

Consuelo Preti

The Irrelevance of Supervenience

abstract

Externalistic theses about the nature of content appear to have untoward consequences for the role of content in folk psychological causal explanation. Given the predominance of Twin Earth formulations of externalism, however, the explanatory role problem is often construed as the problem of content's failure to supervene on intrinsic states of the head.

In this paper I argue that this is misleading. The most general formulation of externalism is shown to be independent of any supervenience claims. The result is that all concepts can be shown to be both externalist in individuation and irrelevant in causal explanation, whether or not they fail to supervene on intrinsic states.

Part I

1. Introduction

Strong words have been written concerning the feasibility of abandoning talk of content in psychology (Fodor 1988, p.xii):

... if commonsense intentional psychology really were to collapse, that would be beyond comparison the greatest intellectual catastrophe in the history of our species; if we're wrong about the mind, then that's the wrongest we've ever been about anything . . .

Strong words notwithstanding, it seems that the explanatory role of content-citing mental descriptions is undermined by externalist theories of content. Externalism claims that content is necessarily individuated by way of objects and properties located in the world outside the head of a psychological subject. But this makes it difficult, on the face of it, to see how content-citing attitude attributions could have a legitimately causal role in psychological explanation. Attitude attributions that involve essential reference to entities outside the intrinsic causal mechanisms that presumably link thought to action raise a causal-explanatory mystery.

Now, it is standardly claimed that content becomes problematic in propositional attitude psychology because it fails to supervene on intrinsic states of the head; complementarily, if content does supervene, there is no explanatory role problem. In this paper, however, I argue that this conception of the explanatory role problem is in need of revision.

Most in need of reconsideration is the widespread tendency to take Twin Earth cases as definitive (even necessary) for both the formulation of externalism and the understanding of the causal/explanatory problem engendered by externalist theories of content. Twin Earth cases are canonically understood as examples of the failure of supervenience of mental states on intrinsic states of the head; and it is failure of supervenience that is taken to be the defining feature of externalist theories of content, as well as the critical difficulty for belief-desire psychology.

But, as we shall show, few concepts are in fact subject to Twin Earth cases — so few kinds of content in fact fail to supervene. But if so few concepts really fail to sustain a viable role in the explanatory claims of propositional attitude psychology, we might be encouraged to believe that the threat of externalism to intentional psychology is too limited to be of any real worry.

However, we will show instead that there is a very general threat to the causal/explanatory claims of intentional psychology that results from externalism. In part I of this paper I argue that there is more to externalism than is implied by Twin Earth cases, so there is more to the problem of content in explanation than its failure to supervene.¹ I go on to show in part II that in fact supervenience is irrelevant to the problem of content in causal explanation.

2. Externalism and Twin Earth

Putnam (1975) originally employed Twin Earth cases as a device in an argument to support the claim that the meaning of a term cannot be wholly determined by internal properties alone. Surrounding a psychological subject are objects and the properties those objects instantiate; these are what she refers to by her words, what her thoughts are about. Putnam posed this question: is it consistent to suppose that two subjects who share all the same internal properties share the same conceptual ones? The answer — illustrated by the now-familiar Twin Earth cases — was no. The relations twins bear to their head-independent environment necessarily contributes to the determination of the meaning of their words. The implications of Putnam's thesis for a theory of mind was obvious; and externalism emerged on the scene.²

Now externalism is meant to be a general individuation thesis about mental content, one that claims that there is a genuinely constitutive relation between the mind and the extra-mental world. But Twin Earth cases simply do not support a general individuation thesis. Twin Earth cases — both those associated with Putnam and those associated with Burge — do not apply to all concepts.

Putnam-type Twin Earth cases are cases where internal facts are held constant and the extra-cranial environment is varied.³ But on closer scrutiny it is clear that such cases are not erectable for content involving, for example, mathematical concepts, concepts of mental states (like pain or pleasure), aesthetic concepts, or ethical concepts. The “environment,” in these cases, arguably consists of properties, abstract entities that can’t be varied in the same way as the natural kind environment that features familiarly in standard Twin Earth cases.⁴

Consider number concepts, the most obvious instance of non-conformity to a standard Twin Earth case. How is it possible to erect a situation where the mathematical environment is varied but the internal factors — what is presented phenomenologically, say — stay the same? It seems very peculiar to claim that the expression “prime number” or “two”, or, for that matter, the logical expression “if and only if,” refer to different things on Earth and Twin Earth, but that both communities have the same qualitative experiences with respect to those entities.

