



Western Michigan University
ScholarWorks at WMU

Master's Theses

Graduate College

12-1996

On Interpreting Strong Supervenience

Zachary J. Ernst

Follow this and additional works at: https://scholarworks.wmich.edu/masters_theses



Part of the Philosophy Commons

Recommended Citation

Ernst, Zachary J., "On Interpreting Strong Supervenience" (1996). *Master's Theses*. 3475.
https://scholarworks.wmich.edu/masters_theses/3475

This Masters Thesis-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Master's Theses by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



ON INTERPRETING STRONG SUPERVENIENCE

by

Zachary J. Ernst

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Arts
Department of Philosophy

Western Michigan University
Kalamazoo, Michigan
December 1996

Copyright by
Zachary J. Ernst
1996

ON INTERPRETING STRONG SUPERVENIENCE

Zachary J. Ernst, M.A.

Western Michigan University, 1996

Jaegwon Kim's definition of strong supervenience has found application in such areas as the mind-body problem, aesthetics, morality, and the relationship between physics and the special sciences. The main reason for the popularity of supervenience is that it purportedly has a long laundry list of virtues. For instance, it has been claimed that supervenience accounts are non-reductive, capable of empirical verification, simple with respect to ontology, and explanatorily powerful.

In this paper, I examine Kim's definition of strong supervenience, arguing that a fundamental ambiguity in the definition makes it impossible for strong supervenience to possess all of these virtues simultaneously. This ambiguity stems from the fact that Kim's definition is written in second-order quantified modal logic, a logic which lacks a standard interpretation. I outline various ways in which we might give a consistent interpretation to this definition. It is seen that each interpretation forces the supervenience theorist to abandon certain purported virtues of supervenience theory.

ACKNOWLEDGMENTS

I owe a debt of gratitude to several people who have lent this thesis what small merit it may possess. First, to the members of my thesis committee, Arthur Falk, Sylvia Culp, and Kent Baldner for their helpful comments on earlier versions of this paper. Second, to Chad Allen for many thought-provoking discussions on this topic. And last, but not least, to Quentin Smith, who graciously lost the bet which resulted in my being presented with a brand-new copy of “Supervenience and Mind.”

Zachary J. Ernst

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
INTRODUCTION	1
VARIETIES OF SUPERVENIENCE	5
Supervenience in General	5
Weak Supervenience Defined	6
Global Supervenience Defined	7
Strong Supervenience Defined	8
A Formal Definition of Strong Supervenience	10
INTERPRETATIONS OF NECESSITY	14
The Ontological Status of Properties	14
Necessity De Dicto	16
Kripke's Modified QS5	16
A De Dicto Formulation of SS	18
Interpreting the De Dicto Formulation	19
Necessity De Re	20
A Platonist Reading of Necessity De Re	21
Natural Kinds	22
CONCLUSION	29

Table of Contents - Continued

ENDNOTES	32
BIBLIOGRAPHY	34

INTRODUCTION

Supervenience has come to play a prominent role in recent philosophical literature in widely diverse contexts. In moral philosophy, aesthetics, and perhaps most importantly, in philosophy of mind, supervenience is often thought of as representing the best hope of solving a broad class of problems. It is hoped that some concept of supervenience can be used to explain the relationship between goodness and the natural properties, beauty and physical properties, as well as between the mental and the physical.

These problems come about when we try to reconcile two highly plausible intuitions. On the one hand, many philosophers are sympathetic to the view that physical properties are, in some sense, all that is 'really real.' Another way of putting this position would be to say that a complete physical description of the world would be complete in every respect. On the other hand, it is quite plausible to maintain that mental, aesthetic, and moral properties are 'real' as well, and would somehow be left out of a purely physical description of the world.

Supervenience theory, by describing the relationship between physical and non-physical, 'emergent' properties, has thus been latched onto as of tremendous potential value. If we could arrive at a satisfactory account of supervenience, a large number of classic philosophical problems could all be solved in one fell swoop.

In addition, supervenience theory, it is commonly believed, can provide an account which is in harmony with other areas of philosophical investigation. For instance, many believe that supervenience accounts can avoid reductionism in the arena of the psycho-physical, can allow us to maintain a simple ontology, can be subjected to empirical verification, and can provide a meaningful, non-trivial explanation.

Eager to arrive at an account of emergent properties, philosophers have begun discussions about these specific purported virtues of supervenience. Discussions about reductionism, the ontological status of supervenient properties, and their roles in causal relations have been heated and numerous. But I believe that these sorts of discussions are premature. It will not be my goal in this paper to discuss the viability of specific supervenience claims, much less to attack or defend supervenience as a whole. Rather, I will attempt to take a step back and examine important questions with regard to the interpretation of Jaegwon Kim's definition of strong supervenience. I will suggest that there are very deep and fundamental questions regarding its interpretation. These questions must be resolved before more specific questions about supervenience claims can be fruitfully addressed.

Another way of putting my programme is this: at present, discussions of supervenience presuppose an understanding of what supervenience claims mean, and set out to determine what they entail with regard to reductionism, the ontological status of supervenience properties, and so on. My claim is that when we focus our attention on the question of how to interpret the definition of strong supervenience,

we find that we have no such understanding of what supervenience claims mean.

Thus, we must first arrive at a satisfactory interpretation of the definition.

