progress by eliminating the latter in favor of the former. How real the progress is seems to depend only on the primitive predicates of L: if they are unproblematic, progress is relatively great; if they are as unclear as the truth predicate, there is no progress at all, except that the remaining unclarity must be accounted to something other than semantics.

Field points out, quite rightly I think, that the main claim originally made on behalf of Tarskian semantics was that it rendered the truth predicate acceptable from the point of view of physicalism. For, to paraphrase Quine, if we have a physicalist language PL, then the recursion will render application of ‘is true’ to sentences of PL as clear as application of the primitive terms of PL. Since a physicalist is someone who thinks that terms like the primitives of PL are already tolerably clear, the progress for him is relatively great.

Or so it would seem. But Field claims that this sort of maneuver constitutes no real philosophical progress at all. In particular, it is his contention that a recursion based on a set of clauses like (1) above, though it does fix the satisfaction conditions of the expressions of the language, and without using semantic terms, it does not reduce satisfaction (hence truth) to a non-semantic relation. Field describes two ‘truth characterizations,’ T1 and T2, which differ only in that T2 eliminates ‘satisfies’ in the manner of (1) while T1 takes ‘satisfies’ (Field’s term is ‘true’) as primitive. He claims ‘that T2 has no philosophical interest whatever that is not shared by T1.’ So Field wants to claim not only that Tarskian semantics does not reduce satisfaction (hence truth) to a non-semantic relation, but that nothing short of such a reduction would have any philosophical interest not shared by an explanation of truth in terms of satisfaction. The remainder of this section concentrates on this stronger claim. I will return later to the claim that such a reduction would have the sort of philosophical interest Field finds absent in theories like T1 and T2.

4 The recursive definition only allows elimination from contexts of the form ‘s is true (in L)’ where ‘s’ is schematic for a description of a sentence of L from which the syntactic analysis of the sentence itself can be recovered. A standard maneuver allows one to upgrade the recursive definition to an explicit definition, but then understanding the set theory required for the upgrading (in the meta-metalanguage) will be part of the bargain.


Field’s suspicion of Tarski’s procedure centers on the case-by-case elimination via clauses like (1) of applications of ‘satisfies’ to primitive terms. Because, ultimately, the elimination turns thus on a “list,” it is felt that we are told only what satisfies what, and left in the dark as to the nature of satisfaction.

By similar standards of reduction, one might prove that witchcraft is compatible with physicalism, as long as witches cast only a finite number of spells: for then ‘cast a spell’ can be defined without use of any of the terms of witchcraft theory by listing all the witch-and-victim pairs.\(^7\)

Of course a list of all the witch-and-victim pairs wouldn’t make witchcraft acceptable to physicalism. But it might very well make the predicate ‘\(x\) casts a spell over \(y\)’ acceptable to physicalism (assuming only witches cast spells). For such an enumeration would fix application of ‘casts a spell over’ without commitments as to what sort of process (or whatever) casting a spell is. If the enumeration could be accomplished without trafficking in terms unacceptable to physicalism, then the conscientious physicalist, armed with such an enumeration, would be in a position to ask, without violating his methodological principles, “What sort of physical process (if any) is casting a spell?”

This is of some importance: the existence of an unproblematic enumeration of the witch-and-victim pairs prevents the investigator from cheating by arbitrarily legislating on cases not fitting his theory. (Suppose we think that to cast a spell is just to hypnotize. Presented with a counter-example, we cannot simply regard it as “some other phenomenon” if the case is on the list. This would at least require being more confident about the theory than the list.) Nor would failure to discover a physical process under such circumstances necessarily damage the physicalist position. Physicalism is surely not committed to the existence of an interesting physical process or regularity or whatever corresponding to the extension of every predicate which receives methodological security clearance.\(^8\) This would be tantamount to requiring that physicalists have an account of spell casting before they can introduce ‘cast a spell’ into the language of science. Scientists would be in the position of trying to open Pandora’s box while standing on the lid.

A corresponding point can be made, in the abstract anyway, for Tarski’s contribution to semantics. Even if we allow (and I think we

\(^7\) *Ibid.*

\(^8\) Surely if \(F\) and \(G\) are acceptable, so is their disjunction. Yet the disjunction need have no physical significance: ‘\(x\) bonds with or dissolves in \(y\)’ identifies no significant physical relation. What is required is that every case of “bonding-or-dissolving” be physically explicable somehow or other.
must) that Tarski has not shown us how to reduce satisfaction to a non-semantic relation, he has perhaps shown us how to launder 'satisfies', and hence 'is true', at least in special cases. What (mere) eliminability shows is that we can claim to know what satisfies what without claiming to know what satisfaction is. Consider how welcome a similar claim concerning synonymy would be: an uncontroversial way of enumerating the synononomous pairs would lay to rest a lot of philosophical scruples. The problem has been precisely that attempts to say what synonymy is collapse into the trivial or indefensible in the absence of any independent way of telling what is synononomous with what. All this should make us suspicious of the claim that, given the explanation of truth in terms of satisfaction, nothing short of a reduction of satisfaction could provide anything of further philosophical interest. There is a distinction to be drawn between making satisfaction acceptable to physicalism, and making 'satisfies' acceptable to physicalism. The latter is not trivial nor does it appear that it must wait on the former. Indeed, the dependence seems to go the other way. Only when there is a physicalistically acceptable way of specifying what satisfies what can the question about the physical nature of satisfaction (if any) be treated non-arbitrarily. Meanwhile, armed with the Tarskian elimination of 'satisfies', physicalists wishing to do model theory, for instance, can do so without trafficking in semantic terms at all. Insofar as physicalists are hesitant about model theory because doing it seems to require using terms like 'true,' and 'denotes' to which no physical meaning has been attached, Tarski's work removes the source of the hesitation by making it possible to do model theory without using the troublesome terms.

