

Daniel Dohrn
Lewis and his Critics on Putnam's Paradox

Philosophie
Universität Konstanz D 22
D-78457 Konstanz
Germany
e-mail: Daniel_Dohrn@yahoo.com
Tel. 0049(0)7531882727

Lewis and his Critics on Putnam's Paradox

The model-theoretic argument known as Putnam's paradox threatens our notion of truth with triviality: Almost any world can satisfy almost any theory. Formal argument and intuition are at odds. David Lewis devised a solution according to which the very structure of the world fixes how it is to be divided into elite classes which determine the reference of any true theory. Three claims are defended: Firstly, Lewis' proposal must be completed by an account of successful referential intentions. Secondly, contrary to Catherine Elgin's criticism of Lewis, natural properties corresponding to elite classes may play a role in sound scientific inquiry. Thirdly, despite Bas van Fraassen's objection that the sceptic cannot consistently maintain doubts about reference, there is a promising sceptical strategy of exploiting Putnam's results which is answered by Lewis' account.

Keywords: Reference, Indeterminacy of Reference, Putnam's paradox, Elitism

1. Putnam's Problem and Lewis' Solution

One of the most spectacular results of the application of formal methods to basic philosophical questions is the model-theoretic argument known as Putnam's Paradox. Putnam develops this argument in order to question the possibility of a realist semantics. The controversy among different philosophers about this reasoning provides an interesting specimen of how to deal with apparent conflicts between formal argument and philosophical intuitions. The original argument goes as follows:

"Let T be an ideal theory, by our lights [...] Now T is consistent (by hypothesis) and has (only) infinite models. So by the completeness theorem [...] T has a model M of the same cardinality as THE WORLD. Map the individuals of M one-to-one into the pieces of THE WORLD, and use the mapping to define relations of M directly in THE WORLD. The result is a satisfaction relation SAT – a "correspondence" between the terms of L (that is, the first-order language in which T is expressed) and the sets of pieces of THE WORLD – such that the theory T comes out *true*.....provided we just interpret 'true' as TRUE(SAT)." (Putnam 1978, 125-126)

Putnam further argues that L may be our natural language so that his argument applies to reference in the latter. The result is that any empirically ideal theory is true. He concludes that metaphysical realism is untenable as it presupposes that even an ideal theory may be false. As Lewis points out, the argument can be generalized: An arbitrary consistent system of sentences can be mapped into a range of objects so that it is true provided the range of objects is sufficiently large. "[...](almost) any world can satisfy (almost) any theory." (Lewis 1984, 229) An even more puzzling result of this argument is the following: Almost any theory can be mapped in more than one way into a range of objects so that it comes out true, provided the range of available objects is sufficiently large.¹ The problem of imposing further constraints on admissible interpretations is that such constraints are just "more theory" (Lewis 1984, 226; Elgin 1995, 290). They can only consist of further sentences which pose the same problem as the original system of sentences. This holds especially for our referential intentions. Putnam hopes to avoid the resulting difficulties by sacrificing metaphysical realism. There is no truth transcending ideal assertibility.

Now we seem to be faced not with a harmonious interplay but with a clash of intuition and formal model-theoretic arguments. The latter threatens the intuitive "[...]realist philosophy we know and love" (Lewis 1984, 221). Common-sense realism yields commonplaces like the following: "There is one and only one objective truth of the matter to be established. Scientific theories may contain sentences which are either true or false. Truth is not easy to attain. A theory about x, e.g. neutrinos, is first and foremost a theory about x and only about x

(although we may realize that it can be used as a model for theories about other objects as well).” I want to discuss strategies pursued by different philosophers to cope with Putnam’s results. I focus on David Lewis’ proposal and reactions to it. Putnam and Catherine Elgin propose to accept as a result of the model-theoretic argument that metaphysical realism is refuted. They embed this view into a pragmatist view of scientific activity which they claim to be intuitively more appealing than realism. Contrarily, Lewis proposes a solution how to reconcile the model-theoretic argument with realism. Elgin criticizes that Lewis’ result is unintuitive in the light of pragmatic arguments. In contrast to all these strategies, Baas van Fraassen denies that the model-theoretic argument applies to our natural language.

I now want to turn to Lewis’ proposal how to deal with Putnam’s paradox. I begin with some important points in which Lewis and Putnam agree. Even if a theory is enriched by as many theoretical constraints as one might wish, its interpretations are not sufficiently delimited. Like Putnam Lewis denies that causal constraints serve to fix reference, but only because he adopts a “causal descriptivism”(Lewis 1984, 227): Causal relationships fix the reference of an expression only by virtue of a description which specifies how this expression and its referent are causally linked. Causal descriptivism does better than pure causal theory to accommodate cases in which the reference of an expression is fixed by a description rather than by a causal chain. Lewis also points out that a pure causal theory is unable to cope with everyday words like “in” in sentences like “I am in Vienna” (Lewis 1984, 235).