As I have said, I will focus my attention on Jaegwon Kim's definition of strong supervenience. The first section of my paper will therefore be devoted to giving a brief overview of his weak, global and strong supervenience. I will give this overview in order to justify my emphasis on strong supervenience. My justification for using Kim's set of definitions is simply that it is this set which has been by far the most influential. After this overview has been given, I will offer a formalized definition of strong supervenience. The formalized definition allows us to more clearly see where problems of interpretation arise.

I will explain in that section that these problems of interpretation stem from the fact that Kim's definition of strong supervenience is written in second-order quantified modal logic, a logic for which no interpretation has been agreed upon. In this logic, modal claims are made with respect to properties. Without an interpretation of this logic, it is far from obvious what exactly these modal claims mean. As will be explained below, this problem is compounded by the fact that Kim's definition employs both the so-called second and third grades of modal involvement. That is, the second-order quantifications lie both inside and outside the scope of modal operators. As Quine has forcefully argued, even in first-order quantified modal logic, we must be very careful to differentiate between first and second-grade modal involvements when interpreting formulae. I believe that a similar caveat applies in this case.

Thus, I will be concerned in the bulk of this paper with outlining how we might interpret Kim's definition of strong supervenience. This will involve sorting out the ontological commitments of second- and third-grade modal involvements in second-order quantified modal logic. Toward this end, I will offer a background logic from which we might begin such an examination. With the proper background logic in place, it will be seen that when we make a consistent interpretation of Kim's definition, we will not be able to simultaneously claim all of the purported virtues of strong supervenience. That is, we will only be able to claim some of them at the cost of abandoning others. Only when we come to accept this fact will we be able to engage in fruitful discussion of other pressing issues regarding supervenience.

VARIETIES OF SUPERVENIENCE

In his paper entitled, “Concepts of Supervenience”, Jaegwon Kim differentiates between weak, global, and strong supervenience. In this chapter, I will review Kim’s distinctions, and justify my emphasis on strong supervenience. This justification will be accomplished by showing how weak and global supervenience are inadequate for most situations in which we would want to employ some concept of supervenience. My justification for concentrating on Kim’s formulations of supervenience is simply that it is this set of definitions which is by far the most often cited in recent philosophical literature.

Supervenience in General

Supervenience in general is a claim that a property (or family of properties) A (call A the supervening properties) covaries with another property (or family of properties) B (call B the supervenience base). Also implied by any concept of supervenience is the claim that for any two objects x and y, and properties A and B, if A supervenes on B then if x and y are identical with respect to B, then they are identical with respect to A as well.

Kim credits G.E. Moore with originating this idea in the following passage, which is quoted from Kim’s “Concepts of Supervenience”:

If a given thing possesses any kind of intrinsic value in a certain

degree, then not only must that thing possess it, under all circumstances, in the same degree, but also anything exactly like it must, under all circumstances, possess it in exactly the same degree. (Kim, Concepts 54)

To adapt an example from R.M. Hare (Kim, Concepts 55), consider two men, Smith and Jones, who are both in happy possession of all the same specific virtues (say, honesty, benevolence, courage, piety, and prudence). If goodness supervenes on the specific virtues, then it follows that Smith is good iff Jones is.

Weak Supervenience Defined

This leads us directly to the concept of weak supervenience which Kim defines as follows:

WS A weakly supervenes on B if and only if necessarily, for any x and y, if x and y share all properties in B, then x and y share all properties in A --- that is, indiscernibility with respect to B entails indiscernibility with respect to A. (Kim, Concepts 58)

It is easy to see that this definition captures the intuition expressed in the above passage from G.E. Moore concerning the relationship between valuational and physical properties.

Why call this 'weak supervenience'? As an answer to this question, consider a situation in which all cars of the same make and model are the same color. All Corvettes are red (say) and all Hondas are blue. It would follow that a car's color weakly supervenes on its make and model. But this is completely compatible with the existence of another possible world in which all Hondas are red and all Corvettes are blue -- for in that world, too, all cars which are identical with respect to make and

model are also identical with respect to color.¹

So this is called ‘weak supervenience’ because it is too weak to support interesting counterfactual claims. In the above case, for example, just because the property of being red weakly supervenes on the property of being a Corvette, and all Corvettes in one world are red, we are still unable to say anything regarding the color of any specific Corvette in another possible world. For instance, we cannot say, ‘had this Honda been a Corvette, then it would have been red.’ Thus, as its name implies, weak supervenience is too weak to be of much use.

Global Supervenience Defined

So if we are to be able to make interesting counterfactual claims from supervenience, we require a stronger definition. The failure of weak supervenience suggests ‘global supervenience,’ which is defined by Kim as follows:

GS A globally supervenes on B just in case worlds that are indiscernible with respect to B (“B-indiscernible” for short) are A-indiscernible. (Kim, Concepts 68)

By making a specific trans-world claim, Kim originally hoped to be able to support interesting counterfactuals. But this, as Kim himself admits, is not the case. (Kim, Strong and Global 82)

To see this, consider the claim that minds globally supervene on brains. Furthermore, let us assume that all people in the actual world have minds. Consider another possible world, exactly like our own, except that in that world, Socrates has a twin. These two worlds are no longer B-indiscernible. Thus, we could consistently

hold that minds globally supervene on brains, but that in this other world in which Socrates has a twin, there are no minds, and Socrates, his twin, and everyone else are mindless robots. This is obviously unsatisfactory, so global supervenience must be abandoned.²

Strong Supervenience Defined

The failure of global supervenience tells us that we require something yet stronger. Thus, we are led to the final stop on our tour, strong supervenience.

