

The Metaphysics of Dispositions

by
Jennifer McKittrick

B.A. Brown University, 1994.

Submitted to the Department of Linguistics and Philosophy in
Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
at the
Massachusetts Institute of Technology

September, 1999.

© 1999 Jennifer McKittrick. All rights reserved.

The author hereby grants M.I.T. permission to reproduce and to distribute publicly
paper and electronic copies of this thesis document in whole or in part.

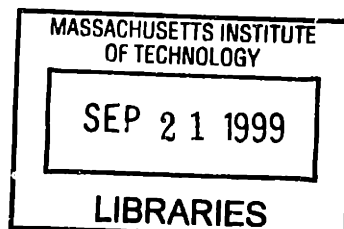
Signature of the Author: _____
Department of Linguistics and Philosophy

Certified by: _____
Alex Byrne
Associate Professor, Department of Linguistics and Philosophy
Thesis Supervisor

Certified by: _____
Stephen Yablo
Associate Professor, Department of Linguistics and Philosophy
Thesis Advisor

Certified by: _____
Robert Stalnaker
Professor, Department of Linguistics and Philosophy
Thesis Advisor

Accepted by: _____
Robert Stalnaker
Professor, Department of Linguistics and Philosophy
Acting Chair of the Committee on Graduate Students



ARCHIVES

The Metaphysics of Dispositions

by

Jennifer McKittrick

Submitted to the Department of Linguistics and Philosophy on July 30, 1999 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Philosophy.

Abstract

As Nelson Goodman put it, things are full of threats and promises. A fragile glass, for example, is prone to shatter when struck. Fragility is the glass's *disposition*, shattering is the *manifestation* of the disposition, and striking is the *circumstances of manifestation*. The properties of a fragile glass which are causally efficacious for shattering constitute the *causal basis* of the glass's fragility. The glass can remain fragile even if it never shatters. One can say of the fragile glass, with certain qualifications, that if it were struck, it would shatter. This much is common ground among philosophers who discuss dispositions. In my dissertation, I defend three claims about dispositions that are more controversial.

Some philosophers have claimed that dispositions are causally impotent. I disagree. In my first chapter, I defend the claim that dispositions can be causally efficacious with respect to their manifestations. Among the arguments I consider is the "no work" argument, according to which a disposition's causal basis causally explains its manifestation, leaving no causal work for the disposition to do. I respond to this argument by challenging the Principle of Explanatory Exclusion, according to which complete explanations exclude competitors.

Furthermore, many philosophers hold that all dispositions must have independent causal bases. In my second chapter, I challenge this view, and defend the possibility of *bare dispositions*. I argue that the concept of a bare disposition is coherent, and show why arguments recently offered against bare dispositions, such as those based on the Truth Maker Principle, do not succeed in demonstrating that they are impossible.

Another common assumption about dispositions is that they must be intrinsic properties of the objects that have them. In my third chapter, I challenge this assumption, and argue that some dispositions are extrinsic properties. Consider the property vulnerability. It seems dispositional in character; something which is vulnerable is susceptible to harm, but is not necessarily being harmed right now. However, it seems as if something could lose the property of being vulnerable without undergoing any intrinsic change. Build a fortress around the vulnerable object and it ceases to be vulnerable.

Thesis Supervisor: Alex Byrne

Title: Associate Professor of Linguistics and Philosophy

Acknowledgments

Thank you to Ned Hall, Michael Glanzberg, Judith Thomson, James Pryor, Daniel Stoljar, Andrew Melnyk, Roderick Long, Philip Bricker, James S. Taylor and especially my thesis committee Alex Byrne, Stephen Yablo, and Robert Stalnaker, for comments on earlier versions of this material.

Thank you to those in attendance at various presentations of material from Chapter 1, including participants at the 1996 Harvard/MIT Graduate Conference in Philosophy, the philosophy departments of St. Louis University, the University of Delaware, Florida International University, and the University of Alabama at Birmingham. Thanks also to participants at the University of Massachusetts at Amherst Philosophy Colloquium for their comments on a presentation which served as a basis for Chapter 3.

Thank you to friends and family for encouragement and support.

This thesis is dedicated to Matthew.

Table of Contents

Chapter 0: Introduction	5
0.1 Dispositions	5
0.2 Dispositional versus Categorical	8
0.3 Properties	12
Chapter 1: A Defense of the Causal Efficacy of Dispositions	18
1.1 Causal Relevance and Causal Efficacy	20
1.2 The Analyticity Argument	26
1.2.1 Hume's Principle	29
1.3 The No Work Argument	34
1.3.1 The Distinctness Thesis	36
1.3.1.1 The Disjunctive Solution	37
1.3.1.2 The Trope Solution	42
1.3.2 The Exclusion Principle	45
1.4 Conclusion	49
Chapter 2: The Bare Metaphysical Possibility of Bare Dispositions	51
2.1 Introduction	51
2.1.1 Causal Bases	53
2.1.2 "Bare Disposition" Defined	55
2.2 The Conceptual Argument	57
2.3 PPJ's Causal and Non-Identity Theses	59
2.3.1 The Causal Thesis	59
2.3.2 The Non-Identity Thesis	63
2.3.2.1 The Arguments from Multiple Realizability	64
2.3.2.2 Swamping the Disposition	66
2.3.2.3 Concluding Remarks about the Non-Identity Thesis	67
2.4 Bare Counterfactuals	68
2.4.1 Smith and Stoljar's Argument	70
2.4.2 The Bare Dispositionalist Response	71
2.4.3 Biting the Bullet	74
2.4.4 "Bare Truth" Revisited	75
2.5 Conclusion	77
Chapter 3: Extrinsic Dispositions	78
3.1 Marks of Dispositionality	80
3.2 Intrinsic, Extrinsic and Relationally Specified Properties	82
3.3 Dispositions Are Relationally Specified	83
3.4 A Possible Qualification	87
3.5 The Case for Intrinsicness	88
3.5.1 The Conceptual Argument	88
3.5.2 Armstrong's Argument	90
3.6 Examples of Extrinsic Dispositions	93
3.7 Conclusion	106

Chapter 0: Introduction

When someone has a disposition, he or she is prone to act in certain ways in certain circumstances. A cowardly person is disposed to flee from danger, for example. A sociable person is disposed to seek the company of others. Physical objects can also have dispositions. Fragile glasses are disposed to break when struck. Elastic bands are disposed to stretch when pulled. Talk about the dispositions of people and things is commonplace.

Scientific language, too, is rife with disposition-talk. A substance is volatile or reactive if it is disposed to enter into reactions with other substances. A material is conductive if it is disposed to transmit an electric charge. An element is unstable if it is disposed to decompose. An organism is fit for a certain environment if it is disposed to survive in that environment.

It is widely acknowledged that disposition-talk is ubiquitous. However, it is a matter of some controversy how this talk is to be understood. In this thesis, I challenge some commonly held views about dispositions. I claim the following:

1. Dispositions can be causally efficacious.
2. Dispositions can be bare.
3. Dispositions can be extrinsic.

The three main chapters are devoted to explaining and defending these claims. (There, I will explain what I mean by "causally efficacious," "bare," and "extrinsic.") This introductory chapter is devoted to introducing basic terminology and sketching the framework within which debates over these issues occur.

0.1. Dispositions

First of all, a disposition is a property, or a quality that things can have, like roundness or tallness. A disposition has a characteristic *manifestation*. For example, the

characteristic manifestation of fragility is shattering, and the characteristic manifestation of cowardliness is avoidance of danger. The manifestation need not occur for the object to have the disposition. A glass can be fragile even if it never shatters. The expression "manifestation" is sometimes used to refer to a particular event, as when I say that a shattering that happened at a particular place and time was the manifestation of a particular glass's fragility. However, "manifestation" is also used to refer to a type of event. When I say that the manifestation of fragility is shattering, I am not referring to any particular shattering, but am making a general claim about the events which are manifestations of instances of fragility -- they are shatterings.

A manifestation of a disposition occurs (at least typically) when the object with the disposition is in certain circumstances. The fragile glass shatters *when it is struck*. I call the circumstances in which the manifestation occurs (oddly enough) the *circumstances of manifestation*. The circumstances of manifestation include not only the salient "trigger" for the manifestation (the striking), but also the necessary background conditions, such as ambient temperature and gravitational forces. In some cases, the manifestation will occur at the location of the disposed object itself. In the case of the fragile glass, the shattering occurs where the glass is. However, in other cases, the locus of manifestation may be elsewhere. For example, if something is provocative, the manifestation will occur in the thing provoked.

Given that a disposition is associated with a manifestation, and with circumstances which trigger the occurrence of this manifestation, there is a natural association between a statement attributing a disposition to a thing and a certain counterfactual statement. For example, the statement "This glass is fragile" bears some important relation to the statement "If this glass were struck, it would shatter." Some say that this relation is one of a priori equivalence. According to this view, to say that the glass is fragile is just to say

that if it were struck, it would shatter. However, it is widely acknowledged that the relationship between these statements is not so simple.¹ If the fragile glass were struck lightly with a feather, or if it had internal supports, it would not break when struck, and so the associated counterfactual would be false. If a counterfactual analysis of disposition statements is going to work, it is going to have to be somewhat more complex. I am not going to develop or defend such an account here; however, it is important to note that there is some sort of conceptual connection between dispositions and counterfactuals. An attribution of a disposition to an object licenses inferences about what will happen in various circumstances. These inferences may be defeasible, but the ability to make these inferences is what makes disposition-talk so useful, if not indispensable.

When an object is disposed to behave in a certain way in certain circumstances, it is thought that there is something about the object "in virtue of which" this is so. Consider the case of fragility yet again. The fragile glass is disposed to shatter when struck. This is due to some feature of the glass. The silicon bonding in the glass is such that, if an excessive strain is placed on these bonds due to warpage of the glass, some of the bonds will break, starting a chain reaction of bonds breaking. Hence, the glass shatters when struck. This type of molecular bonding is the *causal basis* of the glass's fragility. A disposition's causal basis is a property which is causally efficacious for the disposition's manifestation. (I will have much more to say about what it is for a property to be causally efficacious in Chapter 1.)

An important thing to note about dispositions and causal bases is that they do not typically have a one-to-one correspondence. Crystal wine glasses and egg shells are both fragile, but have very different structures and constituents. Presumably, these things have

¹C. B. Martin, "Dispositions and Conditionals," *Philosophical Quarterly*, 44: (1994), 1-8; David Lewis, "Finkish Dispositions," *Philosophical Quarterly*, 47: (1997), 143-58; R. K. Shope, "The Conditional Fallacy in Contemporary Philosophy," *Journal of Philosophy*, 75: (1978) 397-413.

different properties in virtue of which they are prone to break when struck. That is to say that the causal basis of fragility in a wine glass is a different property than the causal basis of fragility in an egg shell. For an example of the reverse, consider electrical conductivity and thermal conductivity, two different dispositions which have the same causal basis in metals.

0.2. Dispositional versus Categorical

Let's switch from talking about properties to talking about predicates and property terms for a moment. Dispositional predicates are expressions such as "is fragile," that have certain conceptual associations for the speakers of the language, with things like shattering and striking. Dispositional predicates are associated with manifestations, circumstances, and counterfactual statements.

A predicate which does not have these features is called a non-dispositional, or *categorical* predicate. As paradigm examples of categorical predicates, philosophers often offer shape predicates. To say something is square is not to say what it would do in particular circumstances. While it is true that a square thing would behave in certain ways in certain circumstances in virtue of being square, "squareness" does not have the sort of conceptual association with a certain type of event that "fragility" has to breaking. That is to say, squareness has no characteristic manifestation.² Consequently, categorical predicates do not have the relevant relation to particular counterfactuals. As Elizabeth Prior notes, "dispositional ascription sentences possess a relationship to certain subjunctive conditionals not possessed by categorical ascription sentences."³

²This can certainly be denied, as we shall see shortly.

³Elizabeth Prior, *Dispositions* (Aberdeen: Aberdeen University Press, 1985), 62.

Some philosophers dispute the claim that we can make out a clear distinction between dispositional and categorical predicates. According to Goodman, for example, "almost every predicate commonly thought of as describing a lasting objective characteristic of a thing is as much a dispositional predicate as any other."⁴ According to some, notably Philip Pettit, all concepts are response-dependent -- that is to say, all concepts are defined via reference to the psychological responses of suitably situated subjects.⁵ According to this account, for example, something is red if and only if it would look red to suitable subjects under suitable conditions. It seems that, on this view, every predicate has a conceptual association with a manifestation and particular counterfactuals. In other words, on this view, all predicates are dispositional predicates.

However, most philosophers recognize a distinction between dispositional and categorical predicates. Hence, philosophers like Elizabeth Prior insist that there is a conceptual association between dispositional predicates and counterfactuals that is lacking in the case of categorical predicates. Obviously, this characterization is rough, and the nature of this conceptual association is left vague. One may resist going as far as Goodman or Pettit, clearly there are some borderline cases. For example, "being negatively charged" seems to describe a lasting, objective characteristic of a thing. However, the expression does have strong conceptual associations with characteristic manifestations, such as repelling negatively charged particles, and attracting positively charged ones. While it may not be clear where to draw the dispositional/categorical line, perhaps we can identify the ends of the spectrum.

⁴Nelson Goodman, *Fact, Fiction, and Forecast*, 4th ed. (Cambridge, MA: Harvard University Press, 1983), 41.

⁵Philip Pettit, "Realism and Response Dependence," *Mind*, 100 (1990), 587-626. See also Mark Johnston, "Dispositional Theories of Value," *Proceedings of the Aristotelian Society*, Supplementary Volume 63 (1989), 85-130.

Some philosophers claim that the dispositional/categorical distinction ends there, as merely a distinction among predicates, which does not correspond to any interesting metaphysical distinction among properties. Shoemaker, for example, says "I think that the term 'dispositional' is best employed as a predicate of predicates, not of properties."⁶ According to Shoemaker, what determines the identity of a property "is its potential for contributing to the causal powers of the things that have it."⁷ Despite Shoemaker's remarks, one might want to describe his view by saying that all properties are dispositional properties. At the other extreme, David Armstrong claims that all properties are categorical properties.⁸ He acknowledges that some properties are picked out by disposition terms, but claims that such terms simply provide us with a useful way of speaking of categorical properties. As Armstrong puts it, the dispositional/categorical distinction is a "verbal distinction that cuts no ontological ice."⁹

Clearly, Armstrong thinks that the distinction between properties makes sense -- he just thinks that one of the categories is empty, and so if a non-trivial distinction is wanted, then it must apply to predicates. Although Shoemaker wants to reserve 'dispositional' as a predicate of predicates, he can agree with Armstrong that the distinction between properties makes sense -- he just disagrees about which category is empty. These views can be contrasted with that of a more extreme theorist who thinks that the distinction between dispositional properties and categorical properties is unintelligible, or that a purely linguistic distinction is being confused with a distinction between properties. For

⁶Sydney Shoemaker, "Causality and Properties," in Peter Van Inwagen, ed., *Time and Cause: Essays Presented to Richard Taylor* (Dordrecht: Reidel, 1980), 109-136.

⁷ *Ibid.*, 211.

⁸ Tim Crane, ed., *Dispositions: A Debate* (New York: Routledge, 1996), 16.

⁹ D. M. Armstrong, *Belief, Truth and Knowledge* (Cambridge: Cambridge University Press, 1973), 15.

example, some property terms such as "yellow" are polysyllabic, others such as "red" are monosyllabic. But of course, there is no interesting distinction between the properties to which these terms refer.

There is widespread agreement, at least, that the dispositional/categorical distinction is not like that, and so my arguments proceed on the assumption that the issues are metaphysical, not merely linguistic. Categorical *properties*, then, are properties which do not have the relevant associations with manifestations, circumstances, and counterfactuals. Along these lines, Stephen Yablo offers as an intuitive characterization, "a property is categorical just in case a thing's having it is independent of what goes on in nonactual worlds."¹⁰

Naturally, there is room for dispute as to which properties are categorical and which are dispositional. Hugh Mellor argues that even shape properties are dispositions. For example, he says that triangularity is the property of being disposed to be counted as three-angled.¹¹ Similarly, Goodman says "a cubical object is one capable of fitting try squares and measuring instruments in certain ways."¹² I am not going to dispute these claims here; I am just trying to explain the distinction between dispositional and categorical properties, without taking a stand as to which (if any) specific properties fall into each category. Furthermore, we need not suppose that this distinction among properties is exhaustive. For example, mathematical properties, and some disjunctive properties, such as being fragile or square, might be neither categorical nor dispositional. According to C. B. Martin, properties admit of degrees of dispositionality.¹³ While

¹⁰ Stephen Yablo, "Identity, Essence, and Indiscernibility," *The Journal of Philosophy* 84:6 (June 1987), 306. Yablo goes on to argue that this characterization is inadequate.

¹¹D. H. Mellor, "In Defense of Dispositions," *Philosophical Review*, 83 (1974), 171.

¹²*Fact, Fiction, and Forecast*, 41.

¹³C. B. Martin, "Power for Realists," in J. Heil (ed.), *Cause, Mind and Reality: Essays*

fragility is highly dispositional, and squareness is highly categorical, negative charge, and perhaps certain disjunctive and conjunctive properties, fall somewhere in the middle.

0.3. Properties

Now that I've explained some of the general characteristics of dispositions and distinguished them from categorical properties, I am going to "back up" and say a bit more about the ontological category in which I place dispositions.

Properties are ascribed to things by predicates, such as "is red". Yet, it is unlikely that there is a one-to-one correspondence between properties and predicates. There can be some predicates that have no corresponding property ("being non-self-instantiating", for example), and there are properties that do not have any corresponding predicate, i.e., there are properties in the world that we have no terms for. I will consider two competing views of properties, as represented by David Armstrong and David Lewis.¹⁴

On Lewis's view, a property is any class or set of actual or possible entities. One property consists of my pencil, Bill Clinton, and some possible apple in possible world W. Another property consists of all the red things in the actual world, and all the possible red things in all possible worlds. On this view, properties are extremely abundant. However, for any class of entities, there is just one property; triangularity and trilaterality turn out of be the same property on this view. So, this account is inclusive, but also coarse-grained.

Entities can be grouped in a large variety of ways. Some of these classes will seem like arbitrary groupings. Other classes will group entities that are similar in some natural respect. Given the abundance of sets of possibilia, the majority of properties will be

Honoring C. B. Martin, (Dordrecht: Kluwer, 1989), 175-86.

¹⁴Armstrong, *A Theory of Universals*, (Cambridge: Cambridge University Press, 1978); Lewis, "New Work for a Theory of Universals," *Australian Journal of Philosophy*, 61 (1983), 343-377.

arbitrary groupings of entities, and the classes with members that are similar in some natural respect will be an elite minority. Classes of similar entities seem to be importantly different from classes of dissimilar entities, and we may want a way of distinguishing these different types of properties. One way that Lewis suggests is to adopt a primitive predicate "--is a natural class", which will apply to some properties, such as, perhaps, the class of red things, but not to others. Implicit in this view is a notion of "natural kinds": natural properties correspond to real distinctions in nature. Arbitrary groupings of dissimilar entities, such as the set {my pencil, Bill Clinton, an apple} are non-natural, or disjunctive properties.

Armstrong presents a different account of properties. While Armstrong does not believe in possible worlds, he would grant there are a great many classes of entities. But Armstrong does not call such classes "properties." Furthermore, Armstrong would agree that some of these classes are special in that their members are similar to one another, i.e., have something in common. However, Armstrong would say that what members of natural classes have in common is that they each instantiate some *universal*, and it is universals for which Armstrong reserves the expression "property". Just as Lewis's "natural properties" correspond to natural kinds, Armstrong's universals correspond to the true natural kinds that would be discovered by a perfect science.

The substantial difference between Lewis and Armstrong (beyond belief in the existence of non-actual entities) is how to characterize this elite set of classes of entities. According to Armstrong, in order to characterize this set we need to bring in the notion of universals. According to Lewis, we can characterize the "elite" properties via the primitive predicate "is a natural class." For the purposes of this thesis, I can remain neutral as to how to characterize this elite set. I will, however, follow Lewis and use the expression "property" in an inclusive and coarse-grained sense: For any collection of possible things, no matter how diverse, there is just one property had by all and only those things. (I'll call this the extensional account of properties.)

In addition to natural properties, another type of property that is important to this discussion is that of a second-order property. Many philosophers think, and much of the subsequent discussion proceeds on the assumption that dispositions are second-order properties. A second-order property is the property of having some property or other which meets some specification.¹⁵ For dispositions, the specification will be a causal one. The having of a first-order property, by contrast, does not necessarily involve the having of another property. For example, suppose that the color properties, such as being red, and being blue, are first-order properties. A second-order property, then, would be being colored -- a property that a thing has just in case it has some color property or other.