Now let me point to some important differences between Lewis’ and Putnam’s accounts. Against the “more theory”- argument Lewis distinguishes constraints functioning by virtue of descriptions and conditions which just factually constrain reference. Whereas the first ones are added to the original theory and need an interpretation themselves, the function of the latter is not mediated by a description (Lewis 1984, 225). Descriptions of causal relations an eligible referent must *satisfy* may be contrasted to real causal relations which must only *hold* between an expression and its referent. The latter might serve to determine reference without being in need of an interpretation themselves. But as already indicated Lewis does not accept pure causal theory.

Lewis doubts that Putnam’s problem can be solved by sacrificing metaphysical realism. He accepts Putnam’s challenge as he regards Putnam’s paradoxical results to be unavoidable unless a special condition is met. The problem and Lewis’ solution can be compared to another sceptical problem about meaning which was developed at the very same time as Putnam’s paradox: Kripke’s rule-following considerations (Kripke 1982). Kripke’s sceptic challenges her opponent to name the mental fact determining an infinity of correct applications as allegedly commanded by a rule. Since all candidates Kripke considers fail, he concludes that here is nothing determining the correct applications of a rule. Authors like Crispin Wright and Paul Boghossian argue that Kripke does not attend to a relevant kind of mental facts: intentions (Boghossian 1989, 542; Wright 2001). Lewis makes quite a similar move: He enlarges the range of reference-determining factors compared to the range considered by Putnam. If reference is not fixed by our theories nor anything else of our doing, another relatum of the reference relationship remains: The world may fix the reference of our expressions on our behalf. Lewis postulates privileged natural things and corresponding elite classes which are to fix the references of general expressions:

“Only an elite minority are carved at the joints, so that their boundaries are established by objective sameness and difference in nature. Only these elite things and classes are eligible to serve as referents.”(Lewis 1984, 227)

Lewis proposes physical objects as eligible candidates for elite classes. But not all theories are physics. What about other theories and their truth claims? Lewis argues:

“Ceteris paribus, an eligible interpretation is one that maximises the eligibility of referents overall. Yet it may assign some fairly poor referents if there is good reason to. [...] overall eligibility of referents is a matter of degree, making total theory come true is a matter of degree, the two desiderata trade off. The correct, ‘intended’ interpretations are those that strike the best balance.”(Lewis 1984, 227-228)

Lewis contends that less eligible qualities like “poisonous” are more or less directly definable by perfectly natural ones:

“The less elite are so because they are connected to the most elite by chains of definability.”(Lewis 1984, 228)

A trade-off must take place between two principles when interpreting a theory:

(P1) Maximize the *eligibility* of referents assigned to the theory.

(P2) Assign referents to the theory so as to maximize its *true* sentences.

This balance of principles allows to account for factors of theory choice which differ from a quest for the ultimate physical theory.

Now I want to briefly discuss what I consider a grievous problem of Lewis’ proposal before coming to critical positions in literature. Assume anthropologist Smith wants to provide a theory about a barter economy in New Guinea. His theory is quite a poor one, but by pure accident it would do much better in terms of eligibility and truth if it were interpreted as a partial description of classical mechanics. I would still be inclined to interpret it as a poor theory of a barter economy and not as a more successful partial version of classical mechanics of which Smith does not have any idea. Of course, once some physicist realizes that upon suitable reinterpretation of its referring expressions, the theory would do excellently as a version of classical mechanics, it is legitimate to use it as such a version. But such a usage presupposes reinterpretation. As philosophers of science we might feel inclined to adopt a notion of theory which implies that the theory remains the same upon such reinterpretation. But I would claim that our everyday life and our scientific practice involve another notion of theory. Normally, economists do not attend to suitability for classical mechanics, astronomers do not care for superconductors and so on. This is reflected in the ways they act, collect evidence, design their experiments, organize symposia etc.. My point does not depend on the notion of theory. It merely requires that we sometimes -like Smith- can direct our theories to objects of attention. This seems crucial for our practice. Smith may desperately need his theory of a barter economy in order to avoid offending the indigene tribes he studies and being killed. Thus there must be some “inner” constraint on reference by what people take themselves to refer. This inner constraint is not adequately captured by Lewis’ principles of interpretation.