Remember that on Twin Earth, what looks like H₂O isn’t H₂O, it’s XYZ, even though it gives rise to the same qualitative experiences; it looks like water, but it isn’t water. There is, in fact, some basis for claiming that standard Twin Earth cases trade on an appearance/reality distinction, a distinction that it is far more difficult to draw with respect to mathematical reality.⁵

The same holds, analogously, for aesthetic concepts like “beauty.” Does it make any sense to claim that on Earth and Twin Earth the aesthetic environment is (really) different but *appears* the same to both sets of people, so that on Twin Earth the word “beautiful” refers to, say, *essentially* ugly things? Similarly, could the *moral* environment really be varied while appearing the same both to me and my twin, so that when she applauds someone’s integrity, she’s doing something I might in fact recoil from?

Putnam-type Twin Earth cases are examples that are meant to illustrate the fact that being identical in every local (non-relational) respect does not always guarantee sameness of mental content. Natural kind concepts vary with the natural kind environment, but such variation is difficult to apply with respect to the ethical, formal, or aesthetic environment. These considerations would appear to cast doubt on the claim that Putnam-type Twin Earth cases apply generally to all concepts. And the situation is no better with respect to Burge-type Twin Earth cases.⁶

Burge cases turn on the idea that concept-possession, in some cases, can be defective, so that variations in a community’s linguistic conventions influence concept individuation.⁷ Both I and my internally indistinguishable twin share a concept of “arthritis,” but not the same concept, given the standards of use

in our respective communities. But this seems to turn on the idea that there is something that arthritis is; something I only imperfectly grasp. So that when I say “I have arthritis in my thigh,” the standards of use in my community dictate that the concept I’m employing is defective: the word “arthritis” is meant to apply to a disease of the joints alone. In the Twin Earth community, by hypothesis, the linguistic conventions accord with my (defective) use of the word; they use “arthritis” to apply to a disease of the muscles and the joints. Burge’s point is that in spite of our internal similarity, my twin and I do not possess the same concept. Variation in the social environment results in a variation of the meaning of the word.

It is implausible, however, that all concepts are subject to defective possession in the way that appears essential to drive Burge-type Twin Earth cases. Consider phenomenological concepts like *experience of color*, or psychological concepts like *experience of pain*. My twin and I, internal duplicates, are in pain. Now suppose that on Earth, the community conventions apply the word “pain” to pain states. On Twin Earth, however, community conventions dictate that the word “pain” applies to states of *pleasure*. The same word, on Twin Earth, picks out a different experiential concept.

But this results in a peculiar situation. When my twin — my internal duplicate — judges that she is in pain, the community standards that allegedly impinge on concept-individuation entail that she is wrong about her own mental state. The community on Twin Earth uses “pain” to pick out states of pleasure — that’s what the word means on Twin Earth. But the experience she’s having is *pain*; since she’s my twin and that’s the experience I’m having.

Similarly, suppose that around me are others who unknown to me have experiences of green which they refer to by using the words “experience of red.” I, however, refer to my experiences of red by using the words “experience of red.” If the community environment dictates the individuation of my mental state, then, on this scenario, I am wrong about my own experience. When I judge that I am having an experience of red, that word, in my community, means experience of *green*. But it very implausible that I could be so wrong about my own mental states — and so, implausible that Burge cases can in fact be erected for all concepts.

Further, as McGinn (1989) points out, there is no logical entailment between Putnam- and Burge-type Twin Earth cases. From the fact that meaning and mental content is subject to variations in an individual’s social environment it does not follow that meaning isn’t ‘in the head.’ What individuates the meaning of an individual’s expressions could just as well be in the heads of those who comprise his social environment — a clear case of anti-individualism but internalism about meaning.