Strong supervenience attempts to remedy the problems explained above concerning counterfactuals by making liberal use of strategically-placed modal operators and second-order quantifications. It is defined as follows:

SS A strongly supervenes on B just in case, necessarily, for each x and each property F in A, if x has F then there is a property G in B such that x has G and necessarily if any y has G it has F. (Kim, Concepts 65)

This definition requires some explanation. Consider the Mona Lisa (x), and assume that the beauty of paintings (A) strongly supervenes on their physical properties (B). The definition of strong supervenience requires that there is a sufficient reason why the Mona Lisa is beautiful, and that this reason can be expressed in terms of its relevant physical properties (i.e. some G in B). Furthermore, if any other painting (any y) had exactly those physical properties, then necessarily, it too would be beautiful in exactly the same way that the Mona Lisa is (that is, ' $\Box(\forall y)(Gy \rightarrow Fy)$ '). Because the concept of necessity is used within this definition, we are

guaranteed the truth of claims in the subjunctive mood such as ‘if my painting had all the same physical properties of the Mona Lisa, then it too would have been beautiful.’³

This concept of supervenience does not fall prey to the problems of weak and global supervenience. Consider our first example of redness weakly supervening on the property of being a Corvette. If redness had strongly supervened on the property of being a Corvette, then all Corvettes in all possible worlds would be red as well. Strong supervenience also does not fall prey to the problems of global supervenience, because it does not restrict itself to worlds that are B-indiscernible. The possible existence or counterfactual properties of other paintings makes no difference to the beauty of the Mona Lisa, or any other possible painting with the physical properties of the Mona Lisa.

It is this virtue of strong supervenience which makes it the focus of so much attention, and thus the focus of my attention as well. For example, for a philosopher of mind, it is a virtue of an account of the mind that we are able to make counterfactual claims about possible, but not actual, minds and bodies. After all, if we could not make such claims, then prospects of artificial intelligence researchers would be dim at best. In ethics as well, assuming that goodness supervenes on the specific virtues, it would be highly desirable to be able to make such claims as, ‘Jones has no virtues at all. But if he had possessed all the specific virtues, then he would have been good.’ Weak and global supervenience do not allow us to unconditionally make such claims; only strong supervenience will do. Hence, for the remainder of

this paper, I will confine my discussion to strong supervenience.

A Formal Definition of Strong Supervenience

A quick reading of SS in the previous section will probably not reveal any potential defects. Applications of this definition will seem at first to be straightforward and contentful. However, I will argue that such applications are never straightforward, and often not contentful. When we make supervenience claims, a shallow understanding of the definition encourages us to not consider the metaphysical and semantical entailments of such claims. The first step toward such an understanding is through the analysis of the formalized definition, which is given below (taken from (Horgan, 567)):

$$SS_f \quad \Box(\forall x)(\forall F_{\epsilon A})\{Fx \rightarrow (\exists G_{\epsilon B})[Gx \ \& \ \Box(\forall y)(Gy \rightarrow Fy)]\}$$

The formal definition SS_f allows us to more clearly see a few unusual features. First of all, the first ‘ \Box ’ is used in the so-called ‘second grade of modal involvement’ in which it is placed outside the scope of a quantification (making the formula ‘wide scope’). Adding to the confusion, SS_f also contains the ‘third grade of modal involvement’ (because the ‘ \Box ’ is inside the scope of ‘ $(\exists G)$ ’).⁴ These quantifications in themselves, aside from their modal involvements, are unusual as well, for they are second-order quantifications. In themselves, neither a second-grade modal involvement nor a second-order quantification are inherently problematic; even Quine will allow the second-grade of modal involvement under some restrictions. The third grade of modal involvement, of course, has been vehemently attacked by Quine as

being incoherent (Quine, Paradox). However, even if we do not oppose the third grade of modal involvement, interesting questions are raised when we place a second-order quantification outside or inside the scope of a modal operator.

First, there is an ambiguity surrounding the interpretation of any necessity claim. When we assert that ' $\Box p$ ' (where p may be atomic or complex), we could intend to use ' \Box ' as expressing nomological, de re, or dicto necessity.⁵ If we intend to employ nomological necessity, then by ' $\Box(Pc)$ ' we mean 'it is a consequence of natural law that Pc .' By de re necessity, we would interpret the same formula as 'it is in the essence of c to display P .' Lastly, if we use de dicto necessity, then we mean 'it is an analytic truth that Pc .'

Assaults on each of these versions of necessity are notorious, especially those due to Quine concerning analyticity (Quine, Dogmas) and de re necessity. But for the purpose of this paper, I prefer not to enter this thicket. I will grant that it is indeed in the essence of human beings to possess a certain genetic structure, that 'all bachelors are unmarried' is an analytic truth, and that it is a nomologically necessary truth that 'unsupported objects, heavier than air, will fall when near the surface of the earth.' Rather, what I would like to accomplish is to outline the ramifications of each of these three kinds of necessity for supervenience claims.

The inconsistent use of wide and narrow scope in SS_1 emphasizes these ambiguities. This is because the interpretations for wide and narrow scope correspond to the distinction between de re and de dicto necessity. When the modal operator falls outside the scope of a quantification (wide scope), the formula is

interpreted de dicto. That is, the necessity claim is said to attach to a sentence, and necessity is then understood as analyticity or tautology. On the other hand, when the operator is within the scope of a quantification (narrow scope), then the formula is interpreted de re. In the latter case, necessity is understood to attach to things, and implies that the sentence asserts of a thing that it has its essence (or denies that it does not have its essence).