Now the problem to account for such inner constraints is that they apparently are just “more theory” and thus in need of an interpretation themselves. Yet I want to suggest that they nevertheless may place real constraints on an interpretation. A way to illustrate this is given by the chains of definability linking elite and less elite qualities. If Smith possessed the ultimate physical theory, he could define some of the less elite referents he intends his economical theory to relate to. I claim that even if Smith’s theory would still do better if interpreted otherwise, we should respect these definitions of his. Thus chains of definitions and principles of interpretation may conflict. However, Smith does not need to know the ultimate physical theory in order to successfully intend to talk about barter economies. How does he manage to do so? I can only hint at a way to deal with this problem. Smith could link his explicit referential intentions descriptively to the rest of his language and belief system. But then the more-theory problem looms. We must recur to Lewis’ distinction of a constraint functioning via a description and a constraint functioning without a description. At least some

inner constraint must function in the latter way. There must be real relations which place some constraint on interpretation without so far being in need of an interpretation themselves. Such relations could be based e.g. on perceptual acquaintance and indexical devices.² In order to fix the reference of his theory, Smith or someone of his community could go to New Guinea and point to the transactions his theory is to deal with. The fine-tuning of such a deictic act surely depends on interpretation, too. But such an act may be claimed to have some rough and rudimentary reference-delimiting function independently of interpretation. If I am right, “internal” reference-determining devices like deictic acts or acquaintance must be added to Lewis’ two principles:

(P3) Pay due respect to “internal” reference-determining devices.

Now one may ask how this principle can be reconciled with Lewis’ principles. For according to the former, reference is determined “from within” by what is accessible to the speaker’s intentions. According to Lewis’ principles, reference is determined externally by natural facts which do not have to be accessible to the speaker. To put it metaphorically: “Internal” devices are anchors people throw in order to hook into reality, whereas eligibility is an anchor nature throws in order to hook into people’s language.³

The result of discussing Lewis’ proposal is that it must be completed by an account of how referential intentions serve to constrain reference.

There is another problem: Elitism is the only way of determining reference Lewis provides. Thus we may expect elite properties to uniquely determine reference. Now Putnam’s problem looms again in a slightly different guise. Although Lewis’ proposal certainly serves to restrict reference, it cannot be guaranteed to uniquely determine it. If the total range of perfectly natural referents is sufficiently large and structured in a certain way, a theory may be mapped in more than one way to eligible referents. Thus one can never be sure whether a theory’s reference is determinate. Again some further constraint on reference may be required. Again indexical mechanisms may supplement eligibility theory although they cannot guarantee either that reference is determined uniquely. Now if eligibility theory proves insufficient to guarantee determinacy of reference, why don’t we give up eligibility theory altogether and look for another real constraint on reference? Such a doubt can be answered by recurring to Lewis’ arguments that alternatives like a pure causal theory do not fare better than eligibility theory. Thus a certain indeterminacy of reference could prove resilient. Grooming requirements of eligibility and overall truth may contribute to indeterminacy.

2. Lewis Critics I –Elgin and Eligibility

I now want to deal extensively with two of Lewis’ critics. Catherine Elgin’s criticism serves to better understand the epistemological and metaphysical requirements of elitism. Elgin denies that privileged qualities can actually be the target of our theorizing. Even granted they exist, they cannot be the subject of methodically guided epistemic activity:

“[...] the factors that distinguish eligible from ineligible referents must be independent of us [...]. But if they are genuinely independent, it is hard to see how they can do any epistemological work. Have we any reason to believe that the considerations that figure in theory choice or assessment key into natural properties?”(Elgin 1995, 294)

If natural qualities are to play an epistemological role, one must have access to them as such. A criterion must be provided how to single out elite classes. Elgin considers a theoretical virtue, simplicity. But there must be some measure of simplicity. In order to provide a criterion, simplicity must not be relativized to capacities, interests etc.. Rather it would have

to be measured against elite qualities which would have to be known before (Elgin 1995, 294). Furthermore, it is by no means sure that nature favours simplicity of theories. Theory formation must be guided by scientific virtues. Elgin's result is that there is no epistemic relationship between scientific virtues and elite qualities. Thus our best theories may totally miss eligible qualities:

"Lewis has contrived a new form of skepticism. Even if we know that the world has the structure our science ascribes to it, we might still be wrong. For although that structure is a genuine structure, it might still be the wrong structure."(Elgin 1995, 299)

Ideal theories may miss the natural structure of the world if the latter proves to be worse than ideal.