Now one might object to this on the grounds that, lacking a reduction of satisfaction, there could be no reason for confidence in a list-type elimination. This is tantamount to objecting that we could have no reason to accept a theory which purports to say what chemically bonds with what, or what characteristics are inherited, unless we also know what physical features of atoms and molecules explain chemical bonds, or what physical facts underlie inheritance. Yet valence theory and Mendelian genetics were in fact well entrenched, and for good reasons, long before the facts these theories codify were explicable at the physical level. So, in general, we must admit that it is quite possible, and very likely necessary to the development of science, to establish theories which fix the extension of a predicate ('bonds chemically with', 'is inherited from') without knowing how the facts codified in

---

9 See, for example, Quine's discussion in chapter one of *The Philosophy of Logic* (Englewood Cliffs: Prentice Hall, 1971), especially the first full paragraph on page three.
such theories (oxygen bonds with hydrogen, offspring of parents neither of which can taste phenylthiocarbamate cannot taste it either) are to be explained beyond their subsumption under the laws of the theory in question.10 It is true that we must be able to establish some of the facts without appeal to the theory; for instance, we must be able to distinguish compounds from mixtures and inherited from acquired characteristics. But it is obviously possible to test the extensional adequacy of a theory in the absence of reductions of its primitives. In particular, the extensional adequacy of a Tarskian truth theory for a language L is established by the fact that the theory has all of the T-equivalences of L as logical consequences.11

III

The above, of course, is a somewhat unsympathetic response to the main thrust of Field’s paper. For what Field and others have wanted to claim is that it is a genuine reduction of satisfaction that is really wanted, and that it is important to realize that eliminability of the sort Tarskian semantics would provide is not reduction. Tarskian semantics would allow us to fix satisfaction (and hence truth) conditions without trafficking in semantic terms, but there is no more prospect than before Tarski of saying what sort of physical relation (if any) satisfaction is. And if we could say what satisfaction is, we could dispense with the case-by-case elimination. If we think we make satisfaction acceptable by

10 Curiously, the objection does seem to have some force in the witchcraft case. It is hard to imagine how we could have any confidence in a list-type elimination of ‘casts a spell over’ without having some idea how spells are to be explained. This may be sheer prejudice, however, deriving from the fact that we cannot in fact formulate a theory which identifies the spells.

11 Field, of course, does not dispute this point. See pp. 361-2 and footnote fourteen of “Tarski’s Theory of Truth.” What Field does dispute is the utility of an extensionally adequate theory which eliminates semantic terms without reducing them.

There is another way to see the benefit that may accrue to a theory in virtue of the availability of a recursive elimination of its otherwise primitive terms. For, in principle anyway, such an elimination can be seen as a technique for promoting certainty about the use of the eliminable term in a finite number of clear cases into certainty about the use of the term generally: if we can find it in ourselves to accept a finite number of clauses like (1), together with a theory like Field’s T1 which gives the satisfaction conditions of complex expressions as a function of the satisfaction conditions of constituents, then we can find it in ourselves to accept statements of satisfaction and truth conditions generally. We thus reduce our total epistemological commitment in semantics to whatever is involved on the right hand side of clauses like (1).
The Problem of Truth-of

making 'satisfies' acceptable, we accord a merit to the case-by-case elimination it doesn't deserve. And if we won't be happy until we know what satisfaction is, then we needn't worry greatly about the elimination of 'satisfies', since that will come automatically and directly when satisfaction is explained. This, I think, is the main thrust of Field's paper, and I would be surprised if it did not find many sympathizers.

If we are to evaluate this line of thought, we must answer two questions. First, in what sense does a Tarskian case-by-case elimination of 'satisfies' in favor of non-semantic terms fail to be a reduction of satisfaction to a non-semantic relation? And second, why is it so crucial that a reduction (and nothing less) be provided?

In order to make persuasive the claim that the Tarskian elimination of 'satisfies' falls short of reduction, we are invited by Field to imagine an analogous case-by-case elimination of 'x has a valence of y' from the chemical theory of bonding. In particular, we are to imagine that scientists adopt a definition of the following form.