Although this inconsistent use of wide and narrow scope does not introduce any outright contradictions into the definition of strong supervenience, it does imply that the supervenience theorist is committed to an equivocal reading of ‘necessity.’ For if wide scope corresponds to de dicto necessity while narrow scope corresponds to de re necessity, then supervenience claims imply de dicto necessity with respect to the supervening properties, and de re necessity with respect to the supervenience base. Thus, without an understanding of how different readings of necessity will impact supervenience claims, the inconsistent use of wide and narrow scope is doubly unwarranted: first, because we do not understand the implications of wide or narrow scope for supervenience claims, secondly because we do not know if an inconsistent use of wide and narrow scope will have undesirable repercussions on supervenience claims.

So the task ahead is to delineate different ways in which this massively ambiguous definition can be disambiguated. This is an important project, for whether we choose de re, de dicto, or nomological necessity, we will find that this choice places severe demands upon our ontology, and will thus impact the claims which

supervenience theorists can consistently make. I will turn to de dicto, de re, and nomological necessity in turn. Prior to this discussion, however, it will be useful to have in place a set of metaphysical conceptions of properties.

INTERPRETATIONS OF NECESSITY

As I have indicated above, the ambiguities which plague interpretation of the '□' operator call for an examination of each possible interpretation, with special consideration on how such interpretations affect the ontological status of properties. I will turn in the first section of this chapter to a consideration of the ontological status of properties. After this is in place, I will use these considerations in a discussion of the interpretations of necessity.

The Ontological Status of Properties

A wide variety of conceptions of the ontological status of properties have been held. These may be characterized as being along a spectrum which runs from Plato to the nominalists. Following Mario Bunge's characterization of these positions, I will call them 'strong,' 'moderate,' and 'weak':

strong: Properties, whether intrinsic [unary] or mutual [relational], are real, nay supremely real, and individuals only exemplify them.

moderate: Whereas intrinsic [unary] properties are real, mutual [relational] properties are not.

weak: All properties, whether intrinsic or mutual [unary or relational] are unreal: only individuals are real. (Bunge, 100)

It is easy to see that the strong conception of properties is Plato's view: properties have an existence which is completely independent of our conceptions of them, and would exist even if they were never exemplified. The moderate view most

likely evolved in an attempt to discredit the so-called ‘Cambridge Changes’ -- purported changes in a thing which result from its losing or gaining a trivial relational property such as ‘being at such-and-such a distance from Alpha Centauri.’ The weak conception is consistent with a nominalist view, diametrically opposed to Plato, but possibly in line with the late Wittgenstein, that denies the existence of propositions entirely (and thus of the properties out of which propositions are constructed).

I will focus on the two extreme conceptions, strong and weak, in order to simplify the discussion. Although supervenience, if it is a viable concept, would most likely employ many relational properties, discussions of supervenience most often are simplified through the exclusion of these properties. Also, we can easily subsume relational properties under unary properties by substituting for ‘Fab’ the sentence ‘F*a’ where ‘F*x’ is read as ‘x has the property of F-ing B.’

These two positions regarding the ontological status of properties immediately suggest the first distinction we must keep in mind when interpreting the definition of strong supervenience, namely:

1. **PY:** Yes, properties exist. They may not reside in a Platonic realm of forms, but they have an independent existence.

2. **PN:** No, properties do not exist. Properties are words which denote a set of individuals which humans have grouped together for some purpose or other. They have no independent existence aside from the use of our language, and the sets into which we collect objects.

At this point, we may disambiguate the language which I have used thus far. From here on in, I will use the word ‘property’ only when operating within the **PY** conception. When operating within **PN**, I will use the term ‘predicate.’ This follows standard usage, in which a predicate is taken to mean a phrase such as ‘is red,’ which can be used to construct a proposition. A property is (loosely speaking) a characteristic which an object might possess; predicates are merely speech acts with which we collect things into sets.

Necessity De Dicto

Because the task of this paper is to examine the metaphysical and semantical entailments of supervenience claims, the definition **SS_r** must be revised. This is because of its use of both wide and narrow scope. As it is plainly incoherent in general to adopt one set of ontological commitments with respect to the properties in **B** and another set of commitments with respect to the properties in **A**, we need to revise **SS_r** in order to make its use of modalities consistent. Therefore, at this point, I must explicate a background logic which will enable us to make these revisions without at the same time conflating wide and narrow scope.

Kripke’s Modified QS5

As I have said above, it is not my intention in this paper to enter into a discussion about the possible meaninglessness of de re or de dicto claims. Rather, I

will assume that such claims can be (and often are) meaningful. My intention is simply to outline how supervenience claims are affected by de re, de dicto, or nomological interpretations of ‘ \Box .’ Thus, it would be helpful to employ as a background logic a system which allows for each interpretation of ‘ \Box ,’ without conflating them.

To do this, I will employ Kripke semantics as revised in 1963. The reason for excluding the original version of Kripke semantics is that it is too permissive for my purposes here. That is, originally, Kripke’s quantified modal logic (call it **QS5_o**) allowed derivation of the Barcan and converse Barcan formulae (Konyndyk, 94-5):

$$\mathbf{BF} (\forall x)\Box Px \rightarrow \Box(\forall x)Px$$

$$\mathbf{CBF} \Box(\forall x)Px \rightarrow (\forall x)\Box Px$$

BF and **CBF** tell us that **QS5_o** is too strong. This is because, if we allow **BF** and **CBF**, then a de re claim will be true if and only if the corresponding de dicto claim is true. But de re claims carry vastly stronger ontological commitments than de dicto claims. Thus, if we allow **BF** and **CBF**, then there will be no real difference between the claim that a sentence is analytically true, and the claim that a thing possesses a property essentially, for **BF** and **CBF** tell us that one claim is true only if the other is true as well. So if the move from analyticity to essentialism (and vice-versa) is permitted, then we will not be able to even begin to address questions of how to interpret the definition of strong supervenience; all interpretations will entail the ontological commitments of all other interpretations.