However, a so-called second-order property might actually be third-order, or higher-order. Perhaps a color property such as being green isn't a first-order property at all, but a property that an object has just in case it has some surface-reflectance property or other. Then, the property of being colored would be a third (or higher) order property, that is, the property of having some property of having some property. To avoid making assumptions about the basicness of any particular property, the expressions "first-order property" and "second-order property" can be thought of as relative terms.

Notice, however, that on the extensional account of properties, the same property may be characterized in first or second-order terms. "The property of having property P" would seem to pick out a second-order property. However, it would necessarily have the same extension as the apparently first-order "property P." So "P" and "the property of having P" are but two names for the same property. In general "the property of having a

¹⁵The expression "second-order property" is used in other ways. Armstrong uses the expression to refer to properties which properties have (*A Theory of Universals*). Kim suggests that second-order properties are those had by second-order entities (entities mereologically composed of first-order, fundamental entities). ("The Nonreductivist's Troubles with Mental Causation," in *Supervenience and Mind* (Cambridge, UK: Cambridge University Press, 1993), 337.)

property meeting specification H" will refer to the same property which consists of a disjunction of all the properties meeting specification H. "Being green or red or blue or yellow, etc." and "having some color property or other" have the same extension, and hence these expressions pick out the same property.

On this account, it might seem as though there are only first and second order *characterizations* of properties which have little, if any, ontological significance. However, this account can do justice to the idea of a second-order property. Certain properties are canonically characterized via second-order characterizations. This would be so if the most salient point of resemblance between members of a class is the having of properties which meet a certain specification. So, suppose that properties P, Q, and R are the only properties which meet specification H. It may be that the most natural or salient aspect of resemblance between members of the class of objects that have P, or Q, or R is that they have some property or other that meets specification H. So, rather than characterizing the property which corresponds to this class as "the property of having P or Q or R," it is more appropriate, and indicative of the nature of the property, to characterize this property as "the second-order property of having some property or other which meets specification H." Even though this property can be characterized as "having P or Q or R," the canonical characterization of this property is second-order, and so the property itself may be called a second-order property.

One important type of second-order property is a functional property. A system has a given functional property if it has a property that fulfills a certain causal role, mediating between inputs and outputs, and internal states of the system of which it is a property. As Ned Block put it:

Functional properties are properties that consist in the having of some properties or other (say non-functional properties) that have certain causal relations to one another and to inputs and outputs. ¹⁶

¹⁶Ned Block, "Can the Mind Change the World?", in *Meaning and Method: Essays in*

The non-functional, or lower-order properties *realize* the functional property. If different lower-order properties can realize the functional property, then the functional property is *multiply realizable*.

As an example of a simple system which has functional properties, consider a Coca-Cola machine. Suppose that one can purchase a coke for 10¢ from this machine, which only accepts nickels and dimes. The machine can be in one of two states, described in the table:¹⁷

	S ₁	S ₂
nickel input	emit no output and go to S ₂	emit a coke and go to S ₁
dime input	emit a coke and stay in S ₁	emit a coke and a nickel and go to S ₁

At any given time, the machine M can instantiate one of two functional properties, being in S₁, or being in S₂. To say that M has the functional property of being in S₁ is to say that it has some property or set of properties, presumably physical, mechanical properties, which are causally efficacious for M emitting nothing and going into S₂ when M receives a nickel input, or emitting a coke and staying in S₁ when M receives a dime input. S₁ is multiply realizable, in that it could be realized by metal gears and springs, or by wooden levers and pulleys, for example.

A disposition can be thought of as a simple sort of functional property. (Or, a functional property can be thought of as a complex type of disposition.) In the case of dispositional properties, the inputs are the circumstances of manifestation, and the outputs are the manifestation of the disposition. For example, an object has the disposition

Honor of Hilary Putnam. ed. George Boolos. (Cambridge: Cambridge University Press, 1990), 155.

¹⁷ Ned Block, "Troubles with Functionalism," in Rosenthal, David M. ed., *The Nature of Mind*, (New York: Oxford University Press, 1991).

fragility if, given a strong enough impact as "input," the "output" shattering will result. An object has a dispositional property if it has some property or other (i.e., some causal basis) which is causally efficacious for the manifestation of the disposition under certain circumstances. Like the mechanical properties of the coke machine, these bases can be said to realize the disposition. And since a single disposition can have various bases, dispositions can also be multiply realizable.

Chapter 1: A Defense of the Causal Efficacy of Dispositions

Disposition terms, such as 'cowardice,' 'fragility,' and 'reactivity' often appear in explanations. We explain why a man ran away by saying that he was cowardly, and we sometimes explain why something broke by saying it was fragile. Dispositional properties like instability, reactivity, and conductivity figure in scientific explanations of certain phenomena. And these look like causal explanations -- they seem to provide us with some information about the causal history of those events.

In recent years, several arguments have been put forward to the effect that dispositions are causally inert. These arguments are due to philosophers such as Ned Block, Jaegwon Kim, Elizabeth Prior, Robert Pargetter, and Frank Jackson. According to this view, the glass's fragility was not responsible for its breaking; the man's cowardice was causally impotent as he fled. This view would call many of the explanations we tend to give into question. By employing a disposition in an explanation, we might have thought we were explaining an event by reference to a property of the cause that was in some sense responsible for the occurrence of the event. However, if dispositions are causally inert, we are explaining the event in some other way, or not really explaining it at all.

The view that dispositions are causally inert suggests that something is amiss with many scientific explanations as well. If properties like conductivity and volatility are causally inert, it is not clear how appealing to them provides us with information about why certain phenomena occur. This is especially problematic if one thinks, as some do, that the fundamental properties that scientists attribute to the ultimate constituents of matter -- things like force, mass, charge, impenetrability -- are dispositional. If, as Simon Blackburn says, "science finds only dispositional properties all the way down,"¹⁸ and

¹⁸Simon Blackburn, "Filling in Space," *Analysis* 50 (1990), 255.

dispositions are causally inert, it would seem that science does not provide us with real causal explanations.

The view that dispositions are causally inert has further worrisome implications if one thinks, as many philosophers do, that mental properties are dispositional. According to a leading account of mental properties, they are functional properties. We can think of functional properties as complex sorts of dispositions. According to functionalism, to be in a mental state is to have a disposition to go into another mental state and to display certain behavior, given the circumstances of preceding mental states and certain inputs. So, if mental properties are dispositional, the view that dispositions are causally inert poses a problem for mental causation; it suggests a form of epiphenomenalism -- the view that mental properties make no difference to what a body does. It is natural to think that the fact that I believe and desire certain things has a great deal to do with my body moving in certain ways. It would take some fairly powerful arguments to cast these beliefs into doubt.

Two commonly offered arguments against the causal efficacy of dispositions are what I call the "Analyticity Argument" and the "No Work argument." In this chapter, I defend the causal efficacy of dispositions by disputing these arguments.

Recall that a disposition is a property, which can be picked out via a second-order characterization, involving a causal specification. The causal specification will involve reference to a manifestation, and the conditions of manifestation. Typically, dispositions are multiply realizable. Also, it is assumed that a disposition has a causal basis. The causal basis is a property which is causally efficacious for the manifestation, given the circumstances. So, for present purposes, we can assume the following account of dispositions:

To have *disposition* D to give manifestation M in circumstances C is to have some property P which is a *causal basis* for giving M in C.
and
To have a *causal basis* for giving M in C is to have a property which is causally efficacious for M in C.¹⁹

Note that it does not follow from this account that a thing can only have one causal basis per disposition. If an object had two properties P and Q, and either P or Q would be efficacious for M in C, then both P and Q would be causal bases.

1.1. Causal Relevance and Causal Efficacy

According to some philosophers, dispositions are "inert", "impotent", "causally irrelevant", and "inefficacious". Presumably, it is thought that there are other properties that do not suffer this inadequacy. Presumably, it is thought that categorical properties have some relation to events which dispositions lack. In this section, I make three attempts to specify this relation that purportedly fails to hold between dispositions and their manifestations. Positing a particular reductive analysis for this relation is not crucial, for the arguments I will go on to consider (or slight variations thereof) could be run under various interpretations of this relation. However, trying to get more clear about what this relation is supposed to be will help us to better understand the claim at issue.

We begin by noting that an event will instantiate a number of properties, as will objects which are involved in that event, and the general environment in which the event takes place. Events can be sudden, involve sweaty cab drivers, and take place in 90 degree heat, for example. I am going to call all the properties of an event, together with

¹⁹This definition would seem to rule out the possibility of bare dispositions -- dispositions with no distinct causal bases -- a priori. However, a bare disposition could technically satisfy this definition, if the base property were just the disposition itself. If bare dispositions are possible, as I argue in chapter 2, then the "No Work" argument against their causal efficacy is a non-starter. For the purposes of the present discussion, I leave this possibility aside.

the properties of the objects involved in that event the "event-properties." Given a set of event-properties, some of those properties will be related in important ways to an effect of that event, and other properties will not. Presumably, an event caused what it did "in virtue of" certain properties. When we are looking at causal relations, we may be interested not only in which events cause certain effects, but which properties are important for that causal relation. This is especially important if we want to make general causal claims, or formulate causal laws.

So, there is an important relation between some event-properties, and a given effect that event has. Intuitively, some of the event-properties will be *causally relevant* to a given effect, and others will not. The fact that an event was sudden and unexpected might be relevant to its surprising me, but the duration of that event might be irrelevant to that effect. If a ball was thrown at a window and broke the glass pane, the mass and velocity of the ball would be relevant to the event of the glass breaking. However, the color of the ball, and its belonging to little Johnny, are properties of the ball, but presumably they are not causally relevant properties.

The relevance of a property to a particular event seems to be a matter of the role that that property can play in explaining that event. So, we can understand causal relevance as follows:

A property P is *causally relevant* to an event e iff
P is apt for being mentioned in a causal explanation of why e occurred.

However, it seems unproblematic to allow that dispositions are causally relevant, in this sense. Earlier, we noted that they are sometimes apt for being mentioned in an explanation of why an event occurred. If someone is asked why something broke, in some contexts it would seem perfectly appropriate to mention the fact that it was the kind of thing that was prone to break --that it was fragile. Properties like cowardliness, shyness, electrical conductivity, and volatility, seem well suited for figuring in explanations. It is doubtful that one who argues that dispositions are causally inert is claiming that they never

figure in satisfying responses to "why-questions." Moliere notwithstanding, dispositions do figure in apparently adequate explanations.

Despite the apparent explanatory relevance of dispositions, some philosophers still want to maintain that they are causally inert. Frank Jackson, for example, says that a property:

... might be causally relevant in some wide sense of 'causally relevant' without being causally potent or efficacious. It is, for instance, clear that being fragile is causally relevant to a glass's breaking. Otherwise it would be vacuous to cite a glass's being fragile in explanation of its breaking. And it isn't -- it tells you that the breaking was due to the glass's nature and not, say to a peculiarity in the way it was struck. But if our earlier argument is right, fragility does not actually cause the breaking.²⁰

We will get to Jackson's "earlier" argument shortly, but the point to note now is that the notion of explanatory relevance is thought to be too broad and permissive to capture the relation we are interested in. The perceived adequacy of an explanation is subject to pragmatic or epistemic considerations. What is appropriate to mention in an explanation is sensitive to context, and relative to circumstances, such as the epistemic position and interests of an audience seeking the explanation. For example, suppose that we know that Johnny had a white baseball and yellow Nerf ball. Someone might explain the window breaking by mentioning that it was the white ball that was thrown at the window. So, the whiteness of the ball could be mentioned in a causal explanation of why the window broke. It would meet the criteria for causal relevance. However, intuitively, something more is desired.

Following Jackson, we can call this "something more" *causal efficacy*, understanding it as a mind-independent, metaphysical relation holding between a property

²⁰Frank Jackson, "Essentialism, Mental Properties, and Causation," *Proceedings of the Aristotelian Society*, 95 (1994-1995), 262. See also Frank Jackson and Philip Pettit, "Program Explanation: A General Perspective," *Analysis*, 50 (March 1990), 107-117.

and an event. Appeal to causal efficacy is an attempt to narrow, or make more precise, the class of properties that play a central role in causal connections between events. While we may be inclined to admit that the instantiation of a property was somehow *relevant* to the occurrence of an event, we might still wish to deny the causal efficacy of that property. Being causally efficacious is one way for a property to be relevant to an event, but perhaps there are others.

If it is granted that dispositions are causally relevant, in some broad sense, isn't that enough? As long as they are relevant, should we care if they satisfy this further condition, being causally efficacious, whatever that may be? Well, perhaps some people will be content with causal relevance, and not be concerned about causal efficacy. However, if we think of the array of properties which seem to be dispositional, it would seem odd, if not disconcerting, to think that they had mere explanatory power, and not some sort of causal power. It seems we want to say more about properties like cowardice, electrical conductivity, impenetrability, or desiring water, than that they are apt to figure in explanations. It seems we want to say that they are apt for figuring in explanations in virtue of making a difference to what happens. It is this sense of "making a difference to what happens" that the notion of causal efficacy is meant to capture.

That's the intuitive idea, but stating necessary and sufficient conditions for a property to be causally efficacious is no easy task, and it is a vexed question in the literature. One approach is to say that a causally efficacious property is an event-property which is counterfactually necessary for the event to cause the effect in question. Intuitively, if an event would have caused what it did even without the instantiation of a particular property, then that property is not a causally efficacious one.²¹ According to this "counterfactual dependence" approach:

²¹Would an event *e* have been the same event, had one of its properties not been instantiated? If we assume a relatively "robust" conception of events, in Lewis' sense ("Events," *Philosophical Papers*, Volume 2. (Oxford: Oxford University Press, 1983)),

If a property P is *causally efficacious* for an event e, then if P had not been instantiated, e would not have occurred.

This looks plausible, but this cannot be what the deniers of the causal efficacy of dispositions have in mind, for it is easy to see that dispositions like solubility and fragility pass this test. For example, if this tablet hadn't been water-soluble, it would not have dissolved when I put it in water. If this glass hadn't been fragile, it would not have broken when I dropped it. It seems unproblematic to allow that a disposition's manifestation counterfactually depends on the prior instantiation of that disposition.

Perhaps a defender of the causal efficacy of dispositions will want to say that this *is* the right test for causal efficacy, and since dispositions pass it, this shows they *are* causally efficacious. However, there are cases where a property will pass this test where we are disinclined to say that the property is causally efficacious. Say two properties, P and Q are always co-instantiated as a matter of law. Any counterfactual test that P will pass, Q will pass as well. However, we might want to allow that P is causally efficacious in some instances where Q is not. Counterfactual dependence is not a sufficient condition for causal efficacy.²²

Another approach, which seems to me more promising, is to say that causally efficacious properties are members of a set of properties which are minimally sufficient for

then the loss of one property would not necessarily "destroy" the event. However, if a property is an essential property of some event, then if that property hadn't been instantiated, that event wouldn't have occurred, though a somewhat different one may have.

²²Ned Block gives an example which illustrates this point. A metal rod heats up and ignites a bomb. The composition of the rod is such that, whenever its thermal conductivity is increased, its electrical conductivity is increased as well. So, it is true that, had the electrical conductivity not increased, the bomb would not have exploded. However, intuitively, it is the increase of thermal conductivity, and not the increase of electrical conductivity, that was causally efficacious for the explosion ("Can the Mind Change the World?" 147).

the occurrence of an effect. Let me explain. Take a time slice of the causal history of some event. Out of the set of all the properties instantiated in that time slice, there will be some subset of causally efficacious properties. These properties are such that, when instantiated together in this fashion, the given effect necessarily occurs (or if the world is indeterministic, the effect has a certain probability of occurring). This set of causally efficacious properties is to be minimal, in that no subset of it would be sufficient for the occurrence of the effect. So, according to the "minimal sufficiency" approach:

If a property P is *causally efficacious* for an event e, then
P is a member of a set of event-properties S which is:
i) sufficient, given the laws, for e; and
ii) such that no proper subset of S has this feature.

Note that there is nothing in this definition which guarantees that there will be only one minimally sufficient set. This will be important later.

This account is not without its difficulties. Like the counterfactual account, the minimal sufficiency account fails as a sufficient condition for causal efficacy. Say that object o has the property of being such that if it has P, it has Q, and furthermore, o has P. If Q is causally efficacious, then being such that (if P, then Q) and having P would be part of a minimally sufficient set, but for all that, P might not be causally efficacious. For example, suppose that a refrigerator is such that, if it makes a humming sound, it lowers the temperature, and that the refrigerator makes a humming sound. On this account, making a humming sound would count as causally efficacious for keeping food cold.

However, I do think that this account is a promising start. The advantage that this account has over the others is that it offers a relatively simple way to place further constraints on what qualifies as a causally efficacious property. This can be done by placing restrictions on the set of event-properties S. We could, for example, add that a property has to be natural to be included in S.

It seems trivial that the relations of explanatory relevance and counterfactual dependence hold between a disposition and its manifestation. Is it trivial that a

disposition will be part of a set of properties that is minimally sufficient for its manifestation? As stated, yes. A set of event-properties, including a disposition and certain properties of the circumstances of manifestation, would be minimally sufficient for the manifestation of that disposition. (If that's not obvious, it will hopefully become clear later.) However, it is acknowledged that further constraints need to be placed on the set of event-properties S. What is not trivial is what that constraint should be, and if it is one dispositions can meet. Clearly, part of the debate over whether dispositions are causally efficacious is a debate over just what causal efficacy is.

I think it is reasonable to proceed on the assumption that those who deny that dispositions are causally efficacious have something like (appropriately constrained) minimal sufficiency in mind. So, what are the reasons for thinking that dispositions are not causally efficacious, in this sense? One argument for that conclusion is the "the Analyticity Argument," according to which there is an analytic or necessary connection between a disposition and its manifestation, and this goes to show that there is no causal connection. I argue that it shows no such thing. I then consider what I think is the greater challenge to the causal efficacy of dispositions -- the "No Work" argument.

1.2. The Analyticity Argument

In light of the preceding remarks, our incomplete, working, account of causal efficacy can be put as follows as follows:

- A property P is causally efficacious for an event e iff
P is a member of a set of event-properties S which is:
- i) sufficient, given the laws, for e; and
 - ii) such that no proper subset of S has this feature; and
 - iii) such that ...?

An adequate analysis will add further constraints, and thus provide necessary and sufficient conditions for a property to be causally efficacious. One suggestion is that the constraint

needs to insure that the causal laws are doing some work in the analysis. So, the further constraint that is needed might be the following: S must not entail e alone, without the laws.

It seems that dispositions would fail to meet this requirement. Suppose a fragile glass breaks, and S is a set of properties which was minimally sufficient for the breaking. If S includes the properties of the circumstances of manifestation, including, say, being a striking, as well as fragility, then these properties alone, without the laws of nature, would entail that a breaking occurred, other things being equal. If something has a disposition, it has some property that would cause the manifestation in the circumstances. So, it follows just from the fact that a disposed object is in the circumstances of manifestation that the manifestation occurs, other things being equal. (It's not clear that we could get rid of the *ceteris paribus* clause even in a case in which the laws were needed, so the fact that it only follows "other things being equal" doesn't help.)

Is this a satisfactory emendation to the account of causal efficacy? I will argue that it is not. Claiming that this constraint needs to be added to the account of causal efficacy amounts to presenting an "Analyticity Argument" against the causal efficacy of dispositions. According to the Analyticity Argument, there is an analytic relation between a disposition and a manifestation, and this precludes any causal connection between them. The Analyticity Argument begins by noting that any adequate definition of "fragility" is going to make some mention of breaking or shattering. This goes for disposition terms generally --their definitions will make reference to their characteristic manifestations. So, there is some sort of definitional, conceptual or logical connection between a disposition term and an event-type --between "fragility" and breaking, for example. The statement "Fragile objects tend to break when struck" is in some sense analytic.

We've seen that a disposition can be relevant to its manifestation. For instance, the glass's fragility was relevant to its breaking. Now, if a property is causally relevant to a certain event, perhaps that gives us some reason for thinking that that property is causally

efficacious for that event. However, this reasoning is defeasible, especially if we have some independent explanation of the relevance of the property. And that is exactly what we have in the case of the disposition and its manifestation. Our sense that the disposition is relevant to the manifestation can be explained by the analytic connection. We need some further reason to suppose a causal connection holds. This is what Block suggests when he writes:

The fact that dormitivity is sufficient for sleep is perfectly intelligible in terms of this logical relation. What reason is there to suppose that there must *also* be a nomological relation between dormitivity and sleep?²³

The Analyticity Argument, as Block expresses it, can be put as follows:

1. A disposition is (explanatorily) relevant to its manifestation.
 2. This relevance can be explained by the fact that there is an analytic connection between the disposition and its manifestation.
- Therefore,
3. There is no reason to suppose that the disposition is also causally efficacious with respect to its manifestation.