(2) (E) (n) (E has a valence of n iff E is potassium and n = +1, or... or E is sulphur and n = -2).\(^\text{12}\)

Since there are a finite number of elements, we can suppose (2) covers all attributions of valences to elements (atoms). Attributions of valences to everything else we may suppose covered by rules yielding valences of complexes given the valences of the constituent atoms. Here is what Field says about this.

But, though this is an extensionally correct definition of valence, it would not have been an acceptable reduction; and had it turned out that nothing else was possible — had all efforts to explain valence in terms of structural properties of atoms \(^\text{13}\) seemed futile — scientists would have eventually had to decide either (a) to give up valence theory, or else, (b) to replace the hypothesis of physicalism by another hypothesis (chemicalism?).\(^\text{14}\)

---


\(^{13}\) Field doesn't explain what he means by "structural properties of atoms." A clear case would be attribution of tab-and-slot or hook-and-eye type shapes. We might then explain that molecules hold together in the manner of jig-saw puzzles. Actual properties of atoms — number and kind of elementary particles together with their quantum states — hardly seem "structural," but I think the point is fairly clear. Potassium bonds with oxygen. To take this as a primitive dispositional property of Potassium atoms is "chemicalism"; physicalists must be able to explain why potassium bonds with oxygen by appeal to laws governing the constituent elementary particles. With this understanding, I will continue to express the matter in Field's way.


109
So the trouble with Tarskian semantics is that it has not shown us that we do not have to choose between (a) giving up semantics, and (b) adopting "semanticalism" and abandoning physicalism. The claim is that Tarski's work does not put us in a position to retain semantics without violating the tenets of physicalism by a commitment to irreducibly semantic facts.

I have no quarrel with Field's contention that Tarskian semantics does not reduce satisfaction, but I do disagree with Field about the philosophical importance of reducing satisfaction, and for this reason it is important that we be as clear as possible about why we are willing to say a reduction has not been achieved. With this in mind, let us see what unsympathetic opposition will force us to say in clarification.

To begin with, we may admit that (2) does not fix the valences of the elements on the basis of structural properties of atoms. If (2) is the best we can do with valence, then we are stuck with irreducibly chemical facts, for we would not be in a position to specify, for each valence, conditions on atomic structure which would be necessary and sufficient for possession of that valence. But (1) and its kin do fix satisfaction conditions on the basis of structural properties of expressions, e.g., properties like shape in the written language which are explicable in terms of the relative spatial distribution of elementary particles. Of course, as Field argues in the early pages of "Tarski's Theory of Truth," it may not be possible to fix satisfaction conditions for expressions in a natural language by appeal to formal properties only. If not, the project breaks down for natural languages, though not for all languages. This beside the present point, however, which is Field's claim that the Tarskian elimination would have no interest even if it could be carried out.

This unsympathetic reply shows that, if Field's argument is to have any force, more must be required of the physicalistic semanticist or chemist than that satisfaction conditions or valences be fixed by reference to structural properties of the expressions or elements involved. After all, (2), by this criterion, could be upgraded to a physicalistic definition by the mere addition of a structural way of discriminating the elements. For suppose that 'Pk', is a predicate applying to things having a structural property, irrelevant to the explanation of bonding, shared by all and only the potassium atoms. Then we could replace 'E is potassium' in (2) by 'E atoms are Pk' without affecting the truth of (2). Similar replacements for each element would lead to (3):

(3) (E) (n) (E has a valence of n if E atoms are Pk and n = +1, or...or E atoms are Ps and n = -2).

(3) "defines" valence for elements in terms of structural properties, but it does not, in the sense Field wants, "explain valence in terms of structural properties of atoms." We are not told what structural property having a valence of, say, +1 is, i.e., it does not identify having
The Problem of Truth-of

a valence of $+1$ with having a certain structural property, and hence, it
does not open the way to a physical explanation of the bonding
behavior picked out by the attribution of that valence. This is why even
(3) would leave a physicalist with a serious problem. For a potassium
atom to have a valence of $+1$ is for it to have a certain disposition to
combine with other atoms, i.e., for a certain regularity to characterize
its chemical behavior. Physicalism requires that this regularity in the
behavior of potassium atoms be explicable in terms of structural
properties of atoms. (3) provides a structural property distinguishing
each element, but provides no hint of the required explanations. On
the other hand, suppose we knew the following.

(4) (E) (n) (E has a valence of n iff (n is positive and E atoms have n
free hooks or n is negative and E atoms have n free eyes)).

Armed with (4) we could readily explain why two potassium atoms
bond with one oxygen atom: one hook to each eye. This, presumably,
would constitute real progress toward vindicating physicalism.\(^15\)

If we take this as paradigmatic, then progress toward making
semantics compatible with physicalism would require progress toward
explaining the regular connections between the “structure” of a
primitive expression and its satisfaction conditions. For, just as valence
theory says what bonds with what, so Tarskian “satisfaction theory”
tells us what satisfies what. And since (4) would be a reduction of
valence because it would explain in a certain way why the bonds that
occur do occur, so a corresponding reduction for satisfaction theory
should explain in an analogous style why the satisfaction that obtains
does obtain. For example, it should explain why ‘$x_1 \subseteq x_2$’ is satisfied by
$\{x\mid(x = x)\}, \{x \mid x = x\}, ..., \text{and not by } \{x \mid (x = x)\}, \{x \mid (x = x)\}, ...$.