So we do not want our background logic to be this permissive. But on the

other hand, neither do we want it to be too austere. That is, we want for de dicto and de re claims (wide scope and narrow scope formulae) to be permissible, but we do not want a de re claim to be derivable from a de dicto claim, or vice-versa.

Kripke's modified semantics walks this line quite nicely. In this system (call it $QS5_r$), **BF** and **CBF** are permissible (that is, do not entail inconsistencies), but are not derivable. So we are permitted to write formulae with either wide or narrow scope (corresponding to de dicto and de re, accordingly), but we cannot derive one from the other.⁶ The system $QS5_r$ will therefore serve our purposes quite well by allowing us to keep de re and de dicto claims separate, yet still allowing us to examine both varieties of formulae. Thus, I will be assuming $QS5_r$ for the remainder of this paper.

A De Dicto Formulation of SS

With $QS5_r$ as a background logic, we are permitted to write the following revised formulation of SS_r :

$$SS_r^d \quad \Box(\forall x)(\forall F_{\epsilon A})\{Fx \rightarrow \Box(\exists G_{\epsilon B})[Gx \ \& \ (\forall y)(Gy \rightarrow Fy)]\}$$

In SS_r^d , we have moved the second ' \Box ' to a position to the left of ' $(\exists G)$,' which makes it wide scope. The definition is now consistent in its use of second-grade modal involvement. Assuming $QS5_r$, we are assured that SS_r^d is a wff. With this assurance, we can now turn to the question of its interpretation.

Interpreting the De Dicto Formulation

Would supervenience theorists be happy with $SS_f^{d?}$? Consider that for a proposition to be necessary de dicto, it must be analytic.⁷ A sentence is analytic only when the meaning of the predicate is contained (as Kant writes in the First Critique) in the meaning of the subject, such as in ‘all bachelors are male.’ Keeping this in mind, let us return to the definition of strong supervenience for a moment, considering the claim that A strongly supervenes on B:

SS A strongly supervenes on B just in case, necessarily, for each x and each property F in A, if x has F then there is a property G in B such that x has G and necessarily if any y has G it has F.

If **SS** is to be interpreted de dicto, then the definition of strong supervenience could be restated as follows:

SD A strongly supervenes on B just in case for each predicate F in A the meaning of ‘F’ contains the meaning of a disjunction, one of whose disjuncts is G in B, and the meaning of ‘G’ contains the meaning of ‘F’.

But if **SD** is the best way to interpret strong supervenience claims, then the concept of strong supervenience has become trivial. Such investigations would be merely investigations into the meanings of words, and could all be resolved by appeal to dictionaries. Supervenience would be a waste of time to at least the degree to which it would be a waste of time to conduct a survey of bachelors to determine how many were unmarried.

The triviality of supervenience claims results from a de dicto reading of

necessity because in such a reading, our ontology has become too impoverished to carry out interesting investigations. If there are no properties, and only predicates, then everything hinges on the meanings of words. If everything is allowed to hinge on the meaning of words, then investigations into supervenience will be rendered trivial.⁸ It would be unfair to supervenience theorists to leave their situation like this; yet it is interesting to note that SS_r 's use of wide scope with regard to the supervening property F gives this reading a certain level of plausibility.

Necessity De Re

In the discussion above, we saw that an interpretation of ' \Box ' as 'it is analytically true that...' is unsatisfactory.⁹ In this section, I will examine a few possible de re interpretations of a revised version of SS_r which is given here:

$$SS_r^f (\forall x)(\forall F_{\epsilon A})\Box\{Fx \rightarrow (\exists G_{\epsilon B})\Box[Gx \ \& \ (\forall y)(Gy \rightarrow Fy)]\}$$

Assuming $QS5_r$, we are again guaranteed that SS_r^f is at least well-formed, and doesn't generate inconsistencies in our background logic. Again, we turn to the question of interpretation. Here, I will focus on two possible interpretations, beginning with a strongly Platonist interpretation, and then moving to a more moderate reading. My goal here is not to exhaust all possible de re interpretations, but to define a spectrum into which most interpretations will fall.

A Platonist Reading of Necessity De Re

The major challenge in interpreting the formal definition SS_f' lies in the fact that, unlike an ordinary de re claim, this one is a claim about properties. That is, this formulation of Kim's definition of strong supervenience forces us to consider the ontological implications of the claim that properties necessarily stand in certain relationships to each other. As we saw above, by reading ' \square ' as expressing analyticity, we would impoverish our ontology to the point at which no interesting non-linguistic claims can be made. One obvious way to avoid this consequence is to go to the opposite extreme and consider a Platonist ontology in which properties are "supremely real" (Bunge, 100).

Operating from within a Platonist perspective, it is easy to see how a claim such as 'goodness strongly supervenes on the specific virtues' might run. In the dialogues, Socrates examines the question of the unity of the virtues. Most notably in the *Protagoras*, Socrates comes to the conclusion that a man is good iff he is in possession of the five specific virtues: wisdom, courage, temperance, justice, and holiness. Without belaboring the point, I think that it is obvious that Plato would hold that it is a necessary truth that only those men in possession of the specific virtues are good. Thus, Plato would be able to accept the thesis that goodness strongly supervenes on the specific virtues, so long as we understand the relevant necessity claim as attaching to the Platonic forms, which exist in a mind-independent reality.