So, even though we can cite a disposition in an adequate explanation of an event, it is possible that the disposition is causally inefficacious with respect to that event. While these considerations may cast some doubt on the causal efficacy of dispositions, to say that it *is possible*, given our language and explanatory practices, that a disposition is causally inert is not to show that *is* causally inert. Furthermore, pointing out *one* reason that something is relevant does not show that that is the *only* reason that it is relevant.

The fact that an event is picked out in terms of its effect does not show that it is causally inert. Consider the statement "The cause of e causes e." The statement is analytic, and there is a logical connection between "the cause of e" and "e." "The cause of

²³Block, "Can the Mind Change the World?" 157.

e occurred" entails that e occurred, and does so without reference to the laws of nature. However, these observations lend no support to the idea that the cause of e didn't cause e. How events are described, what names they are given, should neither determine nor preclude causal connections between them. Real world causal connections hold independently of our descriptions of them. Perhaps our descriptions of events actually track causal connections. Examples of this are familiar: Sunburn is caused by excessive exposure to sunlight; lethal injections and fatal accidents cause death. The existence of conceptual connections between our descriptions of events might actually *support* the idea that these events are causally connected.

The same goes for causally relevant properties. In fact, we can run an analogous argument in terms of properties, considering the claim "The property that was causally efficacious for e was causally efficacious for e." "The property that was causally efficacious for e was instantiated" would entail that e occurred, without reference to the laws of nature. So, the suggested constraint on the set of event-properties S is too stringent.

1.2.1. Hume's Principle

While the above statement of the analyticity argument does not present a very serious challenge to the causal efficacy of dispositions, one might still be troubled by the thought that the connection between a disposition and its manifestation is too "tight" for there to be a causal relation as well. Frank Jackson presents an argument along these lines, appealing to the idea that there is a metaphysically necessary connection between a disposition and its manifestation, and this is incompatible with there being a causal connection. According to Jackson, saying fragility causes glass to break is to violate Hume's Principle about the contingency of causal connections.²⁴

²⁴"Essentialism, Mental Properties, and Causation," 257.

Jackson acknowledges that "there are necessarily true statements asserting causal connections," however, this is compatible with Hume's Principle, according to which what a state causes, or what it would cause, is not an essential property of that state.²⁵

Consider the statement "The lethal injection caused a death." This statement is necessarily true -- in every world in which it is correct to call the injection a "lethal injection," it caused a death. However, an injection of the particular substance that was used does not necessarily cause death. There are worlds in which that substance is harmless. If we take the expression "the lethal injection" to refer rigidly to an injection of a particular substance (rather than non-rigidly to an injection of some death-causing substance or other) then there will be worlds in which "The lethal injection caused death" will be false.

So, according to Hume's Principle, a thing's causal powers are accidental properties, which depend on what world it's in, and which laws obtain. This suggests that if two states have a necessary connection, they are not candidates for being causally connected. According to Jackson, to allow that fragility causes breaking upon dropping--

would be to allow that there are properties that have causal powers essentially: in every world the property of having the property or properties responsible for breaking on dropping in that world is possessed only by objects which are such that were they dropped they would break. There is no way that the second-order property can be instantiated without the relevant causal power being instantiated. So, if we are to respect Hume's insight, we must deny that fragility itself does the causing of the breaking...²⁶

Jackson's argument can be put as follows:

²⁵So, it is a misnomer to call Jackson's argument an "analyticity argument," but old habits die hard.

²⁶Jackson, "Essentialism, Mental Properties, and Causation," 257.

1. Hume's Principle: If P is causally efficacious with respect to events of kind K, it is only contingently so (i.e. there is a world in which P is not causally efficacious with respect to K-events).

2. If a disposition D were causally efficacious with respect to events of kind M (D's characteristic manifestation), it would be necessarily so.

Therefore:

3. No disposition is causally efficacious with respect to its manifestation.

Let us first consider premise (2). Jackson's support for this premise, which is stated in the quote above, is unclear. Jackson assumes a second-order account of dispositions along the lines of the one outlined at the beginning of this chapter. He says "fragility, the disposition, is the second order property of having a first order property, the categorical basis, that fills the role of leading to breaking on dropping."²⁷ So, it follows that every fragile thing instantiates a property which is causally efficacious for breaking.

Now, suppose, for some fragile object, fragility itself is a property which is causally efficacious for breaking. For all Jackson has said, that is compatible with some other fragile object having some other property which is causally efficacious for breaking, while its fragility is inert. While it is true that "there is no way that the second-order property can be instantiated without the relevant causal power being instantiated," it does not follow that there is no way for the second-order property to be instantiated without *that property* having the relevant causal power. Jackson's claims can be set out as follows:

a) Necessarily, for any object x, x has disposition D iff there is some property P such that x has P, and P is causally efficacious for D's manifestation M.

b) Suppose that y has D, and D is causally efficacious for M.

c) Then, necessarily, if an object has D, D is causally efficacious for M.

However, (c) just doesn't follow. Suppose instead (b*) that y has F, and F is causally efficacious for M. We could rightly conclude from (a) that y has D. However, we would

²⁷Ibid., 256.

have no temptation to conclude that, necessarily, if an object has F, F is necessarily efficacious for M. Given supposition (b), it follows twice over that y has D. D satisfies that existential claim on the right side of the biconditional in (a). It follows from the definition of "disposition" that whenever an object has D, the existential claim is satisfied by some property. However, what doesn't follow, on the supposition that the existential claim is ever satisfied by D, is that it is always satisfied by D.

The only way (c) would follow is if (b) is understood as the supposition that D is necessarily efficacious for M. However, that supposition is just the claim that Jackson is trying to establish. Jackson may be attributing to the defender of causally efficacious dispositions with the view that dispositions are *necessarily* causally efficacious for their manifestations. However, in order to prove that dispositions are necessarily inert, one must refute not only the claim that disposition are always and necessarily efficacious, but also the claim that they sometimes are.

Recall that Jackson was trying to establish:

(2) If a disposition D were causally efficacious with respect to events of kind M (D's characteristic manifestation), it would be necessarily so.

Perhaps Jackson doesn't need to say anything quite so strong. A weaker claim along these lines is as follows:

(2*) Necessarily, if an event occurs in which an object has disposition D to M in C and C obtains, then that event will cause an M-event (other things being equal).

This claim follows more directly from the definition of "disposition," and is not subject to the same difficulties as (2). However, this is not in direct violation of Hume's Principle:

(1) If P is causally efficacious with respect to events of kind K, it is only contingently so (i.e. there is a world in which P is not causally efficacious with respect to K-events).

Perhaps Hume's Principle could be modified as follows:

(1*) If events with event-property P cause events of kind K, they do so contingently (i.e., there is a world in which P-events occur, but do not cause K-events.)

So, let's consider Hume's Principle, on either formulation. While Jackson doesn't provide an argument for this premise, perhaps it is supposed to be intuitively plausible. But is it? Consider the property of being negatively charged. Presumably, that property is causally efficacious for repelling negatively charged particles. According to (1), this should be a contingent fact about negative charge -- there should be a possible world in which a thing has negative charge, but is not thereby disposed to repel negatively charged particles. According to (2), there should be a world in which events which involves negatively charge particles being in proximity to other negatively charged particles occur, but these events does not cause the particles to be repelled from one another. But it is not at all obvious that we should admit such possibilities.

It is worth noting that Hume's Thesis is hardly common ground among philosophers. Robert Stalnaker, for example, claims that there are certain properties, such as mass, charge, or impenetrability, that cannot be "separated, conceptually, from the laws in which they occur and from the causal powers they confer on objects that instantiate them."²⁸ Or recall Shoemaker's view, that what determines the identity of a property "is its potential for contributing to the causal powers of things that have it."²⁹ According to Shoemaker, *every* property has its causal powers essentially. So, if we decline to "respect Hume's insight," we are not alone.

²⁸Robert Stalnaker, *Inquiry* (Cambridge, MA: MIT Press, 1987), 160.

²⁹Shoemaker, "Causality and Properties" in P. van Inwagen (ed.), *Time and Cause*, (Dordrecht: Reidel, 1980), 109-35.

In sum, neither the straightforward Analyticity Argument, nor Jackson's appeal to Hume's Principle, succeed in showing that dispositions are causally inert. Now, let's turn to a different attack on the causal efficacy of dispositions --the No Work Argument.

1.3. The "No Work" Argument

In his paper "Mental Causation and Two Conceptions of Mental Properties" Kim offers the following challenge to the causal efficacy of dispositions, or second-order properties more generally. A second-order property is characterized as "having some property or other with specification H." Kim goes on to ask:

given that the second-order property is distinct from each of the properties meeting specification H (call these "first-order realizers"), what causal contribution can it make over and beyond those made by their first-order realizers?³⁰

Given that the base properties are causally efficacious, does this exclude the disposition from being causally efficacious as well? The circumstances of manifestation and the base properties are sufficient for the manifestation, so there is a sufficient set of properties which does not include the disposition. In the case of fragility, Prior, Pargetter, and Jackson (hereafter PPJ) say "this causal basis is a sufficient causal explanation of the breaking *as far as the properties of the object are concerned*."³¹

The question then arises, what role can higher-order properties, such as dispositions, play? PPJ argue that since the base properties are sufficient for the effect, "there is nothing left for any other properties to do."³² Likewise, Kim says of the

³⁰Jaegwon Kim, "Mental Causation and Two Conceptions of Mental Properties," unpublished manuscript, 6.

³¹Elizabeth W. Prior, Robert Pargetter, and Frank Jackson, "Three Theses About Dispositions," *American Philosophical Quarterly*, 19 (1982), 225.

³²Ibid.

provocativeness of a red cape, "given the color of the cape as a full cause, there is no additional causal work for its provocativeness, or anything else."³³ This is what I call the "No Work" argument.

I will focus on Prior, Pargetter, and Jackson's version of this argument, given in their article "Three Theses about Dispositions". Their argument can be summarized as follows:

- 1) Every disposition has a causal basis.
- 2) The Distinctness Thesis: Causal bases are distinct from their attendant dispositions.
- 3) Given the circumstances of manifestation, a causal basis is sufficient for the manifestation of the disposition.
- 4) The Exclusion Principle: If the instantiation of a set of properties is sufficient to bring about a certain effect, then all other properties are causally inefficacious with respect to that effect.

Therefore,

- 5) Dispositions are causally inefficacious.

As applied to the example of fragility, the argument goes as follows:

- 1) Fragility has a causal basis, molecular bonding B.
- 2) B is distinct from fragility.
- 3) Given a striking of the right sort, B is sufficient for breaking.
- 4) If B and striking are sufficient for breaking, then fragility is inefficacious with respect to breaking.

Therefore,

- (5) Fragility is causally inefficacious for breaking.

³³"Mental Causation and Two Conceptions of Mental Properties," 2.

I grant Premise (1) since it follows from our working definition of "disposition."³⁴ We can also grant premise (3) for the sake of argument.³⁵ This leaves two principle ways of attacking the No Work argument --by denying premise (2) The Distinctness Thesis, or by denying premise (4) The Exclusion Principle. First, I will consider denying The Distinctness Thesis, and consider some difficulties involved with that approach. I then go on to pursue my favored strategy, challenging The Exclusion Principle.

1.3.1. The Distinctness Thesis³⁶

According to the Distinctness Thesis, a disposition is distinct from its causal basis. To deny this thesis is to identify a disposition with its causal basis. This would solve the problem of "dispositional causation" and avoid the conclusion that dispositions are causally inert. If the causal basis is causally efficacious, and the disposition is the causal basis, then the disposition is causally efficacious.³⁷

³⁴ Recall: x has disposition D to give manifestation M in circumstances C iff x has some property P which is a causal basis for giving M in C.

³⁵PPJ define causal basis as the properties of the disposed object that would be sufficient for the manifestation in the circumstances ("Three Theses About Dispositions," 251), and so premise (3) follows. We could contest this way of defining "causal basis," but PPJ's argument could be restated easily enough. (Consider a set of x's properties that are sufficient for M in C. Call it S. etc.)

³⁶Elsewhere (Ch. 2), I refer to this thesis as "the non-identity thesis," to distinguish it from what I mean by "distinctness" for the purposes of defining bare dispositions. However, since that distinction plays no role in the arguments presently under consideration, I stick with PPJ's terminology here.

³⁷This may seem to be at odds with the second-order characterization of dispositions. If a disposition is a second-order property, how could it be identical to its first-order realizer? However, on this coarse-grained account of properties, all that matters for property identity is extension. For example, the property of having property P = property P, since they necessarily have the same extension.

The fact that dispositions can be multiply realizable poses a difficulty for this approach. A disposition can have different causal bases in different objects. Crystal, porcelain, and egg shells are all fragile, and presumably they have different micro-structural properties that account for this. So, it seems that you cannot identify a disposition with its causal basis in every instance. For example, you cannot say both that fragility is identical to molecular bonding P and that fragility is identical to crystalline structure Q, because P and Q are distinct.

1.3.1.1. The Disjunctive Solution

One way to respond to this problem is to suggest that a disposition is identical to a disjunction of its realizers. Consider a case where there are only two realizers of fragility, P and Q. According to this Disjunctive Solution, being fragile is identified with having (P \vee Q). The causal efficacy of fragility is just that of P or Q.

However, even if a disposition is a disjunction of its various causal bases, that is not to say that the disposition is identical to any particular causal basis. So, if the causal basis of fragility were P in some instances, and Q in others, even if fragility is identical to (P \vee Q), it still wouldn't be identical to P, or identical to Q. It may seem as though identifying a disposition with a disjunction of its causal bases is not to deny the Distinctness Thesis at all.

However, we have been given no reason to think that a disposition can only have *one* causal basis in any given instance. So, a particular instance of fragility might have P as one causal basis, and (P \vee Q) as another causal basis. The disjunction of various causal bases could itself be a causal basis, and *that* causal basis would be identical to the disposition, on this view.

Notice, however, that this approach to attacking the No Work Argument would run afoul of the Exclusion Principle. It seems that there would be a set of properties which included P, but excluded (P \vee Q) which would be sufficient for breaking. If we

accepted the Exclusion Principle, it would follow that $(P \vee Q)$ is inefficacious with respect to breaking. Identifying the disposition with a disjunction of causal bases won't help to save the causal efficacy of dispositions if the disjunction of causal bases is itself causally inert. So, in order for the Disjunctive Solution to succeed, it will have to be coupled with an argument against the Exclusion Principle, such as the one I will offer in the next section.

Some, notably Lewis and Kim, do not think the Disjunctive Solution can work, for disjunctive properties are suspect. Lewis says that a disposition "unlike the various bases, is too disjunctive and too extrinsic to occupy any causal role."³⁸ Likewise, Kim says "the disjunction of heterogeneous properties can fail to be *projectible, nomic properties*."³⁹ If this is right, and dispositions are disjunctive properties, then the problem is not that there is no causal work for dispositions to do once the base properties do their job; it is that dispositions are not the sort of properties that are capable of doing any causal work at all.

Kim and Lewis seem to be claiming that a property such as $(P \vee Q)$ cannot be eligible for any causal or nomic role, simply by virtue of being a disjunctive property. But what does it mean to say a property is disjunctive? Intuitively, if P is a property and Q is a property, the disjunction of P and Q will be a disjunctive property. But consider, the case where P is the property of being a female cat, and Q is the property of being a male cat. Does this mean that the property of being a female cat or a male cat, i.e., being a cat, is a disjunctive property? No. Being named by a disjunctive predicate does not make a property disjunctive in the relevant sense. (We could make up disjunctive names for any property, e.g., being blue and square or being blue and not square.) To call a property disjunctive is to say that the things which have this property do not form a natural class.

³⁸Lewis, "Causal Explanation," *Philosophical Papers*. Volume 2. (Oxford: Oxford University Press 1983), 224.

³⁹"Mental Causation and Two Conceptions of Mental Properties," 11.

If we had a language in which all the natural properties were named by simple predicates, a disjunctive property would be one that could only be referred to by a disjunctive predicate. But even then, if a property were referred to by a disjunctive predicate, it wouldn't follow that the property was disjunctive, i.e., unnatural. This is because there would be the possibility that it could also be referred to by a simple predicate. One might think that this is impossible, because the disjunction of two natural properties is always a less natural property. However, this view is mistaken. The disjunction of two properties might turn out to be a *more* natural property, or at least natural enough to play some nomic role.

So, even if dispositions can be characterized in disjunctive terms, to what extent are they natural properties? If a disposition D is a property of having some property or other that fills causal role R, even if D is multiply realizable, all the properties that realize D will have at least one thing in common --they fill causal role R. If P and Q both realize D, they play the same causal role. Some philosophers, Shoemaker and Kim for example, think that properties can be classified by their causal roles. Kim says:

kinds in science are individuated on the basis of causal powers; that is, to be recognized as a useful property in a scientific theory, a property must possess (or be) a determinate set of causal powers. To put it another way, the resemblance that defines kinds in science is primarily *causal/nomological resemblance*: Things that are similar in causal powers and play similar roles in laws are classified as falling under the same kind.⁴⁰

On such an account, P and Q would seem to be the same kind. If they are of the same kind, their disjunction is not diverse and heterogeneous. Maybe P and Q can play other causal roles, and this fact can serve to differentiate them. But that does not show that they are of different kinds. Things of the same kind need not be qualitatively identical in every respect.

⁴⁰Kim, *Philosophy of Mind* (Bolder, Colorado: Westview Press 1996), 119.

Think of the role of alleviating pain in humans. There are a number of chemicals that fill that role, that have different constituents and different structures. However, by filling that role, they are considered to be of a kind --opiates. For one reason or another, all of them bind to opiate receptors in the brain. The class of opiates does not seem to be a heterogeneous, gerrymandered group. We are justified in classifying them as the same kind because they have similar effects.

It is important to point out that disjunctive properties may sometimes be natural, causally efficacious properties. However, this approach to saving the causal efficacy of dispositions has its limitations. While this approach may work for some dispositions, like the disposition to relieve pain in humans, it seems less plausible in other cases. Think of provocativeness -- the disposition to elicit an aggressive response. Something can elicit an aggressive response in an animal, suppose, if it is a certain color, if it has a certain odor, or if it is moving in a certain way. But it is hard to believe that a disjunctive property like (being red or oscillating or smelling like blood) is a natural property that is apt for figuring in causal laws.

One route that one might take in response is to say that some disposition terms refer to natural properties, while other disposition terms do not. Perhaps terms like 'provocative' are merely terms we use to make causal generalizations about objects which are not intrinsically similar in any relevant respect. But another disposition could be a natural property if all of its realizer properties formed a natural class. By this strategy, one would not support the causal efficacy of every property we have been calling 'dispositional', but only the dispositions which are plausibly natural properties. Let's summarize the dialectic between the Friend and the Foe of causally efficacious dispositions.

Friend: The No Work argument fails because the Distinctness Thesis is false. A disposition = its causal basis.

Foe: But some dispositions are multiply realizable.

Friend: Then a disposition = a disjunction of its realizer properties.

Foe: But disjunctive properties can't be causally efficacious, because they are non-natural properties, which do not figure in causal laws.

Friend: A disjunction of realizer properties is not necessarily a non-natural property.

Foe: Some disjunctions of realizer properties are very heterogeneous, and it is implausible to think they are natural properties.

Friend: Some disposition terms refer to non-natural (causally inert) properties, but there are some natural (causally efficacious) dispositions.

Perhaps there are other moves that our Friend, the proponent of the Disjunctive Solution could make, but if he takes this route, he winds up with a somewhat qualified defense of causally efficacious dispositions. Without further investigation, and without some guide as to how to determine the naturalness of a property, it is not clear which dispositions would come out to be causally efficacious on this approach.

There is a different reason some philosophers give for being troubled about the causal efficacy of disjunctive properties. It is that admitting disjunctive properties as causally efficacious will lead to overdetermination. Say the glass's having property P is sufficient for breaking when struck. If the property (P v Q) were also causally efficacious, it would seem to *overdetermine* the breaking. If P is responsible for the breaking, as far as the properties of the glass are concerned, what is the role of the disjunctive property (P v Q)? Kim says: "it seems at best gratuitous, if not incorrect, to trot out the disjunction *also* as a cause."⁴¹ But we are getting a bit ahead of ourselves. This argument against

⁴¹ "Mental Causation and Two Conceptions of Mental Properties," 10.

the causal efficacy of properties like $(P \vee Q)$ seems to be invoking The Exclusion Principle, to which we will turn shortly. But first, I will consider another approach to attacking the Distinctness Thesis.

1.3.1.2. The Trope Solution

A different response to the Distinctness Thesis is put forth by Cynthia and Graham MacDonald, among others.⁴² The view can be summarized as follows. While the Distinctness Thesis applies at the level of properties, understood as classes or universals, it does not apply at the level of property instances, or tropes. So, a property instance of a disposition can be identical to a property instance of a base property. Furthermore, it is particular property instances that are causally efficacious for events. So, while it is conceded that 'fragility' isn't causally efficacious, that's not a problem, because particular havings of fragility are. Let me go over that more slowly.