It is obvious that explanations of this sort are not forthcoming
from Tarskian semantics.\(^16\) So if by “reducing satisfaction to a non-
semantic relation” we mean providing an account of satisfaction

\(^{15}\) The actual facts, I think, provide nothing like this clear a case—hence the
fantasy.

\(^{16}\) Unless, of course, one thinks that all that can be done in the way of explaining
this fact is to show that $\{x \mid (x = x)\} \subseteq \{x \mid x = x\}$ whereas $\{x \mid x = x\} \subseteq \{x \mid (x = x)\}$. There is some initial plausibility to this view. Asked why ‘red $x_3$’
is satisfied by $\langle \text{snow, cheese, blood, } \rangle$ a natural response would be to
answer, “Because blood is red.” If we accept a redundancy theory of truth,
then Tarskian semantics provides all there is to provide in the way of ex-
plaining “the facts of satisfaction.” The trouble with this line is that we can
easily construct a language containing the expression ‘Red $x_3$’ such that the
proper answer to the question is, “Because blood is a liquid.” This seems
explicable only if we assume that, in constructing such a language, we
would be establishing a relation between that expression and liquids. If we
are to explain matters in this way, then we must, eventually, explain the ex-
plaining relation.
which would facilitate explanations of this kind in the way the hook-and-eye theory would facilitate explanations of bonding, then it must be admitted that Tarskian semantics does not reduce satisfaction to a non-semantic relation. This answers our first question.

IV

What sort of thing would facilitate the crucial explanations? We don’t know what does explain such matters, but we ought to be able to come up with a fantasy on a par with (4). This turns out to be rather difficult. Consider (5):

(5) For any predicate p, the expression consisting of p followed by the i-th variable is satisfied (in L) by a sequence s just in case speakers of L would (normally) respond to s by uttering p.17

Armed with (5), the semanticist committed to physicalism could dispense with the case-by-case elimination of ‘satisfies’. More importantly, it might seem that he could explain why sequences beginning with red things satisfy ‘Red x1’ and not ‘Heavy x1’: speakers of L would (normally) respond to a red thing by uttering ‘Red’, but not by uttering ‘Heavy’. (Unless all red things are heavy.) But perhaps appearances are misleading here: how are we to know that (5) is “reductive” in the manner of (4) rather than merely “eliminative” in the manner of (3)?18

Once this question is raised, it is surprising how difficult it is. There seems to be no intuitive pull to assimilate (5) to (4) rather than to (3), or vice versa. Of course, we have no general account of what makes something like (4) reductive as opposed to merely eliminative like (3). We do have intuitively clear cases, as the contrast between (4) and (3) shows, but (5) is evidently not one of them. This is not, in itself, very significant, but is suspiciously difficult to come up with a fantasy about satisfaction which, like (4), would be clearly reductive. (5), in fact, is the best I can do along these lines, and this leads me to wonder whether I have any clear idea of what a reduction of satisfaction would be.

Perhaps if we knew why (4) is reductive whereas (3) is not we would be in a better position to say whether (5) is reductive or not. I think there is only one available answer to the question about the difference

17 Of course, (5) could be of no use to physicalists unless a physicalistic account of the psychological notion of response were available. I propose to ignore this as it raises problems tangential to my concerns here.

18 I use “reductive” in a way analogous to the use of “law-like”: (4) is reductive because it is the sort of thing which would actually allow us to reduce valence to a physical property if it were true.
between (3) and (4): (4) would allow us to identify having a valence of +\( n \) with having \( n \) free hooks, and having a valence of -\( n \) with having \( n \) free eyes, whereas (3) would allow no comparable identification of valence with a physical property, because (4) but not (3) would "really explain", and not merely subsume, the chemical bonds specified by valence theory. Of course, (3) is something of a straw man in this respect: we could not identify having a valence of +1 with being \( P_k \), having a valence of -2 with being \( P_s \), and so on. For one thing, both oxygen and sulphur have a valence -2, whereas only sulphur is \( P_s \). But we need not rest the case on that. We might discover, for each valence, a "structural" property shared by just the atoms with that valence, and yet not be prepared to identify having a valence of \( n \) with having the corresponding property. To elaborate on the fantasy of (4), imagine that every atom with \( n \) free eyes has a mean diameter of \( kn \) units, for some constant \( k \), and that every atom with \( n \) free hooks has a mean diameter of \( cn \) for some constant \( c \neq k \). Then we would have the following:

\[
(3^*) (E) (n) (E \text{ has a valence of } n \iff \text{if } n \text{ is positive and } E \text{ atoms have a mean diameter of } cn, \text{ or } n \text{ is negative and } E \text{ atoms have a mean diameter of } kn).
\]