3. *The Formulation of Externalism*

The above considerations thus require that externalism be formulated more precisely. One formulation — sometimes referred to as *strong* externalism — can employ the familiar device of Twin Earth, and, accordingly, will entail a failure of supervenience of mental states on inner states. Content, according to this formulation, is partly determined by a causal or contextual relation between the subject and the world outside her head. From this it follows that the mental states of two internally indistinguishable subjects fail to supervene on those internal states when the subjects inhabit different environments.

The other formulation of externalism — sometimes referred to as *weak* externalism⁸ — involves no entailment of non-supervenience. This more general formulation is best thought of as a claim concerning the existence-dependence of mental states on their objects. On this view, it is a necessary condition on the possession of a thought that the subject be related to extracranial objects; the particulars or properties that are the reference of the expressions he employs. Should the requisite objects and properties fail to exist, or, counterfactually, were they to fail to exist, no thought can be possessed. Weak externalism holds for all content; it applies to content for which strong externalism is true and holds for content to which strong externalism does not apply.

The distinction in formulations of externalism thus expands the issue of the consequences of externalism for the role of content in psychological explanation. We must understand this issue now as dividing into separate but related questions. First: is there a necessary connection between externalism and supervenience? And second: do both weak and strong externalism have untoward consequences for content's causal/explanatory role?

4. *Externalism and Supervenience*

We can begin by considering the notion of supervenience more precisely, focusing in particular on its relation to the notion of externalism. As we have noted, there seems to be a prevailing view that there is an essential connection between formulations of externalism and failure of supervenience. But this is open to a reductio.

The Cartesian theory of mind has it that mental substance and physical substance are distinct; so much so that mental states can be what they are independent of there being bodies (heads with brains in them, say) at all. So on this view, it is perfectly conceivable that one's mental states may be what they are whether or not the objects of those states exist. Clearly this is a view

that entails a failure of supervenience of the mental on inner states; but it is a gross and unwarranted perversion of the Cartesian position to claim that from this failure of supervenience follows externalism about content. Descartes's claims are classically *internalist*: one's mental states are not essentially determined by any relation to the world.⁹

In fact, reflection on the notion of supervenience will suggest that there is no essential connection between supervenience claims in themselves and externalism. Supervenience is a metaphysical relation between two domains, captured in supervenience claims that are usually expressed as follows: one domain supervenes on another if there can be no variation in the supervening domain without variation in the base domain. This notion can be brought to bear in the analysis of any number of things; for example, what relation is borne between moral and descriptive domains; aesthetic and physical domains; or modal and actual domains.

Where externalism is introduced is with respect to the issue of intentionality. Attention is focused on the relation between states of the head and states of mind, and the question becomes: what facts are sufficient to determine the aboutness relation?¹⁰ In other words, given an entity *a*, we can ask what it is about a subject (if anything) that makes her thought about *a* and not about *b*. Is it sufficient that she be in a particular internal state *S*, that is, does her thought being about *a* rather than *b* supervene on her intrinsic states alone? Or is there more to the story?

5. *Supervenience, Externalism, and Intentionality*

Canonical Twin Earth cases make the point that whether you're thinking about water as opposed to something else does not supervene on your intrinsic states alone. But since, as we have seen, Twin Earth cases cannot be erected for many types of content, we must concede that externalism is a notion that goes well beyond Twin Earth cases. We will show, in fact, that there are cases where content supervenes on inner states but is nevertheless relational. This will involve distinguishing between three ways in which the notion of supervenience might feature in the individuation of content and the problem of content's explanatory role. The critical issue now becomes the composition of what we may call the supervenience base.¹¹

6. *The Composition of the Supervenience Base*

As a general thesis about mental content, externalism ultimately concerns the relation between the internal states of heads and the objects and properties out there in the world that feature as objects of thought. Our question is:

what relations between these elements give rise to intentionality? What — if anything — can intentionality be determined to supervene on?

Supervenience claims can be understood as incorporating claims about necessary and sufficient conditions; claims concerning what conditions are necessary and sufficient to determine a set of properties. When a set of properties is said to supervene on another, the base set is sufficient to determine the supervening set.