In addition to talking about redness, something shared by all red things, we can also talk about a particular apple's redness -- a particular instantiation of redness, also called a red trope. Suppose that the apple is a particular shade of red, say russet. Red and russet stand in the relation of determinable to determinate. Now, this apple's "russetness" bears a relation to its redness that it does not bear to, say, its roundness. One might say that the apple's redness is nothing "over and above" its russetness.⁴³ Redness and

⁴²Graham and Cynthia MacDonald, "Mental Causation and Explanation of Action" in *Mind, Causation and Action*, ed. L Steveson, R. Squires and J. Haldane (Oxford: Basil Blackwell, 1986), 38. See also Douglas Ehring, "Mental Causation, Determinables and Property Instances," *Nous*, 30 (1996) 461-80. The target of this attack is epiphenomenalism about mental properties; however, the same arguments can be applied to dispositions.

⁴³Incidentally, I do not share the intuition that the apple's redness is identical to its russetness. I could pick the apple out of a barrel of crimson apples because of its russetness (not its redness). The apple might change its shade as it ripens, losing its russetness, but not its redness. These considerations are perhaps not decisive, but further

russetness are two different properties, having different extensions. However, proponents of this view would say that this apple's russetness is the same thing as its redness. Its red trope = its russet trope. In the same way, it is thought, that a particular glass's fragility is the same thing as its crystalline structure. Its crystalline structure is its particular way of being fragile, just as the apple's russetness is its particular way of being red.

The key to this strategy is the claim that causal efficacy is a relation between a trope and an event, rather than a property and an event. Couple this with the claim that one trope is an instance of a determinate and its determinables, and this view is subject to a serious difficulty, which can be brought out by certain counterexamples.

To adapt an example from Stephen Yablo,⁴⁴ suppose you are shipping packages, and you are constrained by a 20 pound weight limit. You have a crude scale that will only tell you if your package is over 20 pounds. So, you put a package on the scale which happens to weigh 21 pounds. The scale indicates that the package is over the weight limit. Now, it seems that *weighing 21 pounds* was causally efficacious for tipping the scale. By the above account, the package's *weighing over 20 pounds* is the same trope as its *weighing 21 pounds*, so its *weighing over 20 pounds* was causally efficacious. However, *weighing 21 pounds* is also a determinate of other determinables, such as *weighing less than 30 pounds*. It follows that the package's *weighing-less-than-30-pounds* trope is also identical to its *weighing-21-pounds* trope. Do we want to say that the package's *weighing less than 30 pounds* was causally efficacious for tipping the scale? If we bite that bullet, we invite a host of other counter-intuitive causal efficacy claims.⁴⁵ The class of causally efficacious tropes will be much broader than we might have suspected.

discussion of them would take us too far afield.

⁴⁴Stephen Yablo, "Mental Causation," *The Philosophical Review*, 101 (1992), 259 (footnote 32).

⁴⁵This account already invites a host of counter-intuitive identity claims, such as the

What's worse is that, on this view, we lose the sense that it is in virtue of instantiating a certain property that an event has the effect that it does. In our earlier discussion, we saw that some philosophers think that to be causally efficacious, a property must be apt to figure in causal laws. But tropes, being particulars, are not apt for this role, and they provide us with no way of generalizing from a particular causal claim. Assuming regularity of the laws, if an instance of a property was causally efficacious for a certain effect in a certain circumstance, that should give us some reason to think that an instance of the same property in a similar circumstance will be similarly efficacious. However, if we are inclined to think that a *weighing-less-than-30-pounds* trope was causally efficacious for tipping the scale in a particular instance, we would be mistaken in concluding that further instantiations of *weighing less than 30 pounds* will be causally efficacious for tipping the scale. One might think that causal efficacy claims should support counterfactuals, but if causal efficacy claims are claims about tropes, they will not.

Perhaps if we distinguish tropes of determinables from tropes of their determinates, there is some sense in which these tropes can be said to be causally efficacious. However, then the trope solution would no longer stand as a challenge to the Distinctness Thesis. On the other hand, we could say that it is in virtue of being a trope of a particular property that the trope is causally efficacious. However, then we are back to talking about properties, and our introduction of tropes has not advanced our argument. For these reasons, I am inclined to stick with thinking of properties, rather than tropes, as the relation of causal efficacy.

In sum, I have considered two approaches to denying the Distinctness Thesis, the second premise of the No Work argument. One is the Disjunctive Solution, according to which a disposition is identical to a disjunction of its realizers. The other is the Trope

package's *weighing over 20 pounds* = its *weighing less than 30 pounds*.

Solution, according to which instances of a disposition are identical to instances of its realizers. I think that both of these approaches face difficulties, though perhaps they are not insurmountable. Now I turn my attention to the fourth premise of the No Work argument -- the Exclusion Principle.

1.3.2. The Exclusion Principle

As we have seen, even if we deny the Distinctness Thesis via the Disjunctive Solution, we still need to deny the Exclusion Principle. However, even if we are inclined to accept the Distinctness Thesis, a denial of the Exclusion Principle alone can serve as an alternative way to attack the No Work Argument. According to the Exclusion Principle, if the instantiation of a set of properties is sufficient to bring about a certain effect, then all other properties are causally inefficacious with respect to that effect. As PPJ say "a complete causal explanation excludes competitors."⁴⁶ It is claimed that to deny the Exclusion Principle is to allow for spurious over-determination. It would be to say that, given the circumstances of manifestation, both the causal basis *and* the disposition are each sufficient for the effect. If this happened every time any disposition is manifest, we would have as Block says "bizarre, systematic over-determination."⁴⁷

Standardly, an event *e* is causally overdetermined if two or more distinct events occur, each of which is sufficient to cause *e*. Admittedly, there is something wrong with postulating too many coincidences. And if a great many effects systematically had two distinct events that were sufficient for causing them, that would run counter to our understanding of the causal structure of the world. Maybe that would be too high a price to pay for saving the causal efficacy of dispositions.

⁴⁶"Three Theses About Dispositions" 225.

⁴⁷Block, "Can the Mind Change the World?" 159.

However, one should not overlook the fact that these considerations and intuitions about overdetermination apply to two (or more) *events* overdetermining an effect. But what we are concerned with in the case dispositions and causal bases are *the properties of a single event* (or of a single object involved in an event). But it is not clear what it means to say an effect is overdetermined by an object's *properties*. Perhaps it would be to say that there are two different sets of properties, both of which such that their instantiations are sufficient for a certain effect.

However, it is not clear that this kind of "overdetermination" is so worrisome. Most events instantiate a huge number of properties, many of which bear logical or nomological relations to one-another. Recall the account of causal efficacy, according to which a property is causally efficacious if it is a member of a set of event-properties S that is minimally sufficient for a certain effect. This analysis is incomplete and needs further constraints. However, whatever further constraints are appropriate, there is no guarantee that, for any event, there is a unique set S. It seems that any time an event is caused, the cause instantiates several sets of properties, each of which is sufficient for the effect. It depends on how you want to carve things up. Suppose redness was a member of a set of properties that was sufficient for provoking the bull. It seems that there would be another set of properties, that included crimson instead, that would be sufficient.

Can we, in fact, single out one special set of event-properties are minimally sufficient for an effect, and declare the rest causally inert? This may prove difficult to do, especially when properties are so intimately related, as a disposition is to its causal basis. Do we have reason to suppose that, for any effect, there was only one minimally sufficient set of causally efficacious properties?

I suggest we do not. To try to convince you of this, I invite you to think about how the Exclusion Principle would work in practice. Let us make the following suppositions:

An event occurs, which involves an object *o*.
o has property *F*.
F is causally efficacious for some event *e*.
o has *H*, and $H \neq F$.

We are wondering whether *H* is causally efficacious for *e*. *o*'s having one property can entail that it has another property, analytically, nomologically. Suppose that *o*'s having *F* entails that *o* has *H*. It follows from the Exclusion Principle that *H* is causally inert with respect to *e*. A set of properties which included *F* and *excluded* *H* would be sufficient for *e*, and so a set of properties which included both *F* and *H* would not be a *minimally* sufficient set, which is what is required for causal efficacy. Conversely, if *o*'s having *H* entails that *o* has *F*, we know that *H* is not causally efficacious, for if it were, by the same reasoning as above, then *F* would not be, and by stipulation, it is.

So, according to the Exclusion Principle, there is some privileged set of properties which are causally efficacious for a given effect, such that all properties which entail them and all properties which are entailed by them are causally inert.⁴⁸ This implication seems counter-intuitive. Consider the following propositions:

The cape has surface reflectance property *R*.
The cape is red₂₁.
The cape is crimson.
The cape is red.
The cape is colored.

Now, suppose that each proposition entails its successor. If one of the above properties is causally efficacious for a certain effect, it follows that all of the others are causally inert with respect to that effect. How do we decide which property is the efficacious one? It is not clear how to answer, and this does not seem to be just an epistemic problem. It is not

⁴⁸Properly speaking, it is the propositions to the effect that the properties are instantiated that entail or are entailed.

clear what would determine the level of specificity at which the causal action is going on.⁴⁹ One might assume that all of the causal action happens at the most fundamental level. But what is the basis of that assumption? Such an assumption might have serious counter-intuitive consequences, for example, that all of the macro-properties we regularly observe are causally impotent. Furthermore, assuming that all causal action happens at the level of fundamental properties may not serve the foe of causally efficacious dispositions; it is not at all obvious that the fundamental properties are wholly non-dispositional.

Strawson, for example, claims:

It seems that our search for the properties of the categorical base must finally lead us to the undeniably theoretical properties which physics assigns to the ultimate constituents of matter-- perhaps force, mass, impenetrability, electric charge. But these properties seem to be thoroughly dispositional in character...⁵⁰

The implausibility of the Exclusion Principle in practice may be better illustrated by returning to the example of the scale. Recall that our scale will tell you if your package is over 20 pounds, and your package is, in fact, 21 pounds. The scale indicates that the package is over the weight limit. Now, what property of the package was causally efficacious for tipping the scale? Was it the property of weighing 21 pounds, or the property of weighing over 20 pounds? If it was the property of weighing 21 pounds, does that mean that the property of weighing over 20 pounds was causally impotent?

Intuitively, it seems not. I want to be able to say both weighing 21 pounds and weighing over 20 pounds were causally efficacious. It seems to me that there is no reason

⁴⁹One could say that the property which is adequately specific is the most specific one which is such that, if it weren't instantiated, the effect would not follow. However, this move cannot save the No Work argument. As we have noted, if properties which satisfy such a counterfactual were causal efficacious, then dispositions would be causally efficacious with respect to their manifestations.

⁵⁰"Reply to Evans," in Zak van Straaten, ed., *Philosophical Subjects* (Oxford: Clarendon Press 1980), 280.

to suppose that the causal efficacy of one property precludes the other property from being causally efficacious as well. Either could be a member of a minimally sufficient set of properties. Hence I think that the Exclusion Principle is false. If the instantiation of one property entails the instantiation of another, there seems to be no *prima facie* reason to suppose that only one of them can be causally efficacious. I think the worry that we are going to end up with too many causally efficacious properties is an unfounded one. We might be want to limit the number of events that we consider sufficient to cause a given effect. However, the drive to put a cap on causally efficacious properties seems to me unmotivated. This is especially true when the properties in question have some logical or law-like connection. There are many ways to describe an event, many properties of the event you can cite, different levels of specificity which you can appeal to. While I intuitively grasp the idea that there are some causally efficacious properties, and some inefficacious or irrelevant ones, I doubt that there is some privileged minimal set of properties that are causally efficacious for a given effect, to the exclusion of all others. Similarly, I am skeptical that there is some privileged causal explanation that excludes all competitors.

1.4. Conclusion

In conclusion, I think that these worries about the causal efficacy of dispositions can be dispelled. The Analyticity Argument can be countered by pointing out that conceptual connections do not preclude causal connections, and by denying Hume's Principle. The "No Work" argument can be addressed in two main ways. One way is to deny the Distinctness Thesis, and say that a disposition can be identified with a disjunction of its causal bases, or that a disposition trope can be identified with a causal basis trope. The other approach is to deny the Exclusion Principle --that is to deny that there is some privileged explanation or some special set of properties which preclude all other properties from playing a role in the causal story. Despite my attempts to clarify the notions of

causal relevance and causal efficacy, I suspect that these notions are still too vague and ill-defined to give us a clear way of deciding which properties fall into which category.

Barring some explication of causal efficacy which clearly excludes dispositions, I do not think we should be worried, for these reasons at any rate, that dispositions are causally inert.

Chapter 2: The Bare Metaphysical Possibility of Bare Dispositions

It is asked what the cause and reason are of opium's making one sleep.

To which I respond: because there is in it a dormitive virtue whose nature it is to put the senses to sleep.

Moliere, "Le Malade Imaginaire"

2.1. Introduction

As Moliere's jest illustrates, if someone is wondering why opium puts one to sleep, telling him that it has a disposition to do so is not very explanatory. More ought to be said about why opium puts one to sleep, and in fact, we can say more: opium contains alkaloids such as morphine which, being structurally similar to the body's naturally occurring peptides, bind to opiate receptors in the brain, causing sleep. Some people think that all dispositions are like the dormitivity of opium, in that there must always be another property which causally explains the manifestation of the disposition. When someone asks why something produces a certain effect, they are often looking for a deeper explanation than just "because it is disposed to produce that effect." Inability to produce a deeper explanation, on this view, reflects a failure of understanding, or ignorance.⁵¹ It is supposed that there must be something other than the disposition that causally explains the manifestation, or to use terminology that is now common, that every disposition must have a distinct "causal basis."

This chapter takes issue with that supposition. Are all dispositions necessarily like opium's dormitivity, in having distinct causal bases that explain their manifestations? Many have thought so, and not without reason: there is something admittedly puzzling about the idea of a *bare disposition* -- a disposition which has no distinct causal basis. Some have argued that the very idea of a bare disposition is incoherent. I disagree. Bare

⁵¹"We thus expose ourselves to Moliere's ridicule, and, if we did nothing further, we would deserve it." D. M. Armstrong, *Belief, Truth and Knowledge*, 15.

dispositions are possible. Moreover, it is an open question whether any objects have bare dispositions in this world.

Significantly, bare dispositions can figure in larger metaphysical programs, for example, the phenomenalist view that matter is the "permanent possibility of sensation."⁵² More recently, some philosophers have defended the view that the fundamental properties of the ultimate constituents of matter are dispositional.⁵³ On this view, the world abounds with bare dispositions. The possibility of bare dispositions has implications for views like Lewis's "Humean supervenience," according to which everything that is true about the world supervenes on "a vast mosaic of local matters of particular fact."⁵⁴ I take it that, according to such views, these local matters of particular fact are not dispositional. It follows that, if the world contains bare dispositions, Humean supervenience is false.⁵⁵ Clearly, a defense of bare dispositions has broad philosophical significance.

My defense of bare dispositions will proceed as follows. In part 2.1, I explain more fully what I mean by "causal basis," and "bare disposition." In part 2.2, I consider the claim that it follows from the concept of a disposition that it is not bare. In part 2.3, I

⁵²John Stuart Mill, *Collected Works of John Stuart Mill, Volume IX: An Examination of Sir William Hamilton's Philosophy and of The Principal Philosophical Questions Discussed in his Writings*, ed. J. M. Robson (Toronto: University of Toronto Press, 1979), 183.

⁵³For example, Simon Blackburn, "Filling in Space," *Analysis* 50 (1990), 60-65; P. F. Strawson, "Reply to Evans" in Zak van Straaten, ed., *Philosophical Subjects* (Oxford: Clarendon Press, 1980), 273-82; Richard Holton, "Dispositions All the Way Around," *Analysis* (forthcoming).

⁵⁴David Lewis, *Philosophical Papers*, vol. II (New York: Oxford University Press, 1986), ix.

⁵⁵I'm assuming that if a disposition had a supervenience base of categorical properties, that would serve nicely as a distinct causal basis for that disposition. Hence, if a disposition were bare, true statements about that disposition would not supervene on local matters of particular, non-dispositional fact.

consider arguments, due to Prior, Pargetter, and Jackson, that all dispositions necessarily have distinct causal bases.⁵⁶ In part 2.4, I consider arguments, recently articulated by Smith and Stoljar, that there can't be bare dispositions because they would make for unwelcome "barely true" counterfactuals.⁵⁷ In the end, I find no reason to deny the possibility of bare dispositions.

2.1.1. Causal Bases

When a sleeping pill puts someone to sleep, that event can be causally explained in terms of the chemical properties of the pill. When a fragile glass breaks, that event can be causally explained in terms of the microstructural properties of the glass. These chemical and microstructural properties are causally relevant to, and causally efficacious for, the associated manifestations. A property of a disposed object which is causally efficacious for the manifestation of the disposition is called a *causal basis* of that disposition. A causal basis is a property of an object which is causally efficacious for the manifestation of the disposition.

So, a causal basis is a causally efficacious property, but what kind of property? There are three salient candidates. Either causal bases are always categorical, they are always dispositional, or they can be either categorical or dispositional. For example, a causal basis for fragility might be a particular type of molecular bonding. Perhaps, to have a particular type of molecular bonding is to have a dispositional property. As Armstrong says, "To talk of molecular *bonding* is surely to talk again in terms of dispositions of bonded things."⁵⁸ If a type of molecular bonding can serve as the basis of fragility, say,

⁵⁶"Three Theses About Dispositions," 251-257.

⁵⁷Michael Smith and Daniel Stoljar, "Global Response Dependence and Noumenal Realism," *Monist*, 81 (1998), 85-111.

⁵⁸*Belief, Truth and Knowledge*, 13.

then there can be causal bases of dispositions that are themselves dispositions. I use the expression "causal basis" as neutral between dispositional bases and categorical bases, unless stated otherwise.

Note that a causal basis is not, conceptually or by definition, distinct from its associated disposition: if fragility turns out to be causally efficacious for breaking, then fragility is its own causal basis. This will become important in the arguments that follow. As I argued in Chapter 1, dispositions can be causally efficacious. Anyone who claims that dispositions can have dispositional causal bases is supposing that dispositions can be causally efficacious. If it is possible that a disposition can be the causal basis of a disposition, what is to preclude a disposition from being a causal basis of itself?

It might seem counter-intuitive to say that a disposition can be its own causal basis. However, I think this sense of counter-intuitiveness results from confusions which can be dispelled when one gets clear about the how these expressions are being employed. To say that a disposition can be its own causal basis is not to say that a disposition causally explains itself, but only that it causally explains its manifestation. Furthermore, when one says "a disposition has a causal basis," this does not, by itself, suggest that the disposition and the causal basis are distinct. A disposition and a causal basis are both *properties* which are instantiated by *objects*. It is the object with the disposition that *has* the causal basis, in the property-instantiation sense. To say that a disposition has a causal basis is not to say that one property instantiates another; rather, it is to say that an object with that disposition instantiates a property which is causally efficacious for the manifestation of that disposition. Showing that a disposition has a causal basis does not by itself show that it has a *distinct* causal basis. If one object instantiates a dispositional property and a causally efficacious property, for all that has been said, they might be one and the same property.

If a disposition is efficacious for its manifestation, this by itself does not rule out a categorical property of the object being efficacious for the manifestation, too. We need not assume that a thing can only have one causal basis per disposition; perhaps many of an object's properties are efficacious for the manifestation of the disposition. Therefore, a disposition could be its own causal basis, and have a distinct causal basis as well.

2.1.2. "Bare Disposition" Defined

Equipped with this understanding of causal bases, we are now in a better position to understand the concept of a bare disposition. *A bare disposition is a disposition that has no distinct causal basis, neither dispositional nor categorical.*⁵⁹ A disposition whose unique causal basis is itself would count as a bare disposition. If an object has a bare disposition, the object has no intrinsic properties which are both distinct from the disposition and causally efficacious for its manifestation. One might say it is just a brute fact about the thing that it is so disposed. For example, suppose a glass were "barely fragile," and it shattered. The only properties of the glass which could be causally efficacious for the shattering are properties which are not distinct from fragility.