Evidently \( (3^*) \) would be no better than (3) for explaining bonding. Given valence theory purged of "has a valence of \( n \)" via \( (3^*) \), we could deduce that two atoms of potassium bond with one atom of oxygen, but we still could not explain why this is so in the way we could given (4). Valence, from the point of view of physicalism, is something which is supposed to explain bonding in just the way in which hooks and eyes would and diameters would not.\(^{19}\)

If this is right, then (5) could form the basis for a reduction of semantics to psychology only if (5) would "really explain," and not merely subsume, the facts of satisfaction and hence make it plausible to identify satisfying an expression with the disposition to prompt utterances of it. This, I think, would be about as plausible as identifying valences with diameters. This may be simply because (5) is a bad try at make-believe. But I think the problem runs deeper. It is not that (5) obviously could not explain the facts of satisfaction, but rather that, unlike the valence case, we have no idea what it would be like to explain such facts. If this is the correct diagnosis, we won't be able to tell whether alternatives to (5) are reductive either: the question of reduction in semantics lacks a clear sense because we do not know what it would be to explain the facts Tarskian semantics codifies.

---

\(^{19}\) I know of no way to characterize illuminatingly the sort of explanation (4) but not \( (3^*) \) would provide. Plainly, more than subsumption under general laws is involved, since either (3) or \( (3^*) \) would open the way to deductive-nomological explanations of bonding, and in "structural terms" at that!
R. Cummins

I will try to defend this claim shortly. First, however, I want to return to the valence example to make a further point about the connection between physicalism and the possibility of reductions in the manner of (4).

If we could identify having a valence of \( +n \ (-n) \) with having \( n \) free hooks (eyes) then we could explain the facts about bonding codified in valence theory. But this sort of explanatory progress could be had without a completely general account of the physical nature of valence, i.e., it could be had even if having a valence of \( n \) is identified with having one sort of physical property in some atoms, and with having a different sort of physical property in other atoms. Indeed, even if the physical facts underlying having a valence of \( n \) turned out to be different for each element, the explanatory payoff characteristic of (4) but not (3*) might still be forthcoming. In this situation we could say that oxygen’s having a valence of -2 is its having a certain physical property \( \phi_O \), whereas sulphur’s having a valence of -2 is its having \( \phi_S \), but we could not say anything of the form ‘An atom’s having a valence of \( n \) is its having \( \phi \)’. Yet we would still be in a position to explain the facts codified in valence theory in a physically acceptable way. The resulting variety of distinct explanations of bonding would imply that the chemical bonds dealt with in valence theory are physically heterogeneous, but not that the facts codified in valence theory are not physically explicable every one.

Let us suppose, then, that scientists adopt something like the following:

(6) (E) (If \( E \) is oxygen then \( E \)-atoms have a valence of -2 iff \( E \)-atoms are \( \phi_O \), and if \( E \) is sulphur then \( E \)-atoms have a valence of -2 iff \( E \)-atoms are \( \phi_S \), and if \( E \) is hydrogen then \( E \)-atoms have a valence of +1 iff \( E \)-atoms are \( \phi_H \), and ... etc.)

And let us suppose that assuming oxygen atoms to be \( \phi_O \) and hydrogen atoms to be \( \phi_H \) allows us to explain hydrogen-oxygen bonds in the deep way characteristic of (4) but not (3*). This, I think, would vindicate valence theory in the eyes of the physicalist. But I think it would be misleading at best to say in these circumstances that a reduction of valence had been achieved. One fairly natural thing to say would be that while valence theory had been vindicated in the eyes of physicalism, there turned out to be no such thing as valence. Let us distinguish cases like that envisaged in (4) in which a unified account is forthcoming from cases like that envisaged in (6) in which a unified account is not forthcoming by saying that in the former cases we have a reduction of valence, whereas in the latter cases, valence is analyzed away.20 Since it is not always clear whether an account is “unified,” there will undoubtedly be a fuzzy area of intermediate cases, but I don’t think this will affect the succeeding discussion in any important way.
What these remarks about (6) show is that it is the possibility of explaining in a certain way the facts codified in valence theory which vindicates that theory in the eyes of physicalism, and that the required explanations may be forthcoming without achieving anything we should want to call a reduction of valence. If this is right, then we should be skeptical of the claim that nothing short of a reduction of satisfaction could make Tarskian semantics acceptable to physicalism. It is a certain kind of explicable that physicalism requires, and a reduction of satisfaction is no more a necessary condition of the explanations required by physicalism to vindicate semantics than a reduction of valence is a necessary condition of the explanations required by physicalism to vindicate valence theory. Once we allow this, the past, success of physicalism no longer counts as a reason to expect that satisfaction could not, from a physical point of view, turn out to be something rather different for each primitive predicate.