This picture of necessity in second-order quantified modal logic (call it

'2QS5') is diametrically opposed to the view which emerged from a reading of '□' as analyticity. Our ontology has not been diminished in the slightest, and there is no way in which we might hold that the supervening properties are unreal. Rather, they are "supremely real," more real, in fact, than ordinary objects in the world.

However, we pay a high price for the ability to hold onto the reality of supervening properties. Certainly, most contemporary philosophers would be uncomfortable with this solution; the cure may be worse than the disease. For in order to use this conception of properties, we would have to answer objections from everyone from Parmenides to D.M. Armstrong. Operating under the assumption that any philosophical 'solution' which forces us to answer classic philosophical problems is unsatisfactory, we require a more moderate realist view of properties.

Natural Kinds

For a more acceptable view of the ontological status of properties, we may turn to the doctrine of natural kinds. Those who accept that there are natural kinds are committed to metaphysical realism (Wilkerson, 30), but this metaphysical realism is vastly weaker than Platonism. The doctrine of natural kinds is fully compatible with the view that properties have no mind-independent existence, in Plato's sense. Such a metaphysical realist can hold that the entities studied by physics constitute 'all there is.'

Philosophers such as Quine and Putnam who accept that there are natural

kinds require only that nature has ‘joints’ at which it may be carved. The things of the world, which may be as diverse as tigers, gold, and electrons, fall into categories in such a way as to allow science to limn the nature of reality. Putnam sums up the doctrine of natural kinds with the following two ideas:

1. To belong to a natural kind, something must have the same composition, or obey the same laws --- indeed, what makes composition important, when it is, is its connection with laws of behavior --- as model members of the class, and this composition or these laws are not usually known when the natural kind term is introduced, but require an indeterminate amount of investigation to discover.
2. Natural kind terms and proper names are not synonymous with conjunctions of criteria and definite descriptions respectively.
(Putnam, 74)

For example, whether or not a metal is gold depends on whether its essence is the same as the essence of paradigmatic samples of gold. In this example, the essence of gold is its atomic structure. Thus, the property of ‘being gold’ can be cashed out as ‘having such-and-such an atomic structure.’

Although the doctrine of natural kinds revives Aristotelian essences, it does allow us to avoid the extravagant ontology of Plato. This is because it is not the property (say) of ‘being gold’ which is ‘really real.’ Rather, it is the atomic structure shared by sample of gold which is ‘really real.’ So the doctrine of natural kinds permits us to walk a very fine line. That is, some things really are (say) gold, but only because they share a certain atomic structure, and not through participation in a Platonic Form.

So if we are to interpret the properties in SS_f^r as denoting membership in a

natural kind, we are not committed to Plato's two-realm ontology, with its attendant set of difficulties. Furthermore, we are not reducing the necessity claim to mere analyticity, which would trivialize investigations into supervenience. This is because properties, on this view, are 'real,' but only because they denote membership in a set whose objects necessarily have certain essences.

So have we arrived at an interpretation of strong supervenience which allows us to keep all the virtues which supervenience theorists want? Predictably, I think that the answer is an unqualified 'no.' To see this, let us recap the natural kinds interpretation of strong supervenience.

First, on this view, properties are not 'really real'; rather they denote membership in a natural kind. Second, both the supervenient family of properties and the supervenience base are natural kinds.¹⁰ Third, there is a necessary connection between an object having a supervening property and having a base property.

The problem for supervenience theorists is that these are sufficient conditions for reductionism. To see this, it would be helpful to take a step back and consider an argument due to Jaegwon Kim to the effect that reductionism always obtains between a supervenient and subvenient domain (Kim, Nomological). Kim's argument fails in general. However, if we add the premise that the properties in the supervenience and subvenient domains are natural kinds, then the argument will go through.

Kim asks us to consider, for two domains A and B, such that A supervenes on B, the sets of A-maximal and B-maximal properties (call them A* and B*). Consider

two objects, x and y , such that x and y share the same B-maximal property, call it Q . It follows immediately from any definition of supervenience that x and y must share the same A-maximal property, call it R . Assuming that A^* and B^* are finite, then there must be a finite set of B-maximal properties which will ‘cause’¹¹ any given A-maximal property to be instantiated. So we can easily construct one-half of the biconditional which is required for a reduction of A to B.

The converse follows as well. To deny the converse would be to claim that some object x could have a property R in A^* , and yet not have any property Q in B^* , where Q is a sufficient condition for the instantiation of R . But this contradicts the assumption that A supervenes on B, for to make such a supervenience claim is to assert that each R in A^* has a sufficient ‘cause’ which is some Q in B^* . That is, for each property in A^* , at least one property in B^* must be instantiated in x , and it is the subvenient property which accounts for x ’s having the supervenient property. So if we take the disjunction of Q ’s in B^* which could form a supervenience base for R , then that disjunction will be a necessary condition for R .

So what we are left with is a biconditional for each supervenient property R in A^* :

$$Rx \equiv Q_1x \vee Q_2x \vee \dots \vee Q_nx$$

Kim stops here, claiming that the existence of such biconditionals is sufficient for reductionism. But this is not the case. Jerry Fodor, in his article entitled *Special Sciences*, gives compelling reasons why the existence of such a biconditional is not

sufficient for reductionism.