In saying that a bare disposition has no distinct causal basis, by "distinct" I cannot mean merely non-identical. If "distinct" meant just non-identical, finding multiple causal bases distinct from the disposition would be too easy. For example, if *being fragile* is causally efficacious for breaking, perhaps *being a fragile glass* is as well. If we don't treat *being fragile* as distinct from *being a fragile glass*, the glass's fragility could not be a bare

⁵⁹ Some philosophers use the expression 'bare disposition' differently than I am using it here. For example, Mark Johnston defines 'bare disposition' as follows: If *x* has a bare disposition, "*x* would *R* in *S* under *C* and no intrinsic feature of *x* or of anything else is the cause of *x*'s *R*-ing in *S*" ("How to Speak of the Colors," *Philosophical Studies* 68 (1992), 234.) Johnston's definition significantly differs from mine in that it rules out bare dispositions that are both intrinsic and causally efficacious for their manifestations.

disposition, simply because *being a fragile glass* would count as a distinct causal basis. So we need to understand "distinctness" as something other than non-identity. Intuitively, we need a notion of two properties being disjoint or non-overlapping. One way to get at this idea is to say, if the having of property F entails the having of property G, it follows that F and G are *not* distinct. So, *being a fragile glass* is not distinct from *being fragile*.⁶⁰

Provided bare dispositions are possible, it is an open question whether they are actual. While fragility does not look like a good candidate for a bare disposition, perhaps some of the dispositions of fundamental particles are. Consider the property of being negatively charged, and the dispositions of negatively charged things, such as being disposed to repel other negatively charged things. Could there be a possible world in which particles were negatively charged, but not thereby disposed to repel other negatively charged particles? If such a world is not possible, then the connection between being negatively charged and behaving in certain ways in certain circumstances is not accidental or contingent. This suggests that negative charge is itself a dispositional property (which is not distinct from the disposition to repel negatively charged particles). Furthermore, it seems probable that there is no structural, micro-physical property of an electron which accounts for its dispositions to repel and attract other particles -- at any

⁶⁰Defining distinctness this way may seem to have counter-intuitive consequences. Suppose some theorist thinks that all dispositions have categorical bases, and he holds a Shoemaker-type view, according to which properties have their causal powers essentially. Understanding 'distinctness' as explained above, it seems that he thinks that dispositions are not distinct from their categorical bases, and consequently (on my view) he thinks that all dispositions are bare. This may seem counter-intuitive. However, in saying that the having of some categorical property analytically entails the having of a disposition, the theorist would be using the expression "categorical property" in a non-standard way. The standard way of understanding "categorical property" involves the idea that the property is in some sense non-modal, that it doesn't depend on what is going on in non-actual worlds, so to speak. It is not clear that a property that has its causal powers essentially can be categorical in this sense. If the imagined theorist is saying that a fundamental property involves certain causal powers, it is not clear where the disagreement lies between such a theorist and a bare dispositionalist.

rate, current physics does not tell us otherwise. If this is right, then bare dispositions are more than some remote metaphysical possibility. They are posited by our best scientific theories. As Strawson points out:

It seems that our search for the properties of the categorical base must finally lead us to the undeniably theoretical properties which physics assigns to the ultimate constituents of matter-- perhaps force, mass, impenetrability, electric charge. But these properties seem to be thoroughly dispositional in character...⁶¹

Or, as Blackburn claims, "science finds only dispositional properties, all the way down."⁶²

So, there is reason to think that bare dispositions aren't merely possible, but are actually instantiated by the ultimate constituents of our world.

2.2. The Conceptual Argument

Some say that the very conception of a disposition precludes bare dispositions, that it is part of the concept of a disposition that it has a distinct causal basis.⁶³ For example, a disposition can be said to be:

a higher order property of having some distinct intrinsic properties which would cause the manifestation of the disposition in the circumstances of manifestation.⁶⁴

If this definition is correct, bare dispositions are ruled out a priori; it is an analytic truth that all dispositions have distinct causal bases.

⁶¹"Reply to Evans," 280.

⁶²"Filling in Space," 255.

⁶³Gareth Evans, "Things Without the Mind," in Zak van Straaten, ed., *Philosophical Subjects* (Oxford Clarendon Press), 76-116; also Smith and Stoljar, "Global Response Dependence..." 88-90.

⁶⁴Adapted from Johnston, "How to Speak of the Colors," 234. Johnston does not endorse this analysis.

However, the causal relations of an object seem far more central to disposition ascriptions than its intrinsic properties. Consider what we would say in a situation where we felt justified in making a certain disposition claim, but could not find a distinct causal basis for that disposition. Suppose an object *x* reliably produces manifestation *M* under circumstances *C*, but we can find no property distinct from the disposition that is causally efficacious for the manifestation. Perhaps we would assume that there has to be a distinct causal basis, but we just haven't discovered it yet. But what would be our grounds for making that assumption? Suppose we were in a far superior epistemic situation; we have a (nearly) exhaustive list of *x*'s properties, and their causal upshots. We are trying to decide whether to add disposition *D* to the list of *x*'s properties. We find no other properties that are causally efficacious for *M*, and yet whenever *x* is in *C*, it gives *M*. What are we to say? If *x* will give *M* whenever it is in *C*, it seems natural to say that *x* is disposed to give *M* in *C*, and that it has a disposition to give *M* in *C*. It would do more injustice to our linguistic practices to deny that the thing has the disposition than to say that it does, inexplicable as that may be.

One might argue the situation I have described is impossible; however, there is nothing in our language or our concepts that tells us that. We can imagine making a discovery like the one described above, so it is no part of the meaning of "disposition" or of disposition terms in general that there is a distinct causal basis. If the conceptual analysis were correct, we would have to say that the object in our example did not have the disposition that we thought it did. Even if the foes of bare dispositions were right to say this, though, I don't see how they would have achieved any more than a verbal victory. There is a substantive issue still left, concerning what types of properties objects can have. If what I'm calling "bare dispositions" shouldn't be called "dispositions" strictly speaking, they might nevertheless be possible.

2.3. PPJ's Causal and Non-Identity Theses⁶⁵

In "Three Theses about Dispositions," PPJ defend:

- (1) The Causal Thesis: All dispositions have causal bases; and
- (2) The Non-Identity Thesis: Causal bases not identical to their attendant dispositions.

These theses are meant to apply not only to dispositions as they happen to be in this world, but to all possible dispositions. It follows that it is a necessary truth that every disposition has a causal basis which is distinct from itself, and consequently bare dispositions are impossible.⁶⁶ Let's consider PPJ's arguments for these two theses in turn.

2.3.1. The Causal Thesis

According to the Causal Thesis, it is a necessary truth that dispositions have causal bases. PPJ define the causal basis as the property or properties of the object that, together with the antecedent circumstances, is the causally operative sufficient condition for the manifestation of the disposition.⁶⁷ Their argument for the causal thesis proceeds as

⁶⁵Prior, Pargetter, and Jackson call this thesis "The Distinctness Thesis." However, it is clear that by 'distinct', they mean non-identical. As I am using the expression 'distinct', two properties can be non-identical, yet fail to be distinct if one entails the other. To avoid confusion, I translate PPJ's distinctness talk into non-identity talk.

⁶⁶To be more precise, it would follow that there can be no dispositions which have no causal bases, nor any dispositions which are identical to their causal bases. However, PPJ's arguments leave open the possibility that there are dispositions which are neither identical to, nor distinct from, their causal bases. (I have in mind two different properties which are not distinct in the sense I explained earlier, because one entails the other.) These dispositions would count as bare dispositions on my view. However, I am interested in defending the claim that dispositions can be their own causal bases, so I take PPJ's challenge seriously.

⁶⁷This is along the lines of my definition of "causal basis," according to which to have a causal basis for the disposition to giving M in C is to have a property which is causally efficacious for M in C.

follows. First they argue that all dispositions in deterministic worlds have causal bases. In other words, in deterministic worlds, all manifestations of dispositions have causally operative sufficient conditions. That, they claim, follows from Determinism. Then they go on to argue for the Causal Thesis in indeterministic worlds. Having argued that dispositions must have causal bases in both deterministic and indeterministic worlds, PPJ take themselves to have shown that the Causal Thesis is true in all possible worlds.

Because I think that problems for this argument crop up at the first stage, I will focus on the argument for the Causal Thesis in deterministic worlds. PPJ accept a counterfactual analysis of dispositions; they define "A is fragile" as "If A were knocked at t , A would break at $t+\delta$."⁶⁸ They go on to say:

Suppose our world is deterministic, and consider again the case of a fragile glass A. We are interested in the truth of "If A were knocked at t , A would break at $t+\delta$ " and thus of the truth of "A breaks at $t+\delta$ " in the closest possible worlds where A is knocked at t (which may include ours, of course). The closest worlds will be deterministic and have the same laws as ours; to suppose otherwise would be to make a gratuitous departure from actuality. But then it will be either determined that A breaks, or that A does not break. In the latter case, clearly A is not fragile. In the former there will be a causally sufficient antecedent condition operative in producing the breaking --that follows from Determinism. Hence if A is fragile and Determinism is true, there must be a causal basis.⁶⁹

This argument can be set out as follows:

- 1) Suppose our world @ is deterministic and contains a fragile glass A.
- 2) Given the counterfactual account of dispositions, and a possible worlds semantics of counterfactuals, if A is fragile, in the closest possible world w where A is knocked at t , A breaks at $t+\delta$.
- 3) Therefore, if A does not break in w at $t+\delta$, then A is not fragile.
- 4) Since @ is deterministic, and w is close to @, w is a deterministic world.
- 5) If A is fragile, A is determined to break in w at $t+\delta$.

⁶⁸"Three Theses About Dispositions," 252.

⁶⁹Ibid., 251.

- 6) If A's breaking in w at $t+\delta$ is determined, A's breaking in w at $t+\delta$ has causally operative sufficient conditions.
- 7) Therefore, A's breaking in w at $t+\delta$ has a causally operative sufficient condition.
- 8) Therefore, A's breaking in w must have a causal basis.
- 9) "Hence, if A is fragile and Determinism is true, there must be a causal basis."

Two features of this argument that I find problematic are premise (6) and the move from (7) to (8). With regard to premise (6), it is not clear exactly what PPJ mean by determinism, or what they think follows from it. Later on, PPJ say "Determinism is precisely the view that sufficiently alike repetitions must lead to the same result."⁷⁰ However, premise (6) suggests that if an event e is determined, then it follows that e has causally operative sufficient conditions. But this doesn't follow from determinism as stated; it could be that like repetitions lead to the same result, and yet some events happens spontaneously --having no causally operative sufficient conditions. An alternative and attractive account of determinism offered by David Lewis is as follows: if Determinism is true:

the prevailing laws of nature are such that there do not exist any two possible worlds which are exactly alike up to some time, which differ thereafter, and in which those laws are never violated.⁷¹

However, (6) doesn't follow from this principle either; Lewis's account makes no mention of causally operative sufficient conditions. In some possible world, the laws of nature might be such that some event is followed by a particular sequence of events, and yet there is no causation going on. (Incidentally, if there were dispositions in such a world, they would be bare dispositions.) So, on the supposition that Determinism is true, on either

⁷⁰Ibid., 252.

⁷¹"Causation" *Philosophical Papers*. Volume 2. (Oxford: Oxford University Press 1983), 196.

Lewis's or PPJ's own construal of Determinism, it does not follow that every event has a causally operative sufficient condition.

Instead, what PPJ seem to be assuming in premise (6) seems to be something more akin to The Principle of Universal Causation --the thesis that every event has a cause. So, what PPJ have shown is that in worlds where every event has a causally operative sufficient condition, events such as shatterings have causally operative sufficient conditions. However, PPJ's announced strategy is to exhaust all the possibilities by considering the deterministic worlds and the indeterministic worlds. After arguing that the Causal Thesis holds in "deterministic" worlds, they say

Therefore, the problem cases for the Causal Thesis can only lie in indeterministic worlds..., and so if we can show that even there we cannot get a counterexample we will have vindicated the Thesis.⁷²

What they have done instead is consider the worlds in which every event has causally operative sufficient conditions. (Their discussion of the indeterministic worlds suggest that these are not worlds in which some events have no causes, but worlds in which some events have probabilistic causes.) They have said nothing about the deterministic and indeterministic worlds in which some events have no causes. So, they have not ruled out the situation in which the manifestation of a disposition can happen spontaneously, and hence the disposition is bare. Perhaps this isn't such an egregious oversight. If their argument works for worlds in which every event has a cause, perhaps that will do. However, I don't think their argument works even in this restricted domain.

There is a more serious problem with this argument. Even if PPJ have established that A's breaking in w must have causally operative sufficient conditions, it is not clear what this allows them to conclude about A in the actual world. Recall that PPJ define the casual basis as the properties of the *object* that are part of the causally operative sufficient condition for the manifestation of the disposition. But premise (7) says nothing about A's

⁷²"Three Theses About Dispositions," 252.

properties; it just says that A's breaking in w had a causally operative sufficient condition. This leaves open the possibility that the causally operative sufficient conditions for A's breaking in w are entirely extrinsic to A. PPJ say later "We have not here argued that the causal basis is categorical or intrinsic, only that it exists."⁷³ But if all they have shown is that A has some relational property in w, it is not obvious that we can infer that A has that relational property in the actual world. After all, A is knocked at $t+\delta$ in w. Even if the argument shows us something about the properties of A in w, it is not clear what that allows us to conclude about A's properties in the actual world.

Moreover, recall that even if we grant PPJ what they have not established --that some property of A is part of a causally sufficient condition which would be operative in producing A's breaking should A be knocked (i.e., that A has a causal basis for breaking), there has been nothing said so far to show that the causal basis isn't just the disposition. The conclusion still leaves open the possibility that what it is about the object A that causally contributes to A's breaking is A's fragility. That is, PPJ still need to establish the Non-Identity Thesis before they have an argument against bare dispositions.

2.3.2. The Non-Identity Thesis

PPJ offer three main arguments for the thesis that causal bases are not identical to their attendant dispositions. First, I'll consider two arguments from multiple realizability, and then one based on what PPJ call "swamping" the causal basis.⁷⁴

⁷³Ibid., 253.

⁷⁴The Non-Identity, or Distinctness Thesis was previously discussed in Chapter 1 for its part in the "No Work" Argument against the causal efficacy of dispositions. There, I considered the "disjunctive solution" and the "trope solution" to the multiple realizability problem. These approaches obviously won't work here, so PPJ's arguments warrant further consideration.

2.3.2.1. The Arguments from Multiple Realizability

PPJ begin with the observation that a disposition can have different causal bases in different objects. They go on to say:

We cannot say both that being fragile = having molecular bonding α , and that being fragile = having crystalline structure β ; because by transitivity we would be led to the manifestly false conclusion that having molecular bonding α = having crystalline structure β .⁷⁵

PPJ's reasoning seems to be: if one is going to identify a disposition with its causal basis in one case, one must do so in all cases. But since some dispositions can have several distinct causal bases, it is absurd to identify the dispositions with each of these causal bases. However, considering the possibility of bare dispositions, there is no reason to think that, if one is going to identify a disposition with its causal basis, one has to do so in all cases. One could identify bare dispositions with *their* causal bases, while declining to identify other dispositions with their causal bases. What if a disposition like fragility could be bare in some instances, and yet have different causal bases in other instances? In such a case, fragility would be identical to its causal basis in the first instance, but not in the second. There is no reason we have to say that if a disposition is ever identical to its causal basis, then it has to be identical to all of its possible causal bases. PPJ show at most, that in some instances, a disposition is not identical to its causal basis, but this does not preclude the possibility of a disposition being had barely in other instances.

In their second argument, PPJ move from the existential claim that some dispositions are not identical to their causal bases to the modal claim that all dispositions must be. Considering a disposition that has only one causal basis in this world, PPJ claim that:

if "fragility (being fragile) = having α (say)" is true, it is necessarily so, and if false, necessarily so (ignoring worlds where one or the other doesn't exist, if there are any). But there are worlds where fragile objects do not have α , for it is contingent

⁷⁵"Three Theses About Dispositions," 253.

as to what the causal basis of a disposition is. Hence there are worlds where "fragility = having α " is false for the decisive reason that the extensions of fragility and being α differ in that world; and therefore by rigidity it is false in all worlds, including the actual world.⁷⁶

This argument can be set out as follows. Let D be a disposition, and let P be D's causal basis, leaving it open whether P is identical to D.

- (1) It is contingent that P is the causal basis of D.
- (2) Therefore, there is a world in which there is an object x such that:
x has D and x does not have P.
- (3) Therefore, there is a world with respect to which
"having D = having P" is false.
- (4) If "having D = having P" is true, it is necessarily true.
- (5) Therefore, D is not identical to P.

The bare dispositionalist can grant (1). However, (2) just does not follow. What follows from (1) is that there are worlds in which there is an object x that has disposition D, and P *is not a property of x which is causally efficacious for the manifestation of D*. But that is consistent with x *having P*. (1) by itself gives us no reason to suppose that there is *any* world in which some object has D, but not P -- unless of course, we are supposing that P and D are different properties, which would be question-begging.

Consider the argument as applied to a particular example. Suppose that having negative charge is a dispositional property. If "having negative charge = having negative charge" is true, it is necessarily true. Even if negative charge is its own causal basis, perhaps it could have had a different causal basis; there is a possible world in which objects are negatively charged because of some complex structural property. So, it is contingent what the causal basis of negative charge is. So there is a world in which negative charge is not its own causal basis. However, that is not a world in which some object has negative charge, and yet does not have negative charge.

⁷⁶Ibid., 254.

I think that the initial plausibility of this multiple realizability argument trades on the ambiguity of the contested claim: " 'D is identical to its causal basis' is necessarily true." It can mean: (1) "The property which is a causal basis of D is such that, necessarily, D is identical to it." Or it could mean: (2) "Necessarily, D is identical to whatever property fulfills the role of being D's causal basis." The bare dispositionalist can accept (1) while denying (2), and it is only (2) which is subject to the difficulties PPJ raise. I conclude that the arguments from multiple realizability do not show that a disposition cannot be its own causal basis.

2.3.2.2. Swamping the Disposition

PPJ's third argument for the Non-Identity Thesis runs as follows. Even if property P were the only causal basis of some disposition D, a particular object x may have P, but x may have other properties that "swamp" P so that x does not have D. In that case, x would have P without having D. Therefore, $P \neq D$. As PPJ put it:

there is the difficulty that even if there is only one causal basis of fragility, say, bonding α , it may happen that although all fragile objects have α , some objects that have α are not fragile. This would be the case if there were an internal structural property S which swamped the effect of having α .⁷⁷

First of all, it is not clear why this is not just a case of masking.⁷⁸ The fragility of a glass is *masked* when the glass is equipped with internal supports that prevent it from breaking. The causal basis is overwhelmed, so that breaking will not result when the glass is struck. However, in the case of masking, intuitively, the glass still remains fragile. Therefore, this

⁷⁷Ibid., 253.

⁷⁸As discussed in Johnston, "How to Speak of the Colors", and "Dispositions: Predication with a Grain of Salt" (unpublished manuscript).

is not a case where you have the causal basis but not the disposition, and so it is not a counterexample to the claim that the disposition is identical to the causal basis.

But perhaps PPJ take themselves to have described a situation in which an object has the base property α , but fails to have the disposition, fragility. If some objects which have α are not fragile, then α is not a property of the object that, together with the circumstances, would be a causally operative sufficient condition for breaking. But *that* is to say that property α is not the causal basis of fragility. PPJ *define* "causal basis" as the property of an object that, together with the circumstances, would be the causally operative sufficient condition for the manifestation of the disposition.⁷⁹ So, if an object could have some property without having a given disposition, then that property cannot be the causal basis of that disposition. In the scenario described above, the causal basis of fragility has just been misidentified.

As applied to a bare disposition, PPJ's suggestion amounts to the claim that an object x can have a disposition D , but simultaneously x can have some other properties that stop x from having D . I don't know how to make sense of this suggestion. Say some object x is "barely fragile." Suppose x 's properties are changed by adding some fortifying stuff to it, so that it becomes nonfragile. It is not as if, after it has become tough as nails, x has the bare fragility lingering inside of it. If the fragility gets "swamped," then the disposition *and* the causal basis go away. If a disposition is its causal basis, you're never going to be able to lose the disposition and keep the causal basis.

2.3.2.3. Concluding Remarks about The Non-Identity Thesis

If these arguments do not lend themselves to consideration of bare dispositions, perhaps that is because PPJ's target seems to be an identity theorist like Armstrong; their

⁷⁹"Three Theses About Dispositions," 251.

arguments for the Non-Identity Thesis assume that we can distinguish in some way (descriptively, or conceptually, perhaps) between a disposition and its causal basis. But if we are considering a bare disposition which is its own causal basis, we can do no such thing. PPJ define 'causal basis' without making any appeal to non-dispositional, categorical, or micro-structural properties of the thing with the disposition. All they say, basically, is that the causal base is the object's causal contribution to the manifestation of the disposition. Their arguments for the Non-Identity Thesis do not succeed in ruling out the possibility that what it is about the object that causally contributes to the manifestation is just a bare disposition.⁸⁰

2.4. Bare Counterfactuals

Disposition claims bear some important relation to counterfactuals. When something has a disposition, a certain counterfactual is true of that thing. Suppose that I have on my table a normal, water-soluble sugar cube and a cup of ordinary tap water, at room temperature. The prevailing circumstances (atmospheric pressure, laws of nature, etc.) are as one would expect around these parts. Barring inductive skepticism, it is safe to say that if I were to put this sugar cube in this cup of water, it would dissolve. In possible worlds talk, in the closest possible world in which I put the sugar cube into the water, the sugar cube dissolves.

The observation that dispositions are connected with counterfactuals in this way is supposed to cause trouble for bare dispositions for the following reasons. It is reasonable to suppose that true counterfactuals are true because of facts about the actual world.