V

I suggested five paragraphs back that no alleged reduction of satisfaction to a physical relation would be plausible because we have no clear idea what it would be like to explain the facts of satisfaction. I now want to return to this point. In this section, I will argue that it is more likely that the problem Field has identified will be solved (if it has a solution) by analyzing satisfaction away rather than reducing it. I argue this because I suspect that the vision of a unified physical account of satisfaction has encouraged some (including Tarski, perhaps) to assimilate the problem of clarifying satisfaction to the problem of vindicating satisfaction in the eyes of physicalism. Blurring this vision prepares the ground for my claim in VI that the assimilation is a mistake, and that the primary philosophical task is to clarify the concept of satisfaction in a way which would make it possible to distinguish "real explanations" of the facts of satisfaction from mere subsumption. Once we know what it would be like to explain such facts (assuming it can be done), details of the actual explanation will be of little interest to the semanticist.

20 For the purposes of the ensuing discussion, this need be regarded as no more than a terminological stipulation.

21 Field may mean to count as reduction things like (6) which, as I say, analyze valence away. I doubt it, though: I don't think Field could have been thinking that something analogous to (6) could throw any philosophical light on semantics, and he clearly was thinking that a "reduction" would.
R. Cummins

If we are to take seriously the suggestion that satisfaction is to be reduced to a physical relation, we shall do well to make a few preliminary moves. For Tarski, satisfaction is a relation between an expression-type and an infinite sequence. We shall avoid some vexing problems if, instead, we begin by construing satisfaction as a relation between a token of an expression and an infinite sequence: We may then explain this relation as follows:

(7) If x is a token in L consisting of a token of a predicate of L followed by a token of the i-th variable of L, then x is satisfied in L by a sequence s iff x applies to (is true of) the i-th member of s.\(^22\)

This leaves us with a relation — application — which holds between a token of an expression and an entity of whatever sort gets talks about in L. If L is a physicalistic language, then these will be certified physical entities of some sort or other. There is nothing about the relata of application, then, that need inhibit the zeal of the physicalist.

One consideration which might lead one (in one’s metametalinguistic reflections) to suppose that expressions of the form ‘x applies to y’ are not true in virtue of one sort of physical relation holding between x and y is that the physical characteristics of the alleged relata as we know them seem obviously irrelevant. There is, for instance, no physical characteristic a token must have to apply to just the red things. In general, a class of tokens having the same application is not a physical kind. This has seemed to some to undermine the possibility of a reduction of application to a physical relation. There are two points we must clarify if we are to evaluate this line of thought: we must determine whether or not the premise of the argument is true, i.e., whether or not the tokens having the same application conditions are not a physical kind, and we must examine the principle appealed to, i.e., that the physical hetrogeneity of the class of tokens having the same application conditions would block a reduction of application to a physical relation. I will discuss the principle first, then the premise.

To reduce application to a physical relation as opposed to analyzing it away, it is necessary (though perhaps not sufficient) that there be a unified explanation of the facts codified in “application theory” — i.e., the theory which simply tells us what applies to what. Now I think that it must be admitted that if sets of things having the same application conditions turn out to be mixed bags from the point of view of physicalist taxonomy, then we must relinquish the hope of the required unitary account of the facts of application theory. This is simply

\(^22\) This deals only with monadic predicates. Generalization to n-adic predicates involves obvious but (for present purposes) irrelevant complications. Notice, by the way, that application is not language-relative on this construal.

116
because such an account would have no feature of applying tokens which it could appeal to as explaining why tokens having the same application have the application they do and no other. Of course, each token might have a physical feature appeal to which explains why it has the application it does, but this would allow us to analyze application away, not to reduce it.

Let us try to clarify this line or argumentation by illustrating it in a somewhat different physical setting. Consider "dissolution theory," i.e., a theory which fixes the extension of the predicate 'x dissolves in y' but which does not purport to say what dissolving is, or to explain in any deep way why the extension is what it is. To achieve what I have been calling a reduction of dissolution to a physical relation, we require a unified account of the facts codified in this theory. Now if this is to be possible, it must not be the case that, e.g., the things which dissolve in water are, at every level of physical description, a mixed bag. For such an account must, among other things, answer the question, "Why do the water-soluble things dissolve in water?" And it must answer this question in the manner characteristic of (4) but not (3*). (In particular, it must not hold simply that things dissolve in water because they are water-soluble.) Such an explanation cannot get off the ground unless it can begin, "The water-soluble things are all \( \phi \) whereas other things are not, and...," where \( \phi \) is some non-disjunctive property antecedently recognized as physical (presumably a matter of molecular and sub-molecular structure) which can plausibly be identified with being water-soluble. If the best available information indicates that there is no such \( \phi \), then we can either (i) acquiesce in "chemicalism," i.e., regard 'x dissolves in water' (and hence 'x dissolves in y') as a primitive element in our scientific taxonomy, 23 or (ii) try to analyze away water-solubility, i.e., try to explain water-solubility case-by-case.

Although the principle involved is difficult to generalize, I think we must concede that the argument under consideration is valid: if classes of tokens having the same application are physically heterogeneous in the sense in which we have imagined the class of water-soluble things is physically heterogeneous, then a unified physical account of the facts of "application theory" would be block-

---

23 This would constitute an abandonment of physicalism as we know it, for we should have to admit that the behavior of physically complex entities (molecules) is not explicable solely by appeal the laws governing their ultimate physical constituents. (To see that this "componentialism" is central to physicalism, imagine what we would have to say if it turned out that the gravitational acceleration experienced by a hydrogen molecule at a distance \( d \) from a mass \( m \) were not double that experienced by a hydrogen atom at a distance \( d \) from \( m \).)
ed, and we should have to try to analyze application away, or ac-
quiesce in 'semanticalism.'"