What makes eligible properties eligible?⁴ Elgin doubts that this question can be answered, that it can be explained "[...] how the world privileges properties" (Elgin 1995, 291). What constitutes a natural property? Elgin considers a supervenience relation between natural and non-natural properties. But supervenience does not serve to privilege one of the relata:

"Elite properties likewise supervene on a suitable array of plebeian properties." (Elgin 1995, 297)

Elgin further asks why reference should be determined by elite properties at all:

"What is so impressive about perfectly natural properties that we should restrict reference and truth to them and their kin?"(Elgin 1995, 297)

Elgin accepts as Putnam's result that truth is cheap (Elgin 1995, 300). She proposes instead to look for truths which are *interesting* in light of further epistemological and scientific desiderata. Cognitive activity must aim at "[...] an understanding of a certain kind [...]" (Elgin 1995, 301). There is no overarching ultimate aim of scientific inquiry.

I now want to discuss Lewis' resources to meet Elgin's objections. I do not aim at refuting the latter but rather want to indicate that there might be ways of evading them. Thus I grant me the liberty of presenting somewhat vague and tentative suggestions.

Firstly the metaphysical question what privileges elite properties must be addressed. Supervenience is no answer. So it may be most appropriate to regard eligibility as a basic metaphysical category which cannot be further explained by relations like supervenience.

Let us now turn to Elgin's epistemological concerns. Lewis himself grants that scientific theories do not always or even commonly aim at "carving nature at the joints". The range of objects which successful theories must refer to is the result of balancing eligibility and other principles. This trade-off does not need to be a conscious achievement of the theorist. But the intriguing concern remains that since one does not know a criterion of eligibility, one cannot consciously aim at a theory yielding perfectly eligible properties. I want to follow a twofold strategy to deal with this concern. Firstly I want to argue that elite properties may have an indirect pragmatic role in scientific inquiry. Secondly I want to propose that pace Putnam observation may provide a relationship between theories and eligible properties.

Even if there is no direct criterion which allows to conclusively establish eligibility, the latter may have pragmatic significance. We can try to specify what a scientific practice which is guided by the relationship between elite properties and their less elite descendants should look like. The resulting picture of scientific inquiry diverges substantively from Elgin's. However, if it adequately captures our scientific practice, it may be contended that the quest for elite properties does play a significant and rational role in theory choice and assessment.

Elgin draws quite an anarchic, decentralist picture of cognitive activity reaching from natural science to the "[...] cognitive contributions of art [...]"(Elgin 1995, 302). Any discipline sets

its agenda autonomously and independently of other disciplines. This decentralist picture is not in tune with Lewisian elitism. The latter demands a more hierarchic picture of cognitive activity. If there is to be a fundamental discipline which aims at capturing perfectly natural qualities, its aspirations must be inwardly and outwardly hegemonial. Inwardly: It must try to represent a unified system of objects and classes suited to form the full store of the world from which any other objects and classes may be derived by chains of definability. Outwardly: Other scientific disciplines may set their own agenda, pursue their own interests and patterns of salience, develop their own classificatory schemes. But the hegemonial claim of the fundamental science involves that the reference classes devised by other disciplines must be reducible to it in a suitable manner –at least in principle. Such a picture of scientific inquiry follows from eligibility theory. But in turn its hierarchic structure, the hegemonial claims of the fundamental science may depend on some commitment to elitism. If the relationship between this picture and eligibility theory is sufficiently close, there may well be criteria of theory choice and theory assessment which are best interpreted as aiming at “carving nature at the joints”.

In my opinion this hierarchic picture describes our scientific practice better than Elgin’s “Laissez-faire”. Physics seems to be the hegemonial science hinted at. It is more comprehensive and more basic than any other natural science. Take its relationship to biological theories. Surely most scientists are convinced that all biological procedures are based on microphysical procedures but not vice versa. The latter are often used to explain the former but not the other way round. If a biological theory contained statements being in tension with the statements of physics, we would rather try to revise them than the physical ones. The same holds for psychological theories. Elgin points to Fodor’s idea that any theory should set its own agenda (Elgin 1995, 301). But Fodor also wants his cognitive science to be in tune with physical statements about the causal structure of the brain and the surrounding world (cf. Fodor 1987). He proposes that properties devised by other sciences might be defined by disjunctions of physical properties (cf. Saporiti 1996, 26-31; Fodor 1975, 21). A comparable relationship may obtain between elite and other properties. These examples indicate that we ascribe to physical theories a standing which corresponds to what we would expect if Lewis’ contention were true: Non-physical statements should pay allegiance to the statements of physics. Perhaps ascriptions of intentionality are most difficult to reconcile with the supremacy of physical statements. But the very discussion how to reconcile statements about intentionality with physical ones shows the hegemonial claims of the latter. In turn these hegemonial claims may presuppose a commitment to physical objects being naturally eligible. Of course these considerations are far from conclusively establishing that our scientific practice depends on a commitment to elitism. But they show that there are resources to argue for such a dependence.