With our focus on the nature of intentionality, we can take both the intrinsic states of heads and the objects and properties out there in the world as necessary components of the supervenience base; they are necessary conditions for intentionality, given externalist presuppositions. The issue thus turns on the relations heads can bear to the world. Are causal relations a necessary part of the supervenience base? Or can mental states be relational without them?

We can build up to this by considering the case of sensation states as a preliminary example. Sensation states are not intentional; they are not *about* anything. Accordingly, their individuation is wholly intrinsic. All the supervenience base need include, therefore, are heads; pain states are not directed onto objects the way propositional attitudes are. Sensation states like pain can be said to supervene on intrinsic physiological states; physiological properties are sufficient to determine pain, and there are no other necessary conditions.

In contrast to this, there is the notion of supervenience that underlies the standard Twin Earth cases. Here we must understand the supervenience base to consist of *three* things: 1) intrinsic facts about the subject; 2) facts about the world (the existence of the objects of the relevant mental states); and 3) the causal/environmental relations that the subject bears to the objects of her mental states. These three elements are necessary to determine whether someone is thinking about one thing rather than another; and they are jointly sufficient. Thus content in standard Twin Earth cases is said to fail to supervene on intrinsic states, in the sense that intrinsic states alone are not sufficient to fix the aboutness relation; a causal/environmental relation to the objects of thought is *also* required.

But there is also a notion of supervenience consistent with the most general formulation of externalism. Here the key claim is that causal/environmental relations are *not* part of the supervenience base. This makes room for the view that mental states can be relational (externalist) without being causally/environmentally relational. The supervenience base here should be understood as composed by intrinsic facts, and the objects of thought (the

entities out there in the world), but it will *not* include any causal/environmental relational facts.

Recall that externalism is the claim that mental content is essentially determined by relations to non-mental entities that are independent of the subject of the mental state. Clearly, ethical, formal and aesthetic properties are no less capable of being objects of thought than are, for instance, natural kind or indexical properties¹² The obvious inference to draw, therefore, is that, in many cases — even most — intrinsic states are indeed sufficient (along with the objects and properties themselves) to individuate a mental state. That this is so, however, does not undermine an a genuine externalist thesis for the individuation of mental states. My twin and I are in the same mental states when it comes to “not,” “two,” “beautiful,” and the like. Fix the inner non-relational states and you fix the mental states when it comes to these kinds of concepts; they supervene on inner states. But they are externalist all the same. What you’re thinking about when you’re thinking about “two” is necessarily determined by an entity that is not in your head; a perfectly legitimate formulation of externalism.

7. Summary

According to an uncritical view about externalism, intentionality is relational (it necessitates there being an object of the intentional mental state), and it is thus false to suppose that content supervenes wholly on intrinsic states, as Twin Earth cases show. The orthodox position, however, wrongly presupposes an understanding of supervenience such that the only things that are included in the supervenience base are intrinsic states.

A more accurate characterization of the issue involves a more precise understanding of the composition of the supervenience base. With respect to standard Twin Earth cases, there are three necessary and jointly sufficient conditions included in the supervenience base: intrinsic states, states of the world, and the causal relations between them. For content for which Twin Earth cases do not apply, the supervenience base includes intrinsic states of the head and the existence of the objects of the mental state in the supervenience base, but excludes causal/environmental relations to the objects in question.

The critical contrast here should be understood to devolve on the inclusion or exclusion of causal/environmental relations in the supervenient base; are these necessary elements?

What we have shown here is that it is consistent with the broadest understanding of externalism that internal facts *are* sufficient (in conjunction with the existence of the object of a mental state) to fix intentional states. The

key point, however, is that the mental state is no less relational for all that. Intrinsic facts are *not* sufficient, of course, to fix content subject to Twin Earth cases — for these, there is the further condition that causal/environmental relations be a necessary part of the supervenience base. But the vast majority of concepts are not subject to Twin Earth cases. There is thus more to externalism than may have been hitherto supposed.

And if this is so, then it seems that there may be more to the causal role problem for content than may have been hitherto supposed.

Part II

1. *The Problem of Content in Explanation: Singular Thought*

We can start by showing that there is more to a canonical failure of supervenience problem than meets the eye; that, in fact, the orthodox formulation of the problem of content's explanatory role can be absorbed into a more general one.