According to Fodor, there is a feature of such biconditionals which forces them to fall short of providing a full-blown reduction. This is the fact that such biconditionals are massively disjunctive. For Fodor, this indicates that the predicates on at least one side of the biconditional are not natural kinds and that the biconditional will not support counterfactuals. That it will not support counterfactuals is because a predicate which does not denote a natural kind may be instantiated by any number of predicates on the other side of the biconditional. Thus, when constructing such a massively disjunctive biconditional, we can never be guaranteed that the list of predicates on either side is complete. Hence, the biconditional might not be satisfied under counterfactual circumstances. Because a minimal criterion for lawlikeness requires that counterfactuals be supported, such biconditionals fail to be laws. This shows that reductionism does not obtain in these cases.

To illustrate how Fodor's objection blocks reductionism, consider an example from the philosophy of science. The question of whether biology reduces to chemistry is a hotly contested issue. Alexander Rosenberg has recently argued in *Instrumental Biology or the Disunity of Science* that it is impossible for human beings to achieve a reduction of biological kinds to chemical kinds. The reason for this, according to Rosenberg, is that the kinds with which biology is concerned reflect "human needs, interests, and limits" (Rosenberg, 5). By this account, our needs,

interests, and limits force biologists to organize biological kinds according to their function, not their structure. However, biological functions do not correspond in any systematic way to chemical structure. Hence, a single biological kind may be instantiated by a heterogeneous disjunction of different chemical structures. So we see that, because natural kinds are given by their structures, biological kinds cannot be natural.

Chemical kinds, on the other hand, are natural kinds. This is entirely in line with Putnam's characterization of natural kinds given above, for chemical kinds are determined by their structures, and it is their structures which are essential to chemicals.¹² On Fodor and Rosenberg's view, the fact that biological kinds are not natural while chemical kinds are natural is the feature of biology and chemistry which blocks reduction. For it is owing to this fact that any purported bridge law which attempts to reduce biology to chemistry will be wildly disjunctive.

However, Fodor's objection is blocked in this case. First, we are assuming that the predicates denote membership in a natural kind. Second, because we are concerned only with strong supervenience, the support of counterfactuals is guaranteed. So we are left with the fact that the biconditional under consideration is lawlike after all. Therefore, the biconditional does qualify as a bridge law, and we have a full-blown reduction of A to B.

In the light of Fodor's quite plausible qualms about calling massively disjunctive biconditionals 'laws,' it may seem odd to use the previous argument to

imply that (say) the mental can be reduced to the neurological. After all, such a bridge 'law' may very well contain literally a billion disjuncts. But this does not tell against my argument. Rather, I think that this unpalatable conclusion indicates that in the case of the mental supervening on the neurological, it is unreasonable to take neurological predicates as natural kinds.¹³ That is, we are better off treating this argument as a *reductio ad absurdum* of the premise that neurological kinds are natural kinds.

To round off the discussion, we can note that interpretation of predicates as natural kinds entails that we read '□' as nomological necessity. As we have seen, if predicates are natural kinds, then for each strong supervenience claim, we can (in principle) derive a reductionary law reducing the supervenient to the subvenient. Thus, the necessity claim must be taken as meaning that there is a lawlike equivalence between the supervenient and the subvenient. This is entirely in keeping with the concept of natural kinds, for an object is a token of a natural kind iff it obeys certain laws (see Putnam's first criteria for membership in a natural kind, given above).

CONCLUSION

I admit that I am often baffled by supervenience claims, particularly with regard to the mind-body problem. In this paper I have attempted to explain why I am so baffled.

One way to recap my argument is through a thought-experiment. Consider a clever philosopher who has established, beyond a doubt, that the mind strongly supervenes on the body. Suppose that we were able to have a conversation with this philosopher. No doubt, we would say something like, “So now you can answer a burning question. What is the nature of the mind? Is it ‘really real,’ or is mind-talk just abbreviated brain-talk? Does your discovery rule out dualism, or does it imply it? Does the mind have causal powers? And can we reduce the mental to the physical?”

Surely, any good account of the mind should be able at least to address these sorts of questions. However, I am unable to find anything in the concept of strong supervenience which can even begin to. This, I have argued, is because the primitives in the definition, namely, the properties and necessity operators, are left hopelessly ambiguous. But this ambiguity has not prevented supervenience theorists from making grandiose claims about the virtues of supervenience theory.

In this paper, I have outlined in broad form how we might start to disambiguate strong supervenience. It was seen that each disambiguation forces the supervenience theorist to give up certain purported virtues of supervenience theory.

We may summarize the possible positions as follows:

1. Necessity is analyticity, and properties are actually predicates, or speech acts.

2. Necessity is to be understood de re, and because properties necessarily stand in certain relations to one another, properties have a human-independent existence somewhat akin to Platonism.

3. Necessity is to be understood as nomological. Properties denote membership in a natural kind.

Under the first reading, supervenience claims are rendered trivial, because they only apply to speech acts. We also get a sort of trivial ‘linguistic reduction.’ The second view gets around the difficulties of the first, but at a tremendous cost in ontological simplicity. We are committed to some variant of Platonism, and we must therefore resolve all the philosophical problems which have plagued that view for thousands of years. On the third view, our ontology remains simple, and our investigations are not trivial, but we cannot avoid reductionism.