Although this more hierarchic picture of science may better fit the hierarchy of conditions of reference Lewis has in mind than Elgin’s, there remains a basic concern as to whether our theorizing is in accordance with the requirements of aiming methodically at successfully referring theories as Lewis envisages them. We usually do not attend to the definability of our non-physical concepts by physical ones which might be inscrutable for us. How can we reasonably neglect the mechanisms of reference in the way we do? A principal response to this concern could be that we at least implicitly are oriented towards a suitable relationship between physical theory and statements in other fields. Even when we do not aim at representing perfectly natural properties and do not know how the referents of our theories can be determined by their relationship to more eligible ones, we value the accordance of our theorizing and a possible basic physical theory in terms of which all concepts are definable. The pragmatic criteria of such an agreement may be given by relations of hierarchy as those indicated above.

Still there is Elgin’s objection that there is no direct criterion to establish eligibility. Even if the hierarchic system of our theories has all theoretical virtues, it may still be wrong. It seems

to “hang in the air” without keying into natural properties. It is somewhat difficult to see how we could –as Elgin puts it- know the world to have a structure albeit its being the wrong structure if Lewis’ two principles (P1, P2) are applied. To be sure, it is just the definiens of metaphysical realism that our best theories might fail to be true. But this does not mean that ultimate truth cannot play a role in a rational cognitive activity. It may nevertheless be reasonable to assume that these theories capture the world as it is. Lewis’ principles of *maximizing* truth and eligibility place restrictions on the possibility of an ideal theory rendering the wrong structure even if they are supplemented by “inner constraints” as I proposed. Thus a mechanism built into Lewis’ reference-guiding principles ensures that our theories tend to capture eligible properties albeit we do not have a direct criterion of eligibility.

Philosophers like Quine emphasize that theorizing consists in balancing observational results with independent maxims like simplicity or conservativity (Aldrich 1955, 18-19). Thus Elgin hints at a real problem: How can we assume that the world is structured so as to support virtues like conservativity and simplicity? If we mainly aim at as accurate a picture of the world as possible, how can such virtues play a role? Yet note that this difficulty does only directly pertain to the basic physical theory, not to theories setting other priorities. More importantly, even when aiming at the ultimate theory, we must take into account the limits of our cognitive capacities. Presumably virtues like simplicity and conservativity have to do with these limits and not with supposing that simple and conservative theories better fit the world than less simple and less conservative ones. They may be interpreted as rational principles of balancing the quest for the ultimate theory with constraints owed to limited resources. I do not want to deny that Lewis’ two principles (P1, P2) might prove insufficient to fully discard Elgin’s doubts. But they may be supplemented by additional pragmatic principles without being substantially modified.

Finally I want to address Elgin’s own solution of Putnam’s problem. Epistemological constraints do not yield determinacy of reference. Consequently Elgin admits that truth is cheap. But this is exactly what appears unacceptable to Lewis. Van Fraassen’s argument that Elgin’s concession leads to inconsistency will be discussed below. Independently of such worries, we may illustrate the difficulties of this position on a more intuitive level: For Elgin scientific activity is guided by practical concerns like utility. Assume Smith is faced with a cobra and considers it most useful to know whether it is poisonous. What matters to him is not an elegant or virtuous theory, but the bare truth of the matter. But then truth cannot be cheap. Assume Smith endorses a theory involving the sentence “cobras are poisonous”. Assume further that due to the indeterminacy of reference it may as well be possible to interpret this sentence in the sense of cobras being poisonous as in the sense of, say, rabbits being edible or in infinitely many other ways. Nothing fixes that it refers to cobras even if this proves in fact to be the most useful interpretation. Now Elgin could contend that her notion of interesting truth is apt to handle this case. But in order to do so, it must answer the question how it serves to restrict reference so that only the interpretation yielding cobras being poisonous can be the correct one. Elgin’s requirements which are to specify “interesting” are just “more theory”. If truth is cheap, so are they. Assume the constraint of maximizing utility is introduced to cope with the cobra example. Yet as Elgin grants, the whole theory including “useful” can be interpreted in many ways. “Useful” could as well refer to the property of being, say, humorous. Elgin could react by pointing to Lewis’ distinction of a constraint which is just “more theory” and constraints which factually restrict reference. However, this amounts to revising her commitment to indeterminacy of reference. Elgin could adopt a radical pragmatic stance. She could insist that the best way to handle Smith’s theory is to avoid cobras. Successful activity is the hallmark of interesting truth. But if such a radical pragmatism is adopted, it is completely open how theory-building activities could be guided by pragmatic requirements of successful activity. Thus Elgin does not explain how reference can be

restricted in the way necessary to cope with the cobra example. I would suggest that in addition to Lewis' solution an account of successful referential intentions as sketched above is necessary.