Consider the following example. My twin is molecule-for-molecule identical to me, but when she believes that her best friend's husband is unsocialized, her mental state is not the same as mine, since we do not bear the belief-relation to the same items. There is no content-involving psychological generalization that covers our (by hypothesis similar) behavior. My twin does not roll her eyes because she believes *my* best friend's husband is unsocialized.

Now what is the truly relevant factor here? On reflection, it seems clear that content subject to Twin Earth cases fails to supervene because, given externalist taxonomic criteria, *it is individuated in virtue of head-independent entities*. And surely it is this fact *alone* that is sufficient to generate an explanatory role problem for content.¹³ If the individuation conditions of content are held to necessarily involve entities independent of or extrinsic to the inner states that bear the causal burden in explanations of behavior, then such content has no relevant role to play in causal explanation.

2. *The Problem of Content in Explanation: General Thought*

But this entails that an explanatory role problem arises for general as well as singular thought. As we have noted, weak externalism claims that intentional states can be quite naturally thought of as directed onto, and therefore individuated by *properties*; head-independent entities. Thoughts with, for example, arithmetical content are expressed in sentences that contain expressions referring to numbers (number properties). Entertaining this kind of thought puts one in the entertainment-relation to number properties, entities that are

completely head-independent and well beyond a plausible causal range with respect to explaining behavior.¹⁴

This is because commonsense thinking about what matters to a subject's psychology centers on, among other things, what *appears* to her to be the case; how things *look*, what she *believes*. And surely if something extrinsic to or independent of these states can vary without resulting in a variation in these states, then, intuitively, that something will not play an explanatory role with respect to the subject's psychology.

Consider a parallel example concerning the causal/explanatory relevance of a state of knowledge *K*. Even in cases where a belief *cannot* be false (the *cogito*, perhaps) we might still raise the question as to whether it is *K* or the belief *B* (of which *K* is composed) that is the genuinely causally explanatory concept. And on the face of it what appears to be causally relevant to explaining your actions is what belief you possess, *not* whether it is true or justified. So a perfectly coherent question can arise regarding the explanatory role of knowledge, one that is mirrored in our present focus on the formulation of the problem of the explanatory role of weakly external content.¹⁵

So there is more to the problem of content in explanation than its mere failure to supervene. Even when content does supervene, externalistic criteria require that it be individuated with respect to head-extrinsic entities: properties. If we ask what difference the extrinsic, head-independent world of properties makes to behavior, one quite plausible answer is none. Intuitively, one might well behave in exactly the same way in spite of the fact that one fails to be genuinely belief-related to an entity out there in the world. To advert to the extrinsic objects of thoughts in explaining behavior — whether they be particulars or properties — is thus causally irrelevant.

3. *Extrinsicness and the Causal Efficacy of Properties*

Our remarks so far suggest a natural way to describe the explanatory role problem that arises for all content individuated externalistically, either weakly or strongly. Mental states, on this view, are necessarily *extrinsic*: they refer to entities outside the boundaries of what can intuitively be claimed to play a role in the internal causal works — the psychology — of a subject. Extrinsicness is a ubiquitous feature of all content individuated according to externalist criteria.

Extrinsicness should be thought of as a composite idea, involving two claims concerning the causal efficacy of properties. One claim emphasizes the head-independence of abstract properties, the other, their role as content-constituents.

As head-independent entities, properties are outside the causal system whose explanation adverts to beliefs and desires. The claims of intentional psychology are meant to be understood as causal claims: actions are events that are the effects of beliefs and desires. The causal system in question is intrinsic to heads, however; if mental states cause actions, some intrinsic link — including one to physiology — is unavoidable.

But properties don't seem to be where they would need to be to cause action. So it is both difficult to see how mental states that are directed onto properties are capable of causing action, and difficult to see how any explanation that cites content cites any relevant factor in the causation of action. For surely it is a minimal necessary condition (although not sufficient) of the explanatory relevance of a property that the property be *instantiated*; no property not instantiated by a table, say, is going to be relevant to the causal powers of the table. So we might say that a property *P* not instantiated by this table will be “extrinsic” to its causal nexus.