Let us turn, then, to the premise, viz., that classes of tokens having
the same application are physically heterogeneous (in the sense re-
quired by the principle). This premise is, of course, quite obviously
true, even if we restrict our attention to first-order languages. So if
application is construed as relating tokens and objects, a physicalistic
reduction of application seems out of the question. However, a
slight reconstrual will resuscitate the issue. For even if we admit that a
class of tokens having the same application may be as heterogeneous
as you like, it may yet be that, at some level of analysis, they are all
produced in the same way, or all produced under the same psy-
chological conditions, or something of the sort. To allow for this
possibility, let us understand application as a relation between an ac-
tual tokening of an expression (i.e., its actual utterance or writing or
whatever on a particular occasion) on the one hand, and objects (or
events or whatever) on the other. In line with this reconstrual, we may
replace (7) by (8).

(8) If x is a tokening in L consisting of a tokening of a predicate of L
followed by a tokening of the i-th variable of L, then x is satisfied by
a sequence s iff x applies to (is true of) the i-th member of s.24

The advantage of this reformulation is that, whereas a class of
tokenings-in-L having the same application may not be a physical kind
in virtue of their manifest features, they may well be a physical kind in
virtue of their etiology. Here we simply apply a lesson from the
philosophy of psychology. Tokenings are behaviors, and, in general,
what makes two events instances of the same type of behavior is not
their manifest properties, but their etiologies.25 Perhaps, then, this
formulation will allow the reductionist to avoid the objection lately
rehearsed, for perhaps it is the case that a class of tokenings having the
same application will be characterized by one sort of physical etiology.
This, at any rate, seems the most plausible suggestion available to the
reductionist.

Yet the suggestion is not, after all, very plausible. As Fodor remarks,
"...we have no right to assume a priori that the nervous system may not
sometimes produce indistinguishable psychological effects by radical-

24 If (8) is to be useful, we must assume that events which are tokenings in L can be
distinguished from events which are not. Usually, the token produced will settle
this matter, but where it does not we are free to appeal to the etiology of the
event in question.

25 See, for example, J. A. Fodor, Psychological Explanation (N.Y.: Random
ly different physiological means." In general, it seems clear that it will be difficult to defend a materialist theory of mind which insists that every mental state (or event or whatever) which is an instance of a certain mental/psychological type should be a physical state (or event, or whatever) which is an instance of a single corresponding physical type. Presumably, the etiologies leading to tokenings having the same application will be functionally equivalent, but from the fact that a set of etiologies form a functional type we cannot infer that there is any interesting underlying similarity expressible in a physicalistic vocabulary.

We have not yet seen any reason to suppose that the etiologies of tokenings having the same application are not a physical kind (i.e., not such as to allow the required unitary explanation), but we have seen that there is no good reason to suppose they are, and hence no good reason to suppose a reduction of application to a physical relation is possible. More importantly, having come this far, it seems clear that nothing of much philosophical interest turns on whether reduction emerges as the right strategy rather than analyzing application away, and certainly the details of a reduction or analysis could not be expected to throw any philosophical light on application. That vindication of "application theory" in the eyes of physicalism could be achieved somehow or other would be philosophically important, and it must be admitted that Tarski's work shows nothing of the kind. But Tarski's failure to do this should not be regarded as a failure to say anything illuminating about application, since a physicalistic vindication of application theory would not provide that sort of philosophical illumination anyway. Let us try now to get clearer about why this is so.

26 Ibid., p. 118.


28 One might dispute even this: perhaps different persons (or the same person on different occasions) produce tokenings having the same application via processes which have radically different functional analyses (i.e., by executing different programs) just as two adders may produce sums in virtue of realizing radically different algorithms.

29 It seems to me that the well-attested fact of redundancy in the brain, and the fact that linguistic performances are manifestations of learned capacities acquired under widely divergent conditions, together make it likely that etiologies of tokenings having the same application are functionally equivalent at most. But I would not insist on this point.
VI

Some tokenings apply to a thing in virtue of its being red, others apply in virtue of something else, and still others don't apply at all. The details of this matter are what I have been calling "the facts codified in application theory," e.g.,

(F) (x) (Tokenings in English of 'fish' apply to x iff x is a fish)
(G) (x) (Tokenings in English of 'red' apply to x iff x is red)

and so on. What I want to argue in this section is that a reduction of application to a physical relation would throw no appreciable light on the philosophical problem of truth. My contention will be that the philosophical problem is not to explain facts like (F) and (G), but to figure out what sort of facts they are and what it would be like to explain them. Once this is accomplished, the actual physical explanations will hold no philosophical interest beyond the fact that they have a physical explanation of some sort or other.