The result of discussing Elgin's criticism is that although she is right that epistemological concerns must be dealt with, she does not show that Lewis is unable to do so. Her commitment to truth being cheap is untenable for both formal and intuitive pragmatic reasons.

Elgin shares with Putnam the basic conviction that the prima-facie plausibility of metaphysical realism and Lewis' defence of it depend on a mistaken philosophical picture of scientific inquiry mirroring the world rather than pragmatically handling it. She accepts that Putnam's formal argument applies to our natural language and uses it as a means of establishing a pragmatist position which allegedly suits better our natural and unspoiled intuitions than metaphysical realism. Formal methods are used to purify our sense of intuitive plausibility. In doing so, Elgin underestimates the pragmatic importance of reference as captured by intuitive realism. Furthermore, since she focuses on pragmatics, she does not see the formal contradiction of applying Putnam's result to our natural language in the way she does.

3. Lewis' Critics II – van Fraassen on Truth and Interpretation

Van Fraassen doubts the formal model-theoretic argument to apply to our normal language. He rejects the premise accepted by Lewis, Putnam and Elgin that truth is a matter of interpretation. In order to defend their claims, van Fraassen's critique that Putnam's problem cannot even be properly formulated without courting contradiction must be met.

I want to begin with briefly recapitulating van Fraassen's argument that one cannot accept, as Elgin does, indeterminacy of reference without contradiction. He looks at an attempt to define the truth of a theory T for a language L in the sense of "truth under some interpretation":

"(EQ) T is true if and only if T is true under an interpretation that assigns parts of the world as extensions of the terms in L." (van Fraassen 1997, 86)

If "truth" is defined by "truth under an interpretation", theories standing in formal contradiction to each other can be true. A formal contradiction results. Now interpretations could be restricted to admissible ones:

"(EQ*) T is true if and only if T is true under an *admissible* interpretation that assigns parts of the world as extensions of the terms in L." (van Fraassen 1997, 88)

This definition also leads to a contradiction if it is granted that L can be interpreted in more than one way:

"If the language has two genuinely differing interpretations, there must exist two distinct complete theories. Given (EQ*), they are both true. But being distinct and complete, they stand in formal contradiction to each other [...] The reductio is avoided only if correct interpretation is unique." (van Fraassen 1997, 89)⁵

This argument may refute Elgin's commitment to truth being cheap. However, van Fraassen admits that a solution which does allow only one interpretation of a complete theory can avoid this result (van Fraassen 1997, 89 ann.). Lewis' solution may meet this condition as he considers the truth of a complete theory to be uniquely fixed by elite properties.⁶

Van Fraassen directs the rest of his article against the very idea of our own language being in need of an interpretation. The restrictions of reference placed by Tarski's satisfaction scheme are hurt when it is contended that a language may be interpreted in several ways (van

Fraassen 1997, 90). A pragmatic contradiction arises when the ability to refer to something in one's own language is exerted to doubt this very ability. However, this problem is already noted by Lewis (Lewis 1984, 221). The sceptic must play two roles: In order to raise a doubt about reference in her own language, she must at the same time use expressions of this language with determinate reference and pretend that their reference is indeterminate:

"To raise a doubt about whether white is a word for white things, or whether 'Snow is white' if and only if snow is white, *in these very words*, requires two things. To do that, on the one hand, she [the sceptic] must use the words 'white' and 'snow' to talk about white things and snow; on the other, she must treat the quoted expressions as if they belonged to a foreign language."(van Fraassen 1997, 92)

The sceptic commits a pragmatic contradiction when she denies the possibility of what she is thereby doing: successfully referring.

I do not think that van Fraassen wants to argue against Lewis' eligibility theory. He merely wants to question that the sceptical challenge involved in Putnam's paradox must be answered at all. Even if we are caught within the confines of our own language when discussing its reference, this does not contradict the thesis that reference is determined by the principles (P1-P2).

Despite its elegance and simplicity, I consider van Fraassen's argument unintuitive. Although we cannot live with scepticism, van Fraassen's *argumentum ad hominem* seems unduly to restrict our capacities of a self-distancing sceptical reasoning and to undercut the valuable insights it gives rise to like those of Putnam, Lewis, and Elgin. In order to discuss the validity of this argument against the sceptic, we must first look what restrictions it really places on the sceptic's reasoning. Surely sceptical claims like the following are excluded:

(S1) It is indeterminate what the expressions of our own language refer to.