But properties have two roles relevant to the issues here: they can both be instantiated by objects and be content-constituents; that is, be the objects of mental states, and these two characteristics should not be confused.¹⁶ And the deep problem with respect to general content in explanation is this: the properties which are content-constituents are not instantiated by the person, the subject of the mental states in question. In general, of course, one does not instantiate every property one can think about. I do not instantiate numberhood when I think about my bank account.¹⁷

Given these points, it seems that it is the relationality or the extrinsicness of content that results in content's explanatory role problem: supervenience is a red herring. What we have emphasized here is that content is not just a relation to particulars, but also to properties, via the predicate-expressions in contentful sentences. The properties that also individuate content interfere with its causal role in that they themselves are not instantiated by the subject of the mental state in question. Properties fail a minimal condition on relevance in causal explanation — and long before any questions to do with their suitability for being the kinds of things that can be causally relevant can be raised.

Formulating the explanatory role problem for content in terms of *irrelevance* rather than *supervenience* is useful in that it undermines what appears to be the tacit understanding that there are no problematic consequences for content unless it fails to supervene. What we now see is that all concepts to which externalistic determination conditions apply are extrinsic to the machinery to which the fundamentals of material causation must apply.

4. Conclusion

We have argued that all content-citing explanation is irrelevant, whether or not content supervenes on a property-base that includes intrinsic states, states of the world and causal relations to the world. Neither singular thought content nor the content of general thought can play a suitable role in causal mental explanation, given that the individuation relation for such content implicates the extrinsic, head-independent world of particulars and properties.

A real question can be raised now as to whether any of the proposed solutions to this problem — solutions that favor, for instance, the causal role of narrow content, or the ‘shape’ of mental symbols — adequately face up to the fact that all content is subject to a causal role problem.¹⁸ As the explanatory role problem has been shown to be more widespread than has been previously suspected, so further work on this issue must focus on a more general solution.

Notes

*I am grateful to Jerry Fodor and Colin McGinn for much helpful discussion.

1 Certainly the original formulations of externalism *qua* externalism (and many subsequent) do not avail themselves of the notion of supervenience. This way of putting the problem has crept into the externalist literature and has become entrenched most likely as a result of Stich's arguments (1978) concerning content-citing explanation. Stich himself cites Kim's work (cf. 1993) on the supervenience relation to put the point across. The industry generated by Fodor's arguments (1980, 1988) concerning content and explanation has further added to the coupling of the notions of supervenience and externalism.

2 See McGinn (1977); and Burge (1979).

3 The fact is, however, that even for content for which Twin Earth cases are usually thought to apply, a careful examination of the facts reveals that the canonical case — varying the environment while keeping internal states the same — doesn't always make a perfect fit. Take for instance beliefs about oneself. Using ‘I,’ two speakers utter the sentence “I'm hot.” Their internal states are the same, but it is stretching a point to say their *environment* has varied — what has varied is the identity of the speakers. This is sufficient for the purpose at hand, but it isn't a canonical Twin Earth case.

The right thing to say about this is that even for the kind of content for which Twin Earth cases can be erected one should realize that Twin Earth itself is unnecessary in order to formulate externalism, as is varying the environment in some dramatic sense. The more interesting question concerns content for which Twin Earth cases can't be erected, and whether such content can be both externalistically determined and whether it may cause problems for belief-desire psychology — our project here.

4 See McGinn (1989), Chapter 1, for extensive discussion.

5 I pursue this point in another paper.

6 But note that both Fodor and Burge seems to think so (Fodor 1988, 27):

“ . . . I'll assume that the Burge story shows that if the Putnam story raises *any* problems for the notion of content, then the problems it raises are completely general and affect all content-bearing

mental states.”

(Burge 1982, 117): “The Twin Earth thought experiment may work only for certain propositional attitudes. Certainly its clearest applications are to those whose contents involve non-theoretical natural kind notions. But the arguments of “Individualism and the Mental” suggest that virtually no propositional attitudes can be explicated in individualistic terms.”

7 As an added point, note that nowhere does Burge argue that the role of the social environment in the individuation of content is a necessary one.