⁸⁰One might think that this possibility is ruled out by the third of the "Three Theses about Dispositions" -- The Impotence Thesis, according to which dispositions are causally inert. However, the arguments for this third thesis rely on the first two theses. So, unless we have independent reasons for thinking that dispositions are causally irrelevant, PPJ's arguments should not convince us that there can be no bare dispositions.

Given the circumstances described above, the following counterfactual is true: If the sugar cube were placed in water, it would dissolve. But given that I haven't actually placed the sugar cube in the water, the submersion of the cube and the dissolving of the cube are no part of the actual world. What makes the counterfactual true? A plausible answer is that it is something about the sugar cube's properties. The sugar cube is composed of glucose molecules, connected by weak ionic bonds which break when confronted with the bipolarity of H₂O molecules, etc.

But what if the disposition is bare? If there is no distinct causal basis, there are no properties of the sugar cube distinct from solubility which are causally efficacious for its dissolving. A bare disposition is a truly inexplicable causal power. It seems as if the only reason for saying that the object has the bare disposition involves non-actual circumstances and events. Is there anything about the actual world that makes the counterfactual true? If not, that would seem to be a problem for the possibility of bare dispositions. Bare dispositions would seem to run afoul of what C.B. Martin, Armstrong, and others have called the "Truth Maker Principle." Armstrong says "It seems obvious that for every true contingent proposition there must be something in the world (in the largest sense of 'something') which makes the proposition true."⁸¹

However, unless something more is said about what can and what cannot count as something in the world (in the largest sense of 'something') this principle seems vacuous. The bare dispositionalist can say that the fact that something has a disposition is no less a fact about the actual world than any other fact. She can say that the counterfactual "if the sugar cube were placed in water, it would dissolve" is made true by the fact that the sugar cube is water-soluble. Left at this intuitive and abstract level, the argument from Truth Makers does not seem decisive against bare dispositions. Michael Smith and Daniel

⁸¹*Belief, Truth and Knowledge*, 11. See also *Dispositions: A Debate*, 15.

Stoljar offer an argument which promises to sharpen the point suggested by the Truth Maker Principle. However, this argument offers an opportunity for the bare dispositionalist to sharpen her reply as well.

6.4.1. Smith and Stoljar's Argument

Smith and Stoljar begin by noting that some disposition claims are contingent: An object with a certain disposition might have lacked it. That is, there is a possible world in which that object exists, but lacks the disposition. Suppose that:

- (1) x is disposed to give manifestation M in circumstances C .

They give the following semantic analysis of (1):

- (2) " x is disposed to give M in C " is true iff
in the closest x -in- C world, x gives M in C .

Given the disquotation schema:

- (3) " x is disposed to give M in C " is true iff x is disposed to give M in C ,

and from the supposition that (1) is true, they derive:

- (4) In the closest x -in- C world, x gives M in C .⁸²

Smith and Stoljar say that if x 's disposition is a bare disposition, then (1) is not only contingent, but "barely true." That is to say:

If (1) is true of the actual world, say, then there is no further fact about the actual world that makes it true. If someone were to ask "What about the actual world makes (1) true?", the only thing to say is that x is disposed to verb [give M] in C .⁸³

(4) is derived from (1) via the a priori premises (2) and (3). Smith and Stoljar claim that this shows, if (1) is contingent and barely true, then (4) must be contingent and barely true as well. But what, they ask, does it mean to say that (4) is barely true? Consider the worlds @, w_1 , and w_2 :

⁸²"Global Response Dependence...", 98.

⁸³*Ibid.*, 91.

@: x is not in C, and x does not give M.

w₁: x is in C, and x gives M.

w₂: x is in C, and x does not give M.

According to (4), @ is more similar to w₁ than it is to w₂. But if (4) is barely true, there are no intrinsic features of @, w₁, and w₂ which makes it the case that @ is more similar to w₁ than it is to w₂. What this does, according to Smith and Stoljar, is to treat similarity as an extrinsic rather than an intrinsic relation. There is no intrinsic property of @ in virtue of which it is more similar to w₁ than it is to w₂. If bare dispositionalism were true, the relationship between @ and w₁ would be one of "bare similarity." Smith and Stoljar deny that there is any such external relation of bare similarity. Similarity, they claim "is an *internal* relation *par excellence*."⁸⁴ They go on to say:

We ourselves are not sure that *any* external relation between possible worlds that the friends of the Bare Dispositions Theory succeed in characterizing should count as a relationship of *similarity*.⁸⁵

The general point can be characterized as follows. If dispositions can be bare, then there can be counterfactuals that are barely true. If an object had a bare disposition, a certain counterfactual statement would be true, but that statement would not be reducible to or explainable by any categorical facts. If you are going to analyze counterfactuals in terms of similarity to other possible worlds, then you are going to have to say that similarity between possible worlds is not always determined by intrinsic, categorical properties of worlds, and that seems like an odd thing to say.

2.4.2. The Bare Dispositionalist Response

There seem to be basically two ways to respond. One response is to deny that the bare dispositionalist is committed to the bare truth of (4), "In the closest x-in-C world, x

⁸⁴Ibid., 97.

⁸⁵Ibid., 99.

gives M in C." I gave an account of what it means to say a disposition is *bare*. A bare disposition is a disposition with no distinct causal basis. There is no obvious route from there to understanding what it means to say that a statement is *barely true*. As we have seen, Smith and Stoljar elaborate on the claim that (1), "x is disposed to give M in C," is barely true as follows:

If (1) is true of the actual world, say, then there is no further fact about the actual world that makes it true. If someone were to ask 'What about the actual world makes (1) true?', the only thing to say is that x is disposed to verb in C.⁸⁶

It is not obvious that that is the only thing to say. There might be a number of ways to respond to such a question. Consider the "barely" soluble sugar cube. What about the world makes it true that this sugar cube is disposed to dissolve if I put it in this cup of water? Well, for one, the sugar cube is not encased in some water proof coating. For another thing, the water is not frozen, but is at 72 degrees. On another reading, the question could be asking: "what brought it about that x has this disposition?" There might be all manner of things to say about how and why the sugar cube acquired the bare disposition that it has. Perhaps God endowed the sugar cube with this bare disposition. Or, the question might mean: "in virtue of which underlying properties does x have the disposition?" However, if we are supposing that x has a bare disposition, this question has a false presupposition. It is not clear that if something has a bare disposition, any sentence saying so is a sentence which is barely true.

Moreover, even if we grant that "x is disposed to give M in C" is barely true, it does not follow that (4) is barely true as well. Smith and Stoljar's semantic analysis of the disposition claim is in fact a conjunction of two separable analyses, a counterfactual analysis of dispositions, and a possible worlds semantics for counterfactual statements. Smith and Stoljar's argument relies on the plausibility of the claim that the counterfactual analysis of dispositions is a priori true. That is, it relies on the plausibility of the claim that

⁸⁶Ibid. 91.

(2) " 'x is disposed to give M in C' is true iff in the closest x-in-C worlds, x gives M in C" is an a priori truth.⁸⁷ I myself am inclined to agree with Smith and Stoljar that disposition statements can be analyzed in terms of counterfactuals, but the bare dispositionalist is not obviously *forced* to accept that analysis. There are notorious problems for giving a satisfactory counterfactual analysis of dispositions.⁸⁸ A weaker, intuitively plausible claim is that if something has a disposition, other things being equal, a certain counterfactual is true of that thing. But that is not to say that disposition statements are a priori equivalent to counterfactuals. Smith and Stoljar need (2) to be a priori in order for (4) to follow a priori from (1).

Finally, even if we grant that (4) follows a priori from (1), it is not obvious that (1)'s being barely true entails that (4) is barely true as well. It is not clear that bare truth transmits over a priori entailment. To take an example from Evans, if we give the name 'Neptune' to whatever it is that causes the perturbations in the orbit of Uranus, then "Neptune is Neptune" a priori entails "Neptune is what causes the perturbations in the orbit of Uranus."⁸⁹ Plausibly, "Neptune is Neptune" is barely true. However, it is not plausible that "Neptune is what causes the perturbations in the orbit of Uranus" is barely true, for there are surely many facts about our solar system and the laws which govern it that make the second sentence true.

⁸⁷Ibid., 98.

⁸⁸ David Lewis, "Finkish Dispositions," *The Philosophical Quarterly* 47.187 (1997): 143-158; Martin and Johnston, op. cit.

⁸⁹ Evans, "Reference and Contingency," in *Collected Papers* (Oxford: Clarendon Press, 1985), 192.

2.4.3. Biting the Bullet

We have looked at ways in which the bare dispositionalist could resist having to say that (4) is barely true. Now let's look at how the bare dispositionalist could respond if she *accepts* that (4) is barely true. She could admit that barely true counterfactuals and extrinsic similarities among possible worlds are consequences of allowing bare dispositions into one's ontology. Dispositionalism is one of a host of related modal notions, including necessity, possibility, laws, and causal efficacy. We can speak more broadly of the class of modal properties. To say there can be bare dispositions is to say that a thing can have a modal property irrespective of its other properties. Broadening the picture, that would be to say that a modal property of some world does not depend on any other intrinsic features of that world. To insist that, if a certain counterfactual is true at a world it must be on the basis of some non-modal properties of that world, is just another way of denying that bare dispositions are possible. Similarly for Smith and Stoljar's claim that:

if one possible world is similar to another, this must be explained by the intrinsic features of the possible worlds in question. ...But to say that similarity must be explained in terms of intrinsic features of possible worlds is to insist that it be explained... in terms of the non-dispositional properties...⁹⁰

That seems like another way of saying that the modal needs grounding in the non-modal. If my hypothesis can be put "there can be a modal property that is not grounded in non-modal properties," then this has just been denied without argument.

One consideration that Smith and Stoljar raise is that barely true counterfactuals run afoul of the principle that similarity is an internal relation. I'll grant them for the sake of argument that modal properties are relational properties -- they depend on which possible worlds are nearby. However, intuitively, it is not at all clear why two things can't be similar in respect to their relational properties. I'm similar to Ned in that we are both

⁹⁰"Global Response Dependence...", 96.

within a mile of Boston, and that we are both shorter than Alex. Do Smith and Stoljar mean to be denying such claims when they say that similarity is "an internal relation *par excellence*"? And if people can be similar in respect to their relational properties, then why not worlds? There doesn't seem to be anything wrong in principle with talking about worlds being similar to one another with respect to their relations to other worlds. For example, there doesn't seem to be anything wrong with saying that some worlds can be similar to one another in the respect that they are more similar to the actual world than to some far off world.

With this in mind, let's reconsider the possible worlds Smith and Stoljar described.

@: x is not in C, and x does not give M.

w₁: x is in C, and x gives M.

w₂: x is in C, and x does not give M.

The challenge posed to the bare dispositionalist is to say why @ is more similar to w₁ than it is to w₂. Granted, she can't say that they are similar with respect to their intrinsic, categorical properties. But she can say that @ and w₁ are similar in the respect that they are both more similar to other x-Ming-in-C worlds than they are to x-not-Ming-in-C worlds. Similarly (or perhaps equivalently), she can say that in @ and w₁, the counterfactual statement "If x were in C, x would give M" is true. She can say that @ and w₁ are similar in the respect that, in both worlds, x has the disposition to give M in C.

2.4.4. "Bare Truth" Revisited

Perhaps Smith and Stoljar's arguments suffer from a less than clear characterization of what it means to say that a statement is barely true. Michael Dummett gives an alternative account of bare truth in terms of reducibility: "A statement is barely true if it is true, but there is no class of statements not containing it or a trivial variant of it to which

any class containing it can be reduced."⁹¹ One suggestion is that while ordinary disposition statements can be given such a reduction, bare disposition statements cannot. An ordinary (non-bare) disposition statement could be reduced to statements about its causal basis, or could involve translating them into counterfactual statements and then cashing these counterfactual statements out in terms of a less problematic notion, resemblance --that is, resemblance among categorical properties of possible worlds.⁹²

There are two questions which arise with respect to this conception of bare truth. First, are bare disposition statements irreducible in this sense? Plausibly, the answer is yes. Second, is this a problem? Well, it is not clear that it is. If bare disposition statements are irreducible, arguably, they are in good company. If we are to reduce all statements containing problematic modal notions such as dispositions, causes, laws, we need some set of non-modal facts to reduce them to. Unless we can specify a set of facts that will do the job, reductionism is at best a promissory note. As Stalnaker points out:

a reductionist program presupposes that the causal dependencies between events and the causal powers of things in a possible world derive from *relational* properties of the possible world, properties defined in terms of the way the possible world resembles other possible worlds... the project requires isolation of a level of pure categorical particular fact relative to which possible worlds are compared.⁹³

The familiar form of reduction of macro-properties to microphysical properties, which is itself a promissory note, would be inadequate to the task of reducing all modal notions.

⁹¹"What is a Theory of Meaning II", in Evans and McDowell, *Truth and Meaning*. (Oxford: Clarendon Press, 1976), 94. He says that it comes to the same thing as holding "that we cannot expect a non-trivial answer to the question 'In virtue of what is a statement ...true when it is true' " (94). Robert Stalnaker points out in *Inquiry* that these do not come to the same thing; you might not be able to reduce a statement, and yet there might be some non-trivial answer to the question 'in virtue of what is the statement true?'

⁹²This is how Stalnaker characterizes Lewis' view, (*Inquiry*, 155-160).

⁹³ *Ibid.*, 157.

This is because, as I suggested earlier, contemporary scientific characterization of the ultimate constituents of matter seems to be rife with causal and dispositional notions.

If we reject the idea that we are in some sense required to give a reductive analysis of modal statements in terms of non-modal statements, the argument against bare dispositions looks considerably weaker. The bullet-biting response to Smith and Stoljar's argument, which rejects this demand outright, seems to be the right one after all. In sum, it is not clear that the existence of bare dispositions would lead to barely true counterfactuals. And furthermore, even if it did, maybe that is something a bare dispositionalist could happily live with.

2.5. Conclusion

I have considered several arguments against the possibility of bare dispositions: the Conceptual Argument; Prior, Pargetter, and Jackson's multiple realizability and swamping arguments for the Non-Identity Theses; the argument from the Truth Maker Principle; and Smith and Stoljar's Bare Counterfactuals argument. I have explained why I do not think these arguments work. As far as I can see, we have no a priori reason to suppose that there are no bare dispositions.

Chapter 3: Extrinsic Dispositions

In his 1997 paper "Finkish Dispositions," David Lewis appeals to the following principle: "If two things (actual or merely possible) are exact intrinsic duplicates (and if they are subject to the same laws of nature) then they are disposed alike."⁹⁴ The view that dispositions are an intrinsic matter is widely accepted by those who discuss dispositions. In a recent article, George Molnar says "Dispositions are intrinsic properties of their bearers. This is one of the crucial appearances which has to be saved by an analysis."⁹⁵ Considering a fragile glass, David Armstrong says that "the possession of the disposition must depend on upon the *non-relational* properties of the glass."⁹⁶ Similarly, according to Mark Johnston, a disposition must have its basis in the intrinsic properties of the disposed object. Johnston says: "We may therefore think of a constituted disposition as a higher-order property of having some *intrinsic* properties which, oddities aside, would cause the manifestation of the disposition in the circumstances of manifestation."⁹⁷

The idea that dispositions are intrinsic properties has initial plausibility. For example, if two wine glasses came off the assembly line, sharing all their intrinsic properties, it is natural to think that they would share all their dispositions. They would

⁹⁴"Finkish Dispositions," 148.

⁹⁵"Are Dispositions Reducible?" *Philosophical Quarterly*, 49 (1999), 3.

⁹⁶*Belief, Truth, and Knowledge*, 12.

⁹⁷"How to Speak of the Colors," 234. As will be discussed later, there is a sense in which Lewis thinks that dispositions are extrinsic. Strictly speaking, Johnston's and Armstrong's statements do not commit them to dispositions being intrinsic, and their full accounts are somewhat more complicated. However, what's relevant for the purposes of this paper is that all of the above mentioned philosophers would agree to the following: keeping the laws of nature fixed, an object can neither lose nor gain a disposition merely as the result of extrinsic changes.

share the same disposition to break when struck; they would both be fragile. This is thought to be true of dispositions in general.

In this chapter, I take issue with this view. I argue that there are some dispositions that an object can gain or lose without undergoing intrinsic change. In other words, intrinsic duplicates can differ with respect to having certain dispositions. This claim has a number of important implications. First of all, it tells us something about the nature of dispositions, and rules out certain analyses of dispositions according to which they are necessarily intrinsic properties of disposed objects. The claim that dispositions are necessarily intrinsic is wielded against the counterfactual analysis of dispositions, for example. According to the counterfactual analysis, a disposition statement is equivalent to some counterfactual statement. The opponents of the counterfactual analysis offer a number of counter-examples which go to show the following: which counterfactuals are true of a thing can vary merely by changing that thing's environment. If this feature is denied of dispositions, it is easy to see how a disposition statement and the associated counterfactual could come apart.

Furthermore, the view that dispositions are sometimes extrinsic poses new challenges to the claim that dispositions can be causally efficacious. Some philosophers say that a disposition is causally efficacious because it is identical to some intrinsic, first-order property of the disposed object. However, if the disposition is extrinsic, that move is not an option. More generally, it may seem puzzling, and at least some story needs to be told, about how extrinsic properties can play a causal role.

Beyond the relevance to contemporary debates about dispositions, the question of extrinsicness has implications for issues in philosophy of mind. At least some mental properties are dispositional, and they also have particular contents. Think of someone with hydrophobia, a fear of water. This person is disposed to panic when she believes she is near a large body of water. However, if externalism about the mental is correct, her believing that she is near a large body of water is an extrinsic matter -- it depends upon her

standing in the correct relationship to water (not twin water). So, the following three theses are inconsistent: Dispositions are intrinsic properties; mental properties are dispositional; and mental properties are extrinsic properties. Something has to give, and I suggest that it is the intrinsicness of dispositions.

I proceed as follows. First I suggest some criteria by which we may determine if a property is dispositional. Then I explain what I mean by "intrinsic," "extrinsic," and "relationally specified" properties. I go on to explore a number of uncontroversial ways that dispositions can be said to be relationally specified, in order to distinguish these claims from the stronger claim I am making. I then argue that some dispositions are extrinsic in a more controversial sense; that is, they are such that, keeping the laws of nature fixed, an object could gain or lose that disposition without undergoing any intrinsic change. I consider the view that it is part of the concept of a disposition that it is an intrinsic property. Then I consider an argument offered by David Armstrong that dispositions are necessarily intrinsic, and show why I think it falls short of establishing its desired conclusion. I conclude by offering a number of examples of extrinsic dispositions.

3.1. Marks of Dispositionality

In Chapter 0, I outlined some general characteristics of dispositions, and gave some paradigm examples. These should be very familiar by now. These characteristics, or marks of dispositionality, can indicate that a given property is a disposition. These characteristics are widely accepted by philosophers who discuss dispositions, even if they argue about other aspects of dispositions.

First of all, a disposition has a characteristic manifestation. There is some event-type that is associated with a disposition which occurs when the disposition is "triggered." The manifestation need not occur for something to possess the disposition. The manifestation of fragility is breaking. The manifestation of solubility is dissolution.

Secondly, a disposition's manifestation occurs in certain types of circumstances, or circumstances of manifestation. Fragile things break when they are struck. Soluble things dissolve when they are placed in solvents.

Thirdly, when something has a disposition, certain counterfactuals are true of it. If the fragile glass were dropped, it would break. If the soluble tablet were submerged in water, it would dissolve. As Elizabeth Prior says, "What is commonly accepted by *all* those who discuss dispositions is that there exists a conceptual connection between a statement attributing a disposition to an item and a particular conditional."⁹⁸ I'm not claiming that every disposition statement is a priori equivalent to a particular counterfactual statement, only that every disposition is associated with a counterfactual which is typically true of things with that disposition.

Fourthly, a disposition can be referred to by an *overtly dispositional locution*. Fragility = the disposition to break when struck. Solubility = the disposition to dissolve in a solvent. Mark Johnston says "the concept of the property F is a concept of a dispositional property just in case there is an a priori property identity of the form: The property F = the T disposition to produce R in S under C."⁹⁹

⁹⁸*Dispositions*, 5. .

⁹⁹"How to Speak of the Colors," 229. While I suppose that this is a sufficient condition for a property to be a disposition, I do not think it is necessary. If the dispositional/categorical distinction is truly one among properties, and not merely predicates, it would seem that a property may be a disposition, even if speakers of a language do not take their name for that property to be equivalent in meaning to any overtly dispositional locution. (The dispositional theory of color is not doomed by the fact that most people do not take color terms to be equivalent in meaning to overtly dispositional locutions.)