How this trilemma would be resolved will depend upon one’s individual preferences. The Churchlands, for instance, would be quite happy to accept this position, because they could rest comfortably on the first horn of the trilemma. On their view, mind-talk is abbreviated brain-talk, so necessity is analyticity and minds are not ‘really real.’ A moral realist, on the other hand, could accept the second horn, and would say something like, ‘So goodness does turn out to be ‘really real’ after all.’

Anyone trying to naturalize philosophy who accepted reductionism in the particular case under consideration could easily accept the third horn.

These are all reasonable responses to my argument. What is an unreasonable response is to make blanket statements about the many and varied virtues of supervenience theory. Instead, we should face up to the old maxim that 'You pays your money and you takes your choice.'

ENDNOTES

1. Compare with (Kim, Concepts 59-60).
2. Compare with (Kim, Strong and Global 83-4) where Kim writes, “It is difficult to see how, given worlds like these, F’s global supervenience on G could support G--F subjunctives or counterfactuals...”
3. Compare with Kim’s example of Saint Francis in (Kim, Concepts 65-6).
4. At first glance, this formula seems to be using the second grade throughout, because the ‘ \Box ’ appears to the left of ‘ $(\forall y)$.’ But it is important to keep in mind that the relevant quantification is over the predicate G. Therefore, because the necessity operator occurs within the scope of ‘ $(\exists G)$,’ this formula uses both the second-grade of modal involvement as well as the third.
5. Kim identifies four possible interpretations of necessity: “metaphysical, logico-mathematical, analytic, and nomological” (Kim, Philosophical Concept 141) and goes on to say that the meaning of ‘necessity’ should be left “unfixed... to be interpreted to suit specific supervenience claims.” I am subsuming metaphysical necessity under the heading of ‘de re,’ analytic under ‘de dicto,’ and assuming that logico-mathematical necessity is a variety of analytic necessity, and thus can be treated under the heading of ‘de dicto.’ Obviously, I disagree with his second claim.
6. Kripke accomplished this by noting that in order to derive **BF** and **CBF**, it is necessary to employ the reiteration S5 rule on a sentence with a free variable. By prohibiting use of this rule on unbound formulae, derivations of **BF** and **CBF** are blocked.
7. I am excluding tautologies and synthetic a priori truths from consideration. Whatever a supervenience claim means, I think that it is obviously not a tautology. I am excluding synthetic a priori truths from the discussion because if supervenience claims are synthetic a priori, then this would imply that we could determine which properties supervene on other properties a priori, without empirical investigation. This is a highly dubious proposition, especially with regard to the mind-body problem.

8. Of course, some ordinary language philosophers would want to return to the days of “O.E.D. is Q.E.D.”, and this would be a good way to accomplish this. But I will not give such views serious consideration here.
9. Although I will indicate in chapter four that some philosophers of mind would probably not object to this reading at all.
10. Of course, we might be tempted to claim that the supervening family is comprised of natural kinds while the supervenience base is not. However, there is a problem with this which relates to what we might call ‘transitive supervenience.’ For example, consider the claim that goodness supervenes on psychological states, and psychological states supervene on brain states. If we use a double standard, and claim that the supervenient family is composed of natural kinds while the supervenience base is not, then we are committed to the view that psychological states both are and are not to be understood as natural kinds.
11. It is highly problematic to use the word ‘cause’ here, because what Kim calls ‘supervenient causation’ may fail criteria for Humean causation. We may substitute ‘brings about’ or some such phrase for ‘causes.’
12. See (Rosenberg, 20-6) for a more detailed analysis of why reductionism fails in the case of biology and chemistry.
13. And perhaps mental predicates are not natural kinds, either. See (Rorty) for a sustained discussion of the artificiality of mental predicates.

BIBLIOGRAPHY

- Mario Bunge. *Ontology I: The Furniture of the World*, volume 3 of *Treatise on Basic Philosophy*. D. Reidel Publishing Company, Boston, 1977.
- Jerry Fodor. Special Sciences. In *The Philosophy of Science*, pages 429-41. MIT Press, Cambridge, MA, 1993.
- Terence Horgan. From Supervenience to Superdupervenience. *Mind*, 408: 555-86, October, 1993.
- Jaegwon Kim. Supervenience and nomological incomensurables. *American Philosophical Quarterly*, 15(2):149-156, April 1978.
- Jaegwon Kim. Concepts of supervenience. In *Supervenience and Mind*, pages 53-78. Cambridge University Press, Cambridge, 1993.
- Jaegwon Kim. Strong and global supervenience revisited. In *Supervenience and Mind*, pages 79-91. Cambridge University Press, Cambridge, 1993.
- Jaegwon Kim. Supervenience as a philosophical concept. In *Supervenience and Mind*, pages 131-60. Cambridge University Press, Cambridge, 1993.
- Kenneth Konyndyk. *Introductory Modal Logic*. University of Notre Dame Press, Notre Dame, Indiana, 1986.
- Hilary Putnam. *Realism and Reason*. Cambridge University Press, Cambridge, 1961.
- Willard van Orman Quine. Two dogmas of empiricism. In *From a Logical Point of View*, pages 26-46. Cambridge University Press, Cambridge, 1961.
- Willard van Orman Quine. *The Ways of Paradox*. Random House, New York, 1966.
- Richard Rorty. *Philosophy and the Mirror of Nature*. Princeton University Press, Princeton, New Jersey, 1979.

Alexander Rosenberg. *Instrumental Biology, or the Disunity of Science*. University of Chicago Press, Chicago, 1994.

T.E. Wilkerson. *Natural Kinds*. Ashgate Publishing Company, Brookfield, Vermont, 1995.