8 McGinn (1989; 8) writes: “Note that the strong/weak terminology does not mark a difference in the degree of externalism being claimed; the distinction concerns the conditions under which externally determined contents got to be in the mind, not the degree of their externality. One way to think of the contrast between weak and strong externalist claims is to think of it logically; weakly externalist claims have fewer logical entailments than strongly externalist claims.”

9 Of course, it might be thought that Descartes is an externalist with respect to God-thoughts, for these have their content in virtue of God, the only being of sufficient magnitude to cause those ideas in us — and God is an extra-cranial entity. The relevant thing to remember, however, is Descartes’ mentalism. Ideas in the mind — whose origin is innate — are the objects of thought and knowledge. That these Ideas happen to correspond to the world (thanks to the non-deceptive nature of God) is not to say that these Ideas are vulnerable to fluctuations of any sort in the world. God’s job is to ensure veridicality, and hence certainty; the *individuation* of content is not His concern.

10 We need to stipulate that ‘relational’ should be taken to *include* causal/environmental relations, but not to be exclusively restricted to these. This is imperative because there are those, unfamiliar with the weak/strong externalism distinction, who might take ‘relational’ to be equivalent to ‘bears causal/environmental relations to’

11 The issue I am discussing here is orthogonal to those discussed by Kim (1993).

12 Property-realism is presupposed here.

13 Note that adhering to a Twin Earth formulation of externalism will interfere with a grasp of this point.

14 McGinn argues (1989, chapter 2) that an interesting parallel argument, in effect, has been given by Field (1980). His argument that science ought to be done without numbers can be understood along these lines: physics is certainly replete with numerical references, but it is scarcely credible that it is *numbers* that are driving the causal mechanisms that are at work in, say, what happens to water when it reaches 0° Celsius. Numbers are abstract, so they have no causal powers, and, even if they *did*, they are *extrinsic* to the entities whose doings they are employed to characterize.

15 Consider the still more extreme example: even if Descartes were wrong, and you didn’t know that you existed, but you still believed it, your not knowing, on the face of it, would make no difference whatever to your behavior. I am grateful to Colin McGinn for this point.

16 I expand this point in another paper.

17 The property that *is* instantiated by a subject is the property of *having belief b about a property P*. Call this property property P²; it is, of course, is not the same property as property P — the property one has the belief about.

18 See McGinn (1989) Chapter 2 for a fairly comprehensive summary of the main contenders.

Michael N. Liston

Externalist Determinants of Reference

abstract

According to externalism, reference is a relation between uses of an expressive and features of the environment. Moreover, the reference relation is normative (constitutive of correct semantic use), and the referential relations of our expressions are explanatory of successful language use. This paper largely agrees with the broad conception underlying externalism: it is what people do with words that makes them have the references they have, and the world constrains what people can successfully do with words. However, the paper strongly disagrees with the details (at least as usually presented). A centrally important feature of what people do with words is how they use them in inferential contexts. When due attention is given to the reference-determining role played by inferential properties of expressions, I argue, we arrive at a more satisfactory account of semantic norms and explanations. Much of the argument is based on a detailed look at the language of chemical classification used in the late 19th century.

It is customary to think of a Tarski-style definition of truth for a formal language as an abstract model of truth for a natural language.¹ Given a formal language, *L*, whose syntactic specification distinguishes logical (operator) from extralogical (names and predicates) expressions, Tarski showed how to define ‘true-in-*L*’ along the following general lines. For each syntactically simple non-logical expression (name or predicate), ‘*t*’, the definition supplies a primitive reference (or satisfaction) condition. For each logical expression ‘*’, the theory supplies a rule determining the semantic value of a complex expression in which ‘*’ occurs as major operator in terms of the semantic values of its non-logical constituents. For each sentence of the language, the theory supplies a truth condition recursively generated by means of the rules from the primitive reference and satisfaction conditions of its simple constituents. A sentence is true-in-*L* just in case its truth condition obtains. Ideally, when applied to a natural language, this abstract model of truth should enable us to derive the truth conditions of each sentence of the language from a specification of primitive reference and satisfaction for the simple constituents of the language. On the basis of the abstract theory we know that *if* the natural language is treatable as a Tarski-like language, and *if* the specification of primitive reference and satisfaction is correct, then the truth