One may have reservations that this is a sufficient condition for the following reason. Just because speakers of English take some property term to be equivalent to "the disposition to so and so," the property they are actually referring to might fail to be a disposition. This might be so because they are mistaken about the property being a disposition, or for some bizarre reason just call the property "the disposition to so and so" even though it does not satisfy that description. (Just because something is referred to by

If a property exhibits enough of these characteristics, then there is good reason to think it is a disposition. Consider a property. Does it have a characteristic manifestation? Are there certain circumstances that would bring about the occurrence of this manifestation? Is there a certain counterfactual that is typically true of things that possess this property? Can the property be named by an overtly dispositional locution? If one can find plausible candidates for the manifestation, circumstances of manifestation, the associated counterfactual, and an overtly dispositional locution which names the property, then it is plausible to think that the property is a disposition.

3.2. Intrinsic, Extrinsic, and Relationally Specified Properties

Intuitively, a property is intrinsic if it is a property that a thing has or lacks regardless of what is going on outside of itself. Intrinsic properties are necessarily shared by perfect duplicates. Both of these ways of characterizing intrinsicness suffer from problems of circularity. Consider x's property F. One thing that is going on outside of x is that things are accompanied by something that has F. So, "what is going on outside" must be understood as what the outside is like *intrinsically*. And how are we to understand "perfect duplicates" but as things that share all of their *intrinsic* properties? How to characterize intrinsicness is a difficult philosophical problem I will not attempt to solve here, but I think we can make due with an intuitive understanding for the purposes of this paper.¹⁰⁰

a definite description, or by an expression that is a priori equivalent to a definite description, it doesn't follow that it satisfies that description.)

¹⁰⁰For a discussion of intrinsic properties see Langton, Rae & Lewis, "Defining 'Intrinsic'," *Philosophy and Phenomenological Research* 58, 1998; Yablo "Intrinsicness," forthcoming in *Philosophical Topics*.

Extrinsic properties, by extension, are simply those that are not intrinsic -- if a property is *extrinsic*, then whether a thing has or lacks the property *can* depend on what is going on outside of itself. Perfect duplicates *can* differ with respect to their *extrinsic* properties.

A type of properties that I want to distinguish from extrinsic properties are *relationally specified properties*. A property can be specified relationally even if it is an intrinsic property. "Jennifer's favorite property" is specified relationally, but it may be an intrinsic property of the things that have it. Say my favorite property were roundness. Even if we refer to roundness relationally as "Jennifer's favorite property," roundness might be something that an object can neither gain nor lose due to changes in its environment. It seems that any property can be specified relationally, and so this doesn't mark an ontological distinction among properties, but a semantic distinction among property terms. A property is relationally specified relative to a description, or a way of picking it out. A property expression can make reference to extrinsic factors, or rely on a broader context of utterance, in order to successfully refer to a property.

3.3. Dispositions Are Relationally Specified

First, I will explain a number of ways in which dispositions are relationally specified, in order to distinguish my more controversial claim that some dispositions are extrinsic. Canonically, dispositions are specified by reference to causal relations to other states of affairs. Consider the case of dormitivity, the disposition to induce sleep upon ingestion. Here we have reference to two states of affairs that are extrinsic to the object with the disposition: the circumstances of manifestation -- ingestion by a creature of the relevant sort, and the manifestation itself -- the slumbering of said creature. Everyone should grant that the actualization conditions, or circumstances of manifestation, are often extrinsic to the object with the disposition. Whether or not the manifestation occurs depends upon the environment of the disposed object. But that is not to say that the

disposition itself is extrinsic. We mustn't confuse a disposition's circumstances of manifestation with the circumstances of its *possession*.

Another way that a disposition can be relationally specified is to be a member of a pair of dispositions that work in tandem. As David Lewis says:

Dispositions ... can come in pairs: x is disposed to respond to the presence of y, and y is disposed to respond to the presence of x, by a response r given jointly by x and y together. ... we can express this by a counterfactual: if x and y were to come into one another's presence, they would jointly give response r. ... For example, I and a certain yellow disc are so disposed that if I and it came together, it would cause in me a sensation of yellow. We could say it is disposed to influence me; or that I am disposed to respond to it. Or both. Or we could say that the two-part system consisting of me and the disc is disposed to respond to the coming together of its parts.¹⁰¹

So, x has the disposition to produce r, but not all by itself; y is also needed. This seems to be just a matter of specifying the circumstances of manifestation. Perhaps all dispositions could be cast in these terms. A dense object, such as a hammer, has the disposition to produce shatterings of fragile objects, in the event that the hammer and the fragile object enter into a certain relationship, namely, a high speed collision.

In any event, even if y is mentioned in specifying x's disposition, x's disposition might still be an intrinsic property. More generally, if a disposition D1 has circumstances of manifestation which involve some other thing having disposition D2, D1 might still be an intrinsic property -- it might be the case that all intrinsic duplicates share or lack D1 regardless of their proximity to objects with D2. Or, if you are going to say that two objects jointly have a disposition, there is a sense in which that disposition can still be intrinsic to those two objects considered as a system, or mereological sum.

A disposition can also be relationally specified if the referent of a particular disposition term is context sensitive. If a property term is context sensitive, which

¹⁰¹"Finkish Dispositions," 144-45.

property is picked out will depend on extrinsic factors, usually the context of utterance. What is considered fragile on a construction site might not be considered fragile in a China shop. The expression "fragile" does not have a single fixed extension, but one that varies across contexts of utterance. Whether or not something is considered fragile is relative to a certain comparison class. One might say, as a very old house is being moved to a new location, that this procedure has to be done quite carefully, because the house is very *fragile*. More than likely, it would not be prone to break when subject to same force that breaks crystal glassware. The old house is *fragile for a house*.

Consider an analogy with other properties which make implicit reference to a comparison class, such as tallness. We can say that Spud Web is not tall, relative to the class of NBA basketball players, but that is not true if the reference class is the general population. If we say that Spud Web is not tall, in most contexts it will be understood that we mean he is not tall *for a basketball player*. But once the comparison class is fixed, by the conversational context or whatever, we can talk about the property of being tall for a basketball player, and ask of *that property* whether it is an intrinsic or an extrinsic property. Could someone cease to be tall for a basketball player, merely by changing his environment? Perhaps, if the composition of the NBA changed so as to be populated by taller people. However, that is essentially to change the comparison class. We can keep the contrast class fixed while considering a change in environment. Keeping the comparison class fixed, it does not seem that someone could cease to be tall without undergoing some intrinsic change.¹⁰²

Back to the case of dispositions, once the referent of a disposition term is specified with the help of a certain context or comparison class, we can ask, of that property so

¹⁰²A slight complication: If space is relative, there may be no difference between (a) my getting larger while everything else stays the same, and (b) my staying the same while everything in the universe shrinks.

specified, is it an intrinsic or an extrinsic property? In the case of a crystal wine glass, we can consider its property of being fragile for a drinking vessel, or what have you. Could the wine glass cease to have this property merely by changing its environment? If not, then fragility is an intrinsic property.

The context of utterance serves to determine the reference of disposition terms in other ways. Often, it is implicit in the meaning of a disposition term that the circumstances of manifestation involve standard background conditions -- the way things are around here most of the time. When we talk about things that are disposed to shatter when struck, we don't usually mean to include all the things, including steel bars, which would shatter when struck in circumstances of extremely low temperatures. This implicit assumption can be overridden by a more precise description of the disposition.

Sometimes the circumstances of manifestation of a disposition will implicitly involve a particular reference class. For example, cyanide is not poisonous simpliciter, but poisonous to a certain class of creatures. Perhaps on some other planet, the lakes are filled with cyanide, and it is a staple part of the diet of the life forms there. We could say "on planet X, cyanide is not poisonous." So whether or not cyanide is poisonous might seem to depend on what creatures are in the environment. However, as above, we can fix the relevant reference class. We can specify the property of being poisonous *to humans* (actual humans). Then, if we consider the cyanide off on this other planet, we can see that, like the cyanide on earth, it has the property of being poisonous to humans, even though there are no humans in that environment.¹⁰³ The class, relative to which the

¹⁰³Suppose all actual humans were genetically altered in some way that rendered the effect of cyanide harmless. It would seem that then, cyanide would no longer be poisonous to humans. Does this make poisonous to humans an extrinsic property? Or does this example involve a change in the reference class, and hence a shift in meaning of the disposition term? Surely not every change in the properties of the members of the reference class results in a change in meaning of the disposition term. Poisonousness to humans might turn out to be extrinsic after all.

disposition is picked out, can be specified without respect to the disposed object's actual environment.

These examples show the ways in which dispositions are relationally specified. Perhaps for many dispositions, in order to define or specify it, we have to make reference to things that are extrinsic to the thing with the disposition --the manifestation, circumstances, the context, the comparison class, etc. However, once we have specified the property, it is perfectly compatible with a disposition being relationally specified that the property we have succeeded in referring to is in fact an intrinsic property. Showing that a disposition is picked out relationally does not go to show it is extrinsic.

3.4. A Possible Qualification

So, the preceding considerations do not show that dispositions can be extrinsic. However, one reason for thinking that dispositions are extrinsic, advocated by David Lewis, is the following: which dispositions an object has depends not just on the intrinsic properties of that object, but also on the prevailing laws of nature.¹⁰⁴ Imagine taking some object and transporting it to another possible world which has different laws of nature, such that the object's dispositions would change. Or, consider two intrinsic duplicates, each in a different possible world where different laws of nature prevail. We can imagine that the two objects possess different dispositions. In this world, the glass is disposed to break when struck. In another possible world, perhaps it is disposed to melt when struck.

It seems right that the possession of dispositions is not independent of the laws of nature. However, what is not clear is whether the possession of *intrinsic properties* is independent of the laws of nature. Lewis is supposing that something could be

¹⁰⁴David Lewis, *Philosophical Papers*, vol. II (New York: Oxford University Press, 1986), ix.

intrinsically the same if the world had radically different laws. What would such a world be like? Are there gravitational forces? Are the same chemical compositions found? Do particles exert and respond to the same forces? Or do electrons go flying away from protons? If that is so, in what sense can we say that a crystal glass in that world is intrinsically the same as it is in this world? It would seem that we need to consider worlds in which the laws are similar enough to ours for intrinsic duplicates to be able to exist. However, it is not clear that such a world would be one in which things have different dispositions.

The idea that in another possible world there could be a perfect duplicate of a fragile crystal glass that is hard as nails, or a perfect duplicate of an elastic rubber band that is not elastic, involves some substantive metaphysical assumptions. It involves assuming that there is some set of core, categorical "true-intrinsic" properties, that can be considered apart from the forces, powers, and causal relationships that a thing has, under a certain set of laws. However, I'm going to put this matter aside, and focus on the question of whether dispositions are intrinsic-keeping-laws-of-nature-fixed, which I will just call "intrinsic."

3.5. The Case for Intrinsicness

I will consider two arguments that dispositions are intrinsic, The Conceptual Argument, and (for lack of a better name) Armstrong's Argument.

3.5.1. The Conceptual Argument

One may argue as follows. It is part of the concept of a disposition that it is an intrinsic property. This is supported by the etymology. The word 'disposition' is derived from 'ponere', as are the words 'pose' and 'position'. Saying something has a disposition is to say something about the way it, or its parts, are arranged. If a thing is prone to act in a certain way because of the way its parts are arranged, then it has a disposition. However,

if something is prone to act in a certain way for some other reason, then we don't say it has a disposition.¹⁰⁵

In response, even if this is the right way to interpret the etymology of 'disposition', it is not clear how that should govern our current understanding of the concept. After all, 'terrific' has the same root as 'terrible,' but they have come to mean very different things.

Some people have the opposite intuition about the concept of a disposition. On some views, to say that something has a disposition is to say something about what it would do in certain circumstances, but not to say anything about what it is like intrinsically. As Gilbert Ryle puts it in *The Concept of Mind*: "to classify a word as signifying a disposition is not yet to say much more about it than to say that it is not used for an episode."¹⁰⁶ (116) Even Armstrong, who thinks dispositions are intrinsic, says "dispositional concepts leave us in ignorance concerning the properties of the disposed object which give it that disposition."¹⁰⁷ Armstrong thinks that the intrinsicness of dispositions is something he needs to argue for, not something that falls out of the concept of a disposition.

Like many expressions, the word 'disposition' doesn't have a fixed and precise usage in ordinary English. In philosophical discussions, it is something of a term of art. I have outlined a characterization of dispositions that seems to be standard in the philosophical literature. On this understanding, having a disposition is not distinct from being prone, having a power, a tendency, or a potentiality. One could insist that we add to the characterization "...and the property has to be intrinsic." We could restrict our usage of the expression 'disposition' to properties so specified. However, as I will try to

¹⁰⁵This argument was put forth by Fred Feldman in personal correspondence.

¹⁰⁶Gilbert Ryle, *The Concept of Mind* (New York : Barnes & Noble, 1949) 116.

¹⁰⁷*Beleif, Truth, and Knowledge*, 13.

convince you, there is an interesting group of properties which have manifestations, circumstances of manifestation, and associated counterfactuals, that are not *intrinsic*. It seems natural to me to use the expression 'disposition' to encompass this broader class.

3.5.2. Armstrong's Argument

David Armstrong offers the following argument for the thesis that dispositions are intrinsic properties:

Consider an occasion where a disposition is manifested. A brittle piece of glass is struck and, as a result, it breaks. In any causal sequence the nature of the effect depends upon three things: the nature of the cause; the nature of the circumstances it operates in; the nature of the thing it acts upon. The glass breaks because it is struck, it is not carefully packed around with protective material, and it is brittle. Cause of a certain nature + circumstances of a certain nature + disposition = effect of a certain nature. Now a disposition is something which the disposed thing retains in the absence of both a suitable initiating cause and of suitable circumstances for the cause to operate in. A brittle piece of glass is still brittle, even although it is not struck and is so packed around with protective material that striking would not cause breaking. But the presence or absence of the initiating cause, and the presence or absence of suitable circumstances for its operation, are the only *relational* properties of the piece of glass which are relevant to its breaking or not breaking. The possession of the disposition must therefore depend upon *non-relational* properties of the glass.¹⁰⁸

I'm going to ignore the distinction, if there really is one, between a disposition being an intrinsic property, and a disposition depending on intrinsic properties. I am not arguing that a thing's intrinsic properties make no difference whatsoever to its dispositions, only that a thing's *extrinsic* properties *can* make a difference to its dispositions. The crucial claim I want to dispute is that something can neither gain nor lose a disposition merely by changing its environment, and clearly this is something that Armstrong holds. Armstrong's argument can be set out as follows:

¹⁰⁸Ibid., 11.

(1) When a fragile glass shatters, the fragility, the striking, and the background conditions (including a lack of protective material on the glass) together result in the shattering.

(2) An object need not manifest a disposition to have it. An object can have a disposition outside of the circumstances of manifestation.

(3) A fragile glass that is packed with protective material and unstruck differs from one that is unprotected and struck only with respect to background conditions, and the triggering cause. All the relevant extrinsic properties are different, and yet the glass is still fragile.

(4) Therefore, the fragility of the glass depends upon the intrinsic properties of the glass.

(5) Therefore, (since there is nothing special about fragility, we can conclude in general that) dispositions depend upon intrinsic properties of the things that have them.

The following reply to Armstrong has been offered by A. D. Smith:

A sturdy wooden block is taken on a space flight to Neptune. The inner constitution of the block does not change at all, but when the astronauts who are flying the spacecraft arrive on Neptune and one of them inadvertently knocks the block to the ground it shatters. ...we can certainly make sense of the supposition [that the block of wood is now fragile]. What does this example show us about Armstrong's argument? I think we must agree with him that if an object acquires or loses a dispositional property at t then there must have been some change in the world at t , and indeed, a 'categorical' change which is the 'basis' of the acquisition or loss of the dispositional property. What our example shows is that Armstrong goes too far in insisting that the change in question must concern the intrinsic nature of the object in question. In our example the change consists in the translation of our block of wood for the earth's environment to one which causes the wood to behave surprisingly different from the way it behaves on earth.¹⁰⁹

I do not think that this reply succeeds in refuting Armstrong's argument. Smith's reply to Armstrong involves an equivocation on the sense of fragility that is in play. There is not a

¹⁰⁹A. D. Smith, "Dispositional Properties," *Mind*, 86 (1977), 441.

single property, fragility, such that the block of wood lacks it on earth, but gains it on Neptune. Recall that a disposition is relationally specified, by reference to the salient circumstances of manifestation. The referent of a disposition term is determined in part by the prevailing circumstances, or background conditions which are part of the circumstances of manifestation. For example, if you say something is disposed to break when dropped, you do not mean that is disposed to break when dropped in near zero gravity conditions. Again, one does not normally say that steel is fragile, but when chilled to low enough temperatures it will shatter when struck. A full specification of the circumstances of manifestation would include background conditions, such as ambient temperature.

Consider two very different sets of circumstances, C1 and C2. Something can be disposed to shatter in C1, and a very different object may be disposed to shatter in C2. One may apply the term "fragility" to both of these dispositions, but two different properties are in fact being denoted. If an object which is disposed to shatter in C2 is moved from C1 to C2, it does not suddenly gain the disposition to shatter in C2. It had that all along. The wooden block didn't acquire the disposition to shatter when dropped on Neptune when it arrived on Neptune. It had that all along.

So, I don't think that Smith's reply to Armstrong works. However, I do think that Smith is putting pressure on the right place. Armstrong has described a case in which some changes in the extrinsic properties of an object do not make a difference to its disposition. However, he fails to show that no extrinsic changes could make a difference to the glass's fragility. His claim that only the striking and the lack of protection are the only relevant extrinsic properties of the glass is open to doubt, and this is just what Smith is denying. So, Armstrong's move from (3) to (4) is problematic.

Furthermore, Armstrong's move from his specific example to his generalized conclusion is hasty. Even if we grant Armstrong that he has described a case in which no extrinsic properties make any difference to the possession of the disposition, all his

opponent needs is *one* case in which extrinsic properties *do* make a difference to the possession of a disposition. It seems to me that Smith hasn't provided such a case, but I do think there are examples in the neighborhood. If Armstrong had considered another type of example, as I will consider below, then the analogue of premise (3) would not look at all plausible. In such a case, only extrinsic properties would be different, and the object would thereby lose the disposition.

As I will try to show, some dispositions are such that objects can lose or gain them as a result of changes of environment. Lewis offers the intuitively plausible example of two glasses that come off the same production line, intrinsic duplicates. It is hard to doubt that they are equally fragile. But this is a case where philosophers' limited repertoire of examples is skewing the analysis of the concept, or so I shall argue. The paradigm cases of dispositions, such as fragility and dormitivity, are not as representative of all dispositions as they have been taken to be.

3.6. Examples of Extrinsic Dispositions

Example 1: The Power to Open a Door

A familiar example of an extrinsic disposition comes to us from Sydney

Shoemaker:

A particular key on my key chain has the power of opening locks of a certain design. It also has the power of opening my front door. It could lose the former power only by undergoing what we would regard as a real change, for example, a change in its shape. But it could lose the latter without undergoing such a change; it could do so in virtue of the lock on my door being replaced by one of a different design. Let us say that the former is an intrinsic power and the latter is a mere-Cambridge [extrinsic] power.¹¹⁰

First question: Does the key have a disposition to unlock a certain door? That seems hard to deny. It has a manifestation, the opening of the door, which occurs in the

¹¹⁰Shoemaker, *Identity, Cause, and Mind*, (London: Cambridge University Press, 1984), 221.

circumstances of manifestation, the insertion and turning of the key. We can also presume that it is generally true of the key that, if it were placed in the lock and turned, the door would open. This property doesn't correspond with a simple predicate, but we can successfully refer to it by an overtly dispositional locution, "the disposition to open door x."

Second question: Is this property extrinsic? Yes. As Shoemaker points out, it is a property that the key can lose merely by changing its environment, i.e., by changing the lock on the door.

Objection: The key has a disposition to open locks of a certain type, not a disposition to open a certain door. When we talk about the key's disposition or power, we mean its disposition to open locks of certain type. The disposition is merely relationally specified by reference to a certain door. Its disposition to open locks of a certain type does not change when the lock on a certain door changes.

Reply: It is true that the key has a disposition to open locks of a certain type, and this is an intrinsic property. However, the key also has another disposition, the disposition to open a particular door. Not only do we talk about a key's extrinsic disposition (though not ordinarily in those terms), for most key users, the extrinsic disposition is the most salient. I am most concerned that the key to my front door retains its power to open *my front door*; its disposition to open locks of a certain type is merely instrumental to that end. (Suppose the lock on my door and my key were simultaneously altered, but the key still opens the lock. My key will no longer open locks that are the way that my lock used to be, so my key lost the power to open locks of a certain type. However, that does not concern me, because it still has the power to open my front door.) Moreover, even if it were uncommon to do so, we could specify such a property of the key, and this property is an extrinsic disposition.

Example 2: "Weight"

In his recent paper "Intrinsicness," Steven Yablo presents the following example of an extrinsic disposition.¹¹¹ Let's assume the following account of weight:

x has *weight* n iff x has a disposition to depress a properly constructed scale so as to elicit a reading of n pounds in the local gravitational field.