Let us return again to the valence case. The analogy between valence theory and application theory is misleading in that there is no philosophical problem of valence. We already know what it means to say of something that it has a valence of n, viz., that it combines with this and that in such-and-such proportions, where we could actually give a list (or a procedure for generating it) together with the correlative proportions. To say of something that it has a certain valence is just to attribute a very complicated but precisely specifiable disposition to it. There is a scientific problem here, namely, in virtue of what (presumably structural) features do things have the valences they do? This problem is solved when we establish something like (4) which explains why the combinations specified by valence theory do occur.

Since the scientific problem is the only readily apparent problem raised by valence theory, we have been misled by this analogy to suppose that the philosophical problem about application is a problem analogous to this scientific problem about valence.

It is not. The philosophical problem is that we don't know, or anyway cannot say explicitly, what 'application' means. A consequence of this is that, unlike the situation with respect to valence, we don't know what sort of facts application theory codifies, and hence we don't know what it would take to provide an explanation of them which would be explanatory in the manner characteristic of (4) but not (3*). This is why we are in the dark as to whether (5) belongs with (4) or (3*). If we could say what 'application' means, then scientists could get on with the scientific problem. The details of their solution, assuming they find one, would have exactly the philosophical interest that the solution to the problem about valence has: physicalism would have passed another test. How the test was passed would not concern semanticists (qua semanticists) at all.
With this in mind, let us take another pass at some not so fantastical fantasy and see how our intuitions fare.

(F) and (G) and the like are true, it seems to me, because they are rules of English. They are true, therefore, in much the same way a stipulation is true, or a rule of a game. This double-faced aspect of rules is familiar enough, but it won’t hurt to have an illustration in mind. If someone says, “Bishops that begin on white squares move only along white diagonals,” he expresses an empirical truth (somewhat idealized) about the movement of white bishops. But the facts underlying this truth are a bit special: it isn’t as if starting on a white square constrains a bishop to remain on white diagonals in the way that starting on a certain track constrains a train to remain on that track, but rather the people who move the bishops have accepted as a stipulation that bishops starting on white squares move only along white diagonals. Accepting this stipulation is a necessary condition of being a player of the game. So far, this isn’t very helpful because we cannot set out explicitly what it means to say someone accepts a stipulation of this sort. In order to see how this problem of philosophical analysis hooks up with the issue of reduction, let us cover our ignorance about stipulation with some symbols.

Let us suppose we have discovered a general behavioral description $B_x$ and a general description of psychological conditions $C_x$ which can be specialized in systematic ways to explain what it is for someone to accept a stipulation thus:

(9) For someone to accept the stipulation that bishops starting on white squares move only along white diagonals is for it to be the case that his behavior would satisfy $B_{white\, bishop}$ were he to satisfy $C_{white\, bishop}$

(10) For someone to accept the stipulation that rooks move only along rows and columns is for it to be the case that his behavior would satisfy $B_{rook}$ were he to satisfy $C_{rook}$.

And so on. Having done this much, there would be a well specified scientific problem concerning the physical facts underlying the acceptance of such stipulations. Every case of accepting a stipulation is doubtless some physical state or other, though probably not a state of the same physical sort in every case. The problem raised by (9), for instance, is to find some physical state $\phi$ such that someone’s being in $\phi$ explains why his behavior would satisfy $B_{white\, bishop}$ were he to satisfy $C_{white\, bishop}$. For, given (9), this would be to identify the physical fact in virtue of which someone accepts the stipulation that bishops starting on white squares move only along white diagonals. If it turns out that this problem has a different solution for each stipulation, or for each case of a distinct person accepting a distinct stipulation, then physicalistic scientists will have to be content with analyzing away accepting a stipulation rather than reducing it. How this issue
turns out is philosophically uninteresting. The philosophical problem is to construct and defend real analyses to play the role of the likes of (9) and (10). Until this is achieved, there isn’t a well specified problem for physicalistic scientists to attack.30 This discussion applies mutatis mutandis to the likes of (F) and (G). We can go some way towards understanding (F) when we see that it holds because it is a rule of the language. And perhaps it is not far wrong to explain this by saying that a necessary condition of speaking English is accepting the likes of (F) and (G) as stipulations.31 But this does little more than indicate a possible line of inquiry, for we can no more set out what it is for someone to accept a stipulation in the sense required than we can set out what it is for something to be a rule of language or, for that matter, what it is for a tokening in English to apply to x just in case (and because) x is a fish. If we could do these things, we could specify a scientific problem which would constitute a test of physicalism. But a solution to that problem would cast no additional philosophical light on semantics at all.

February 1975

30 Since the scientific problem is not well-specified, there is no way to assess alleged solutions to it. Hence our feeling of being at sea when confronted with (S).

31 Of course, accepting (f) as a stipulation is not strictly speaking a necessary condition of knowing English. One may know English without understanding, or even encountering, tokenings of 'fish'. So a rule of English is not simply something one must accept as a stipulation to count as knowing English. This is an aspect of the fact that knowing a language is a matter of degree.