(S2) It is doubtful that "white" does refer to white and not to non-white things.

But what about the following sceptical consideration? Take an arbitrary consistent theory in an artificial language. Putnam's model-theoretic argument shows that this theory under certain circumstances has more than one interpretation making it come true. In so far this theory has no determinate reference. Pointing to the analogy to our own language and our own theories, we can ask sceptical questions like the following:

(S3) How can a theory have determinate reference?

(S4) What fixes the reference of *our* theories in *our* language so that they are referentially determinate?

The sceptic may hope that her opponent accepts her challenge and fails to explain how reference is fixed. But why should the opponent fail to do so if she must only point to the Tarskian satisfaction scheme? The sceptic could reject such an explanation as missing the point. Her request is not merely to state a general condition which must not be violated if successful reference is to take place, but rather to explain how successful reference is possible, how language can hook into a reality of infinitely many things. The sceptic about reference resembles Kripke's sceptic about rules who just asks the seemingly innocent question what mental fact can bear rules. The sceptic about rules courts pragmatic contradiction, too, if she maintains that one cannot follow rules and in doing so follows linguistic rules. But as long as she points to the difficulty of naming rule-bearing facts, she remains perfectly consistent. The same holds for the sceptic about reference when she only asks how we manage to refer. Now what if the sceptic's opponent simply declines to answer the challenge? One cannot be blamed for using one's own language as any critical assessment must start "in the midst of content" (McDowell 1987, 74). But this does not necessarily mean

that one is justified in doing so. In any case it seems legitimate and valuable to try to answer the sceptic not merely *ad hominem* (by pointing to the pragmatic contradiction involved in questioning reference of one's own language within that language), but by providing an understanding of how reference works. However, can scepticism have any pragmatic significance if we cannot but use our normal language? There remains a desperate measure for a sceptic convinced by Putnam's argument: He could simply fall silent for good.

Finally I want to discuss van Fraassen's resources to answer this moderate sceptical challenge. I am not sure whether he actually does so but he could endorse a minimalist thesis about reference: There is nothing to be said about reference than what is expressed in satisfaction schemes like the following one:

(R) 'Green' refers to green things.

This position corresponds to minimalist accounts of truth according to which there is nothing more about truth than what is expressed by the T-sentences of a language. Crispin Wright resumes this position as follows:

"[...] the Disquotational Schema

(DS) 'P' is T if and only if P

is (all but) a complete explanation of the truth predicate [...] truth is not a 'substantial property', whatever that means, of sentences, thoughts, and so on, but merely a device for accomplishing at the metalinguistic level what can be accomplished by an assertoric use of the mentioned sentence." (Wright 1992, 15f.)

On this basis a minimalist account of reference could be developed. Talk of reference is only a device of accomplishing at the metalinguistic level what is accomplished by using referring expressions.

Such a minimalist theory of reference provides an answer to the sceptic: Her asking for more explanation misses the point as it unduly inflates reference. Minimalism could be the real alternative to the efforts spent by Putnam, Lewis, and Elgin. We may wonder whether one can content oneself with minimalism or whether there is some real achievement involved in successful reference which calls for further philosophical analysis. This grand issue cannot be discussed here. But the following can be maintained: If a controversial commitment to minimalism must be added to meet moderate scepticism, pace van Fraassen the model-theoretic argument allows to raise sceptical worries about reference in our natural language.

In conclusion, although van Fraassen presumably refutes Elgin's thesis that truth is cheap and an immoderate sceptic who preposterously denies that reference is possible, he does not succeed in discarding a moderate sceptical position and with it an issue Lewis wants to settle by his eligibility-theory. While accepting the formal validity of the model-theoretic argument, van Fraassen denies that it can be used to establish scepticism about reference. In doing so he underestimates the resources of Putnam's argument. Although it does not directly apply to reference in our natural language, the model-theoretic argument can serve as an indirect dialectical tool to question and to scrutinize demands of reference.