This can be contrasted with mass:

x has *mass* n iff x has a disposition to depress a properly constructed scale so as to elicit a reading of n pounds in a gravitational field of strength *f*.

Whether this is a satisfactory account of weight and mass is not the issue. The properties referred to by the overtly dispositional locutions on the right of the biconditional can be called "weight" and "mass" for short, even if these overtly dispositional locutions fail to capture what we ordinarily mean by "weight" and "mass."

Is the disposition to depress a properly constructed scale so as to elicit a reading of n pounds in the local gravitational field a disposition? Of course. Is this property extrinsic? Yes. While mass may be intrinsic, weight is extrinsic. A person's weight on earth is different than her weight on the moon. Moving from the earth to the moon, she can remain intrinsically the same, however, a different gravitational field becomes local, and so her weight changes.

First objection: Weight is in fact an intrinsic disposition that is merely specified relationally. The expression "local gravitational field" is to be understood rigidly, and is fixed by the context of utterance. So, when we say that someone weighs 120 pounds, in most contexts, we are to be understood as saying that she weighs 120 pounds *on earth*. The circumstances of manifestation for eliciting a reading of 120 pounds involve not only

¹¹¹Yablo, "Intrinsicness," 611.

stepping on a properly constructed scale, but also being in a particular gravitational field. *Weighing 120 pounds on earth* is a property a person retains even if she travels to the moon.

Reply: According to the above account, a person's disposition to depress a properly constructed scale so as to elicit a reading of 120 pounds in a gravitational field of strength f (i.e., the earth's gravitational field) is her *mass*, not her weight. While it is true that *weighing 120 pounds on earth* is a property that a person retains even if she travels to the moon, she also has the property of *weighing 120 pounds* simpliciter, as defined above. The expression "local gravitational field" can be understood rigidly, but it need not be. Our assumed account of "weight" is in accordance with ordinary language in this respect. In most circumstances that we ordinarily encounter, the salient gravitational field remains constant, and so the default assumption that we are talking about the quantity that a properly constructed scale would register on earth is a safe one. However, with the advent of space travel, there are circumstances in which a person moves from one gravitational field to another, and we say that she no longer weighs the same. The astronaut is "weightless" in outer space, and is no longer so when she returns to earth.

Second objection: This example, like the first one, involves a cooked-up gerrymandered property, not a natural property. It's telling that these properties do not (correctly) correspond to any simple predicates in English. The claim at issue is that dispositions like fragility, dormitivity, and solubility are intrinsic. Surely, you can make up a dispositional expression, including a context-sensitive term, so that that expression will apply to a thing at one time and not another. The predicate "is disposed to provoke the nearest animal" (let's suppose) applies to the red cape while it is nearest to a bull, but not while it is nearest to a rabbit. But these predicates do not refer to dispositions, properly conceived.

Reply: First of all, depending on what one means by "natural property," it is a matter of some controversy whether dispositions are natural properties.¹¹² So, even if the properties considered are non-natural, that does not show that they aren't dispositions. By what criteria of naturalness does "the disposition to provoke bulls" refer to a natural property, but "the disposition to provoke the nearest animal" fail to do so? Secondly, the correspondence of a property to a simple predicate is not decisive. It is implausible that the only dispositions there are are ones that we have simple predicates for. But if it is dispositions that are referred to by simple predicates that are wanted, the following examples should be illustrative.

Example 3: Vulnerability

A military target, a city, is protected by a Star Wars-like defense system. The system has sensors that bring out the defenses when there is a threat, to render the city invulnerable. However, the sensors and the anti-aircraft weapons are all located outside the borders of the city, and are built, maintained, and staffed by a foreign country. Should the defense system be disabled, or should the foreign power withdraw its protection, the city would change from being invulnerable to air attack to being vulnerable. However, the city might remain intrinsically the same, or internally the same in all ways that are relevant to its vulnerability.

First question: is vulnerability a disposition? Yes. It has a manifestation, i.e., harm to the vulnerable object. It has circumstances of manifestation, in this case an air attack. It is associated with certain counterfactuals, such as "if it were attacked, it would be harmed." Furthermore, vulnerability can be referred to by an overtly dispositional locution. Something is vulnerable if it is disposed to suffer harm, as a result of an attack.

¹¹²See Chapter 1, p.36 of this thesis for a discussion of the question: are dispositions natural properties?

Second question: Is vulnerability an extrinsic property, in this case? Yes. Perfect duplicates can differ with respect to their vulnerability. Whether or not something is vulnerable can be changed by changing its environment. By adding the defense system, the city changes with respect to its vulnerability. Nor is this phenomenon peculiar to this example. Walking alone in Central Park at night, I would be vulnerable. Accompanied by an entourage of bodyguards, I would be less so.

First objection: The city is not just what exists within its borders. The thing referred to as "the city" includes the defenses as well. When the defenses are disabled, the city undergoes intrinsic change. Therefore, its in/vulnerability is an intrinsic property.

Reply: Perhaps sometimes, something that is not contained within the geographical boundaries of a city can be considered part of the city. But equally, sometimes, forces and weapons which defend a city can be distinct from that city, not part of it. That is just the sort of case I am imagining. In this case, a foreign power owns and operates the defense system. It seems plausible that there can be something which is not part of a city which defends that city.

This response seems even less plausible in the bodyguard case. It is implausible that my entourage of bodyguards are part of me, or that I would undergo any relevant intrinsic change if my bodyguards abandoned me. Considering my bodyguards and I, it might be correct to say that *we* are invulnerable. However, considering just me, surrounded by my bodyguards though I am, it is also correct to say that *I* am invulnerable. So this objection fails.

Second objection: With or without the defenses, the city remains vulnerable. It is just that when the defenses are up, the city's vulnerability is *masked*. If a fragile glass were packed with protective material, it would not break when struck, but it would still remain fragile; it's just that its fragility would be masked. Likewise for the vulnerable city.

Reply: True, if a glass were packed with protective material we would still say that it is fragile. I also concede that we can make sense of the idea of vulnerability being

masked. Speaking of psychological vulnerability, we sometimes say of people "he has a tough exterior, but inside he is really vulnerable." Furthermore, we can make sense of the idea of vulnerability being had intrinsically. We say that newborn infants are vulnerable, even if they are as protected as can be.

However, this does not show that vulnerability is an intrinsic property. Recall that a property is extrinsic if, for any x , x 's having or lacking it *can* depend on what is going on outside of x . But that is not to say, for any x , x 's having an extrinsic property *does* depend on what is going on outside of x . Some extrinsic properties are not always *had* extrinsically. For example, the property of being loved by someone is an extrinsic property that one can, sadly, lose through merely extrinsic changes. However, being loved by someone can be an *had* intrinsically if that someone is yourself.

So, I'm not claiming that vulnerability is always had extrinsically. However, in the cases I have described, I think that it is. This accords with my linguistic intuitions. When external protection is added, vulnerability is not masked, but reduced. Consider one city with anti-aircraft defenses and one without. Would we say that they are equally vulnerable? Consider me with my bodyguards and my intrinsic duplicate walking alone in central park at night. Would we say that she and I are equally vulnerable? No. Such attributions of vulnerability are sensitive to the relational properties of the object under consideration. If a predicate is applicable to an object in one environment, and inapplicable to an intrinsic duplicate of that object in another environment, that is reason to think that the objects differ with respect to a certain extrinsic property.

Example 4: Visibility

Something is visible if it is disposed to be seen, or capable of being seen. The same thing might be visible to some perceivers and invisible to others. Very tiny objects, or objects which emit certain wavelengths of light, might be visible only to certain creatures. So, it might seem as though whether or not something is visible depends on

who is looking at it, and hence the property is extrinsic. However, a particular sense of visibility can be relationally specified, its referent fixed by reference to a certain reference class. In this way, visibility is like poisonousness. Nothing is visible simpliciter, but visible to a certain class of perceivers. Once visibility has been relativized to a class of perceivers, it doesn't seem like that property can be gained or lost by an object merely by changing its environment. So, it seems that visibility is intrinsic.

A standard philosopher's understanding of visibility is roughly associated with the following counterfactual: If this object were placed in a certain proximity to the relevant type of perceiver, in good light, with no obstructions in the way, then the perceiver would see the object. However, the expression 'visible' is not always used in a way that corresponds to such a counterfactual. Consider the report of the astronomer: "Saturn is going to be visible tonight for the first time this year." He does not mean that, for the first time this year, Saturn is such that, if it were in proximity to a perceiver in good light with no obstructions, it would be seen. Presumably, Saturn has fulfilled that description all year long. Consider also: "The top of the Eiffel tower will not be visible today." Or, suppose one were asked, in a darkened room, "is anything visible?" There is a clear sense in which one could say "no," and be speaking truthfully.

The idea that the expression "visible" is used in these different ways is supported by the fact that there are two different ways of negating it. Note, there is a difference in meaning between the expressions "not visible" and "invisible." To say something is not visible is often consistent with saying that if the viewing conditions were different, it would be visible. To say something is invisible suggests that it is something that can never be seen in standard viewing conditions, such as when we say electrons are invisible.

So, consider this other sense of "visible," such that the things in a room are visible only as long as the lights are on. Is this property a disposition? Yes. It has a manifestation, being seen, which occurs in the circumstances of manifestation, being looked at by the appropriate type of perceiver. Furthermore, it is typically true of a visible

object that, if the appropriate type of perceiver looked at it, he would see it. (It's not clear if "visibility" is equivalent to "the disposition to be seen when looked at," or any other overtly dispositional locution, but this is not a necessary condition for dispositionality in any case.)

Is visibility extrinsic? Yes. An object can cease to be visible without undergoing intrinsic change, merely by changing the lighting conditions, or by introducing an obstruction between the object and the perceiver.

First objection: The relevant property is not "visibility simpliciter." The dispositions in question are more accurately described as "visible from vantage point P" or "visible under level of illumination L," and these properties are intrinsic. They are properties that Saturn, and the objects in a darkened room, retain with changing conditions.

Reply: To say that an object is visible is not to say anything so specific. It is just to say that one can see it. Suppose the astronomer says that Saturn will be visible tonight, and will be again in six months, when the earth moves half way around its orbit. It's not the Saturn will be "visible from vantage point P" again, but visible from some vantage point or other.

Suppose that a room has both bright spots and dark shadows in it. All of the objects in the room have the following properties in common: they are visible in the bright light, and not visible in dark shadows. But as things stand, there is a respect in which some of the objects differ from one another -- some of them are capable of being seen, and others aren't. If someone asked whether something in that room were visible, it would not do to say "If you mean visible in bright light, then yes, but if you mean visible in shadow, then no." Presumably, the person is asking how the object stands in relation to the actual conditions of visibility, not whether the object would be seen in some possible, counterfactual lighting conditions. Change the lighting conditions, and the objects in the room change with respect to visibility.

Second objection: Granted, there are different senses of visibility. Let's call the first sense of visibility "visibility₀." Visibility₀ is just the philosopher's sense of visibility, associated with the counterfactual: if the appropriate type of perceiver were in proximity to the visible₀ object, with the right lighting conditions, no obstruction, etc., then the perceiver would see the object. Another sense of visibility, which the objects in a well-lit room have, is visibility₁. Visibility₁ is a conjunctive property, the property of having visibility₀, and being in the right lighting conditions. Similarly, we could specify visibility₂ --visibility₀, plus no obstructions, and so on. Visibility₀ is a disposition, and an intrinsic one at that, but visibility₁ and the rest are merely conjunctions of a disposition and a relational property, not dispositions themselves.¹¹³

This seems to me a nice analysis of the different senses of visibility. However, it is not clear why visibility₁, for example, is not a disposition. The conjunction of a disposition and a nondispositional property might itself be a dispositional property. Assuming transparency is non-dispositional, is it so clear that the property "being fragile and transparent" is non-dispositional? (Perhaps we have to start admitting degrees of dispositionality?) Furthermore, visibility₁ seems to have the marks of dispositionality --it has a manifestation, circumstances of manifestation, and an associated counterfactual.

Example 5: Recognizability

People can have dispositions relative to their social context. Consider recognizability. People who are recognizable are prone to be recognized. Bill Clinton is recognizable. When he goes out in public or makes a speech, people see his face and know that they were looking at the man called "Bill Clinton," the President of the United States.

¹¹³This objection is due to Steve Masterson in personal correspondence.

Is recognizability a disposition? Yes. It has a manifestation, being recognized, which occurs in the circumstances of manifestation, roughly, going out in public without disguise or concealment. Furthermore, certain counterfactuals are true of the recognizable person, such as: If he were to go out in public without any disguise or concealment, he would be likely to be recognized.

Is recognizability extrinsic? Yes. It is not the case that any intrinsic duplicate of Bill Clinton would share Clinton's recognizability.¹¹⁴ If it weren't for all the people in his environment that recognize him, Clinton would not be recognizable. In another possible world where Clinton became a chicken farmer in Arkansas, he can wander through Little Rock anonymously.

First objection: To say that someone is recognizable is just to say that his looks are such that someone who had seen him before could re-identify him. Certain people have this property, and social context has nothing to do with it. Therefore, a person's recognizability does not change with her social context.

Reply: Perhaps this is the way "recognizable" is sometimes used. However, we do say that people like Michael Jordan, Madonna, or Bill Clinton are recognizable, as opposed to your next door neighbor. In any case, we can consider the property of being prone to be recognized by a great number of people, whether the word "recognizable" is the proper name for this property or not. Furthermore, even this "re-identifiable" sense of recognizable seems to depend on social context. How easy it is for me to re-identify a person depends in part on how distinctive his looks are, and how distinctive his looks depends upon how they compare with the looks of those around him --his social context.

¹¹⁴This case is complicated by the fact that "recognize" is a success term. If you truly recognize someone, then they must be the person that you think they are. No one can *recognize* someone other than Bill Clinton *as Bill Clinton*, not even an intrinsic duplicate. If they think the duplicate is Bill Clinton, they would be in error. It would be a case of mistaken identity. So there is a sense in which, even Bill Clinton's duplicate were in the same social environment, he would not share Bill Clinton's recognizability.

Second objection: Being recognizable is a matter of being such that, if you were in a particular social context, people would recognize you. If there is some possible social context in which many people recognize you, then you are recognizable. Therefore, changing of one's actual social context would not change one's recognizability.

Reply: If this objection is to be distinct from the first, we must not take "recognizable" to mean merely "re-identifiable." If there is some possible situation in which your next door neighbor is recognized everywhere she goes, that doesn't make her recognizable as things stand. If that were the case, then everyone would be as recognizable as Bill Clinton. Part of our understanding of the word "recognizable" is that it applies to people like Clinton, and not to those of us who haven't even had our fifteen minutes of fame.

Third objection: With respect to recognizing, the disposition isn't Bill Clinton's, but the people who recognize him. They must have memories and associations in their minds in order to recognize him. The disposition to recognize is an intrinsic disposition of the individuals in the society.

Reply: It is true that individuals in society have a disposition to recognize Bill Clinton, and that may be an intrinsic property of each individual. However, when one says that Bill Clinton is recognizable, one is saying something about *Bill Clinton*. One is attributing a property to him and, as I argued, it is a dispositional property.

Fourth objection: Recognizability is one of a pair of dispositions, the other being the disposition of a person to recognize. Just as a thing can have a disposition D1 without anything having the second half of the disposition-pair, D2, so someone can be recognizable without there being anyone around with the disposition to recognize.

Reply: If the claim is just that someone can possess recognizability without being recognized, that is granted. A disposition need not be actualized for it to be possessed. That does not go to show that the disposition is intrinsic. If the claim is that someone can be recognizable even if there is no actual person that would recognize him, then it is a

repeat of the second objection. As I said, it seems that there is more to being recognizable than being such that one would be recognized in some merely possible social situation.

Fifth objection: Recognizability is an intrinsic property that is merely relationally specified. As will be the case of tallness, or poisonousness, we can specify recognizability in relation to a certain reference class. No one is recognizable simpliciter, but recognizable to a certain group. In the case of Bill Clinton, a relevant reference class might be the residents of the United States in the 20th century. So, Bill Clinton has the property of being "recognizable to residents of the United States in the 20th century" (and others of course, but we can consider just this group). His social context may change any way you like, transport him through time or space and leave him among strangers, but he will still have the property of being recognizable to residents of the United States in the 20th century.

Reply: It is true that we can specify a property that the recognizable person would not lose if he changed his social environment. However, just as "weight" could be defined making reference to an object's local gravitational field, whatever it may be, "recognizability" can be defined by making reference to a person's local social environment. On this account, a person is recognizable iff he is disposed to be recognized by people in his local social environment. This does not seem to be a drastic deviation from ordinary usage. This is a property a person can lose if his local social environment changes.

Other Examples

In the centuries before 1800, Egyptian hieroglyphics were indecipherable. In 1799, the Rosetta Stone was discovered. Not long after that, Egyptian hieroglyphics were decipherable. Is decipherability a disposition? Yes. It has a manifestation, being deciphered, or translated, which occur in circumstances of manifestation, which involve someone trying to read or understand what is written. Certain counterfactuals are true of

decipherable inscriptions, such as: if someone (with the appropriate skills) tried to decipher it, they would succeed.

Is decipherability extrinsic? Yes. Consider an inscription of Egyptian hieroglyphics on a certain tablet. Suppose it remained in a vault, untouched during the period of discovery and study of the Rosetta stone. During that time, it went from being indecipherable to being decipherable.

There are a number of expressions in our language that are both dispositional and dependent on social institutions and social contexts. A coupon is redeemable. A device is marketable. A position is enviable. An event is memorable. A statement is humorous, provocative, or inflammatory.

3.7. Conclusion

There are basically two ways to reply to these proposed examples of extrinsic dispositions. 1. The examples I mentioned aren't really dispositions. 2. Contrary to appearances, these dispositions really are intrinsic after all. Perhaps they can be cast as properties that are merely specified relationally, in one of the senses I explored earlier. I have examined various attempts to cast these properties as merely relationally specified, and shown why these attempts are not successful.

A proponent of the first type of reply might say, more generally, it is a mistake to assume that all predicates ending in -able or -ible are dispositional, and that the "...is disposed to..." locution is not a reliable guide for determining dispositionality. Words like "desirable," "valuable," "contemptible," and "reprehensible," for example, have a normative component. Some say that being desirable is not a matter of being prone or disposed to be desired, but of being apt for or worthy of desire. However, I did not rest any claims of dispositionality of the fact that a predicate has a certain suffix. I did not claim that all predicates that end in "ible" or "able" refer to dispositions, though I admit

that such a suffix serves to suggest that a predicate is a candidate for being considered dispositional.

Various reasons could be offered as to why a proposed example is not in fact a disposition. It might be claimed that, if there is no precise counterfactual associated with the term, then that is reason to think that the term is not dispositional. Consider a coach of one basketball team saying to his players that the opposing team is "beatable." Presumably, he doesn't mean that if the opposing team scores fewer points, it will lose. That would be true of unbeatable teams as well. It might be claimed that the coach is not attributing a disposition to the opposing team at all, and as above, we should not be misled by the "able" suffix on the predicate. The same goes for properties like decipherable. It is difficult, if not impossible, to state a precise counterfactual that neatly corresponds to decipherability.

However, it is widely acknowledged that few, if any disposition terms have a *precise* counterfactual associated with them. The precision of associated counterfactuals is a matter of degree. That is only to be expected, since some disposition terms are more precise than others. Furthermore, this objection would deem not only beatable, decipherable, and some of the extrinsic properties I mentioned as non-dispositional, but also, cowardly, shy, irritable, and provocative, to name a few.

Some may remain unwilling to call extrinsic properties like visibility or recognizability dispositions. However, if they are willing to concede that there are extrinsic properties that have manifestations, circumstances of manifestation, and the relevant associations to counterfactuals, then the disagreement between us is merely a verbal disagreement. Perhaps I apply the expression "disposition" more liberally than most. If my application of the term "dispositional" to these extrinsic properties is somehow nonstandard, I would welcome a more precise characterization of the concept of a disposition, and a fuller specification of the marks of dispositionality.

In sum, I began by looking at paradigm cases of dispositions, and considered the features they have in common in virtue of which we consider them dispositions. This gave us rules of thumb for deciding if a certain property is dispositional. I considered The Conceptual Argument and Armstrong's Argument that dispositions are intrinsic, and found them unconvincing. I then gave examples of extrinsic properties that bear the marks of dispositionality. I considered a number of objections to the claim that these examples are in fact examples of extrinsic dispositions, and found none of them decisive. In the end, it seems that those who insist that dispositions are intrinsic must accuse me of misapplying the expression "disposition." If that is so, then I am owed some criteria according to which we may appropriately call a property dispositional. If one merely adds "they must be intrinsic properties," then I have succeeded in identifying a set of properties that are just like dispositions, except they are extrinsic.