References

Aldrich, V. (1955). Mr. Quine on Meaning, Naming and Purporting to Name. *Philosophical Studies* (6), 17-26

Boghossian, P. (1989). The Rule-Following Considerations. *Mind* (98, 392), 507-549

- Boghossian, P. (1994). The Transparency of Mental Content. *Philosophical Perspectives* 8, 33-50
- Bonjour, L. (1998). *In Defense of Pure Reason*. Cambridge: Cambridge University Press
- Chambers, T. (2000). A Quick Reply to Putnam's Paradox. *Mind* (109, 434), 195-197
- Dummett, M. (1973). *Frege. Philosophy of Language*. London: Duckworth
- Elgin, C. (1995). Unnatural Science. *Journal of Philosophy* (92, 6), 289-302
- Fodor, J. (1975). *The Language of Thought*. Cambridge, Mass.: Harvard University Press
- Fodor, J. (1987). *Psychosemantics*. Cambridge, Mass.: MIT Press
- Kripke, S. (1982). *Wittgenstein on Rules and Private Language*. Oxford: Blackwell
- Lewis, D. (1972). How to Define Theoretical Terms. *Journal of Philosophy* (67, 13), 427-446
- Lewis, D. (1984). Putnam's Paradox. *Australasian Journal of Philosophy* (62, 3), 221-236
- McDowell, J. (1987). In Defence of Modesty. In B. Taylor (Ed.), *Essays on the Philosophy of Michael Dummett* (pp. 59-80). Amsterdam: Nijhoff
- Putnam, H. (1978). *Meaning and the Moral Sciences*. New York: Routledge
- Putnam, H. (1981). *Reason, Truth and History*. Cambridge: Cambridge University Press
- Saporiti, K. (1997). *Vergleich einer repräsentationalistischen und einer syntaktischen Theorie des Geistes*. Berlin: de Gruyter
- Searle, J. (1983). *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press
- van Fraassen, B. (1997). Elgin on Lewis's Putnam's Paradox. *Journal of Philosophy*, (94,2), 85-93
- Wright, C. (1992). *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press
- Wright, C. (2001). *Rails to Infinity. Essays on Themes from Wittgenstein's Philosophical Investigations*. Cambridge, Mass.: Harvard University Press

¹ In Putnam (1981) this consequence plays a more important role than in Putnam (1978).

² Searle suggests that perceptual states have a reflexive indexical structure (Searle 1983, 57-78). Their satisfaction conditions involve causal chains linking them to their reference objects. Such a structure could serve as a constraint on reference which does not need an interpretation.

³ There is an interesting parallel to internalist and externalist theories of meaning. According to the former, content is determined "from within", meanings are "in the head" (cf. Searle 1983). According to the latter, content is determined by external facts, e.g. by causal relationships (cf. Fodor 1987). I would suggest that there is a middle way between content internalism and content externalism which allows to reconcile Lewis' principles (P1-P2) with the intention-principle (P3). Moreover, it allows to support Lewis' contentious proposal by almost universally held externalist opinions and vice versa. I agree with Lewis in rejecting the idea that meaning can be

fully determined by external causal relations. My main reason is a weak accessibility claim: Entertaining contentful thoughts necessarily requires an awareness of content (cf. Bonjour 1998, 164; Boghossian 1994). Probably referential intentions determine external meaning by allowing natural structures to hook into it: I (or my community) have the intention that my (or our) word “water” refer to whatever is structurally identical to a specimen indexically referred to. By virtue of this intention, I give nature the opportunity to determine the extension of “water” by the most eligible structural properties underlying a specimen. Meaning externalism could even require elitism. Our claim that water is identical with H₂O and not with, say, an odorless liquid, is a matter of how to define structural identity. This issue is settled by eligibility. If the extension of “water” is determined by its being H₂O, this scientific expression must yield the most eligible interpretation of “water”. Now what about the holism implicit in Lewis’ demand to balance overall eligibility and truth of a theory? Apparently I (or we) do not attend to maximizing overall eligibility and truth of a theory when indexically defining water. But I see no problem in interpreting my (or our) intention as follows: “Water” is to refer to whatever is structurally identical to an indexically identified specimen. Structural identity is to be understood in terms of maximizing overall eligibility and truth of an optimal physical theory relating to water. Thus the intention-principle (P3) can not only be reconciled with Lewis’ principles (P1-P2), but serves to sustain them and vice versa.

⁴ Elgin may also have another question in mind: Given elite properties exist, why should we try to refer to them and only to them and their kin? Lewis’ answer to this question is obvious: There is nothing to be referred to except such properties and their kin.

⁵ Another *reductio* is provided by Chambers (2000) who tries to show that Putnam is committed to the claim that any ideal theory is *necessarily* true.

⁶ However, since Lewis admits that statements about less eligible things may be true, the threat of contradiction may occur again. Nothing seems to exclude that a theory containing “cobras are poisonous” may as well be true as another theory containing “cobras are not poisonous” if interpreted according to Lewis’ two principles of balance.