The Southern Journal of Philosophy (2001) Vol. XXXIX

Causally Inefficacious Moral Properties

David Slutsky
University of Connecticut

1. Introduction

According to moral realists such as Richard W. Miller, Nicholas Sturgeon, Peter Railton, Richard Boyd, and David Brink, various parts of the world have certain moral properties independently of whether anyone believes that they do, and these moral properties help explain certain things that people do.[1] These moral realists seem to have causal explanations in mind, and sometimes they state explicitly that moral properties have effects on human behavior.[2]

In this paper, I will try to motivate skepticism about the causal efficacy of moral properties in two ways. First, I will highlight a tension that arises between two claims that moral realists may want to accept. The first claim is that physically indistinguishable things do not differ in any causally efficacious respect. The second claim is that physically indistinguishable things that differ in certain historical respects have different moral properties. The tension arises to the extent to which these different moral properties are supposed to have different effects on people. I will introduce a class of cases in which this tension arises and suggest that the moral properties in these cases have no causal power. I will also question whether there are differences between the moral properties in these cases and moral properties in other cases that do not involve physically indistinguishable things that could make the latter moral properties causally efficacious.

The second way that I will try to motivate skepticism consists in pointing out a unique feature of cases in which

---

David Slutsky is a Ph.D. student at the University of Connecticut. In addition to working in metaethics and defending anti-realism about moral properties, his interests include normative ethics and metaphysics. In normative ethics, he works on consequentialism and egalitarianism, defending demanding accounts of moral and political obligation. In metaphysics, he works on the nature of historical properties, defending anti-realism about social, psychological, and biological properties that are supposed to supervene on historical properties.

moral properties are supposed both to supervene on historical properties and to be causally efficacious. These cases allow us to change moral properties with alleged effects while we hold constant the nonmoral candidates for causal contribution to those effects. This feature of these cases is unique because in most other cases the moral properties supervene on the physical properties in the nonmoral candidates, such that we cannot change the former while holding constant the latter. This way of changing moral properties provides empirical grounds for testing their causal efficacy.

2. An Assumption about Historical Properties

I assume that historical properties are causally inefficacious and that physically indistinguishable things that differ in their historical properties do not differ in any causally efficacious respect. Let me clarify how I will be using these claims and begin by quoting two people who have made them.

> Beds that were slept in by George Washington may command a higher price than those that lack this historical property, but presumably this is a result, not of any causal potentialities in the beds themselves, but of the historical beliefs and interests of those who buy and sell them. (Shoemaker 1980, 121)

> The cashier accepts my $100 bill because she thinks it is genuine, and she thinks it *is* genuine, not because it is genuine (a fact V [for value] about the history of this paper), but because it *looks* genuine (a fact C [for cause] about its intrinsic properties). Even if it was an authorized government agency that *caused* it to look this way, it is the fact that it looks this way, not the fact that an authorized agency made it look this way (the fact on which the value of the paper supervenes) that explains the cashier's coming to believe that it is genuine and, hence, her reaction to it. (Dretske 1990, 8)

> Uncles, dollar bills, and divorcees can cause lots of trouble, but the fact that they are (as opposed to being believed to be) uncles, dollar bills, and divorcees, does not explain their various effects on the world.... These effects are to be explained, instead, by the intrinsic properties of uncles, dollar bills, and divorcees.... So, since counterfeit—good counterfeit (i.e., paper having the same intrinsic properties)—is as effective in obtaining coats and groceries as is real money, the property of being real money adds nothing to the paper's causal powers. (Dretske 1998, 521)[3]

It is not easy to find in print disagreements with these claims and assertions that historical properties have causal power. Michael Tooley (1972, 286–287) and J. L. Mackie (1973,

130–131) seem to be two examples. Let us call this the historical view. For purposes of illustration, consider two physically indistinguishable paintings that are each in the centers of physically indistinguishable rooms at a time T. Suppose that these paintings have different historical properties, such as having been touched by different people and produced in different ways. The issue here is not whether different methods of production, for example, can affect present causal powers. The issue is whether such differences can do so even if there are no physical effects or traces of them involved in the present.

Now suppose that we bring a mirror (or a camera or a camcorder or a computer with a video camera or a human being) into one of the rooms and place it in front of one of the paintings. According to the historical view, we should take seriously the possibility that the different historical properties of the paintings could produce different images in the mirror. Of course, it takes time for light to travel from the surfaces of the paintings to the mirror. In these terms, the historical view suggests that we take seriously the possibility that the different historical properties of the paintings could have different effects on the quality of the light that reflects onto the mirror.

On an alternative view, the physical properties of the paintings, combined with such contemporaneous properties as the atmosphere in the room, wholly determine the causal powers of the paintings and the effects on the mirror. Let us call this the physical view. As Prior, Pargetter, and Jackson (1982) point out, this view does not require determinism. We can restate it in terms of the physically indistinguishable paintings having the same probabilistic dispositions regardless of their historical properties.[4]

Now suppose that we bring the mirror in front of the first painting to see what image it produces, and, then, we produce a physical change in the mirror, such as a crack in it. Next, we bring the mirror in front of the second painting, and we see that it produces a different image. The historical view suggests that we take seriously the possibility that the different historical properties of the paintings are causally responsible for the differences in the effects that the paintings have on the mirror. On the physical view, even though the physical properties of the paintings may be causally responsible for the different images, it is the physical difference in the mirror that makes the different causal contribution to those different effects. Since the physical properties of the paintings make the same causal contribution to the different effects, it is the physical difference in the mirror that is causally responsible for the difference in the effects. The historical properties of the paintings (and the mirror) contribute nothing to, and do no causal work in bringing about, those effects.[5]

As Shoemaker and Dretske assume in the examples with which this section began, these claims apply to humans as well as mirrors. Suppose that someone has different beliefs about two physically indistinguishable paintings (or beds or $100 bills)—namely, she believes that one is an original, one is a replica, and the original is worth more money than the replica. These different beliefs are analogous to the physical difference in the mirror. Suppose that this person offers more money for the original than the replica. Although the physical properties of the original and the replica may make causal contributions to, and be causally responsible for, the person's different offers, it is something about the physical properties of the person that is causally responsible for the difference in the offers. I shall continue to assume in these terms that the historical properties themselves do no causal work, particularly in cases that involve human behavior.

I propose to show how analogous claims apply to certain moral cases. My strategy is to describe physically indistinguishable things that differ in their moral properties in much the same way that these beds and paintings differ in their historical properties. If a person forms a belief about or responds in a certain way to these subjects of moral properties, it is something about the person's brain or psychology (combined with the nonhistorical properties of the subjects) that is causally responsible for that belief or response. More particularly, the historical beliefs and interests of the person, rather than the historical properties of the beds< do the causal work in bringing about the different money offers in the bed case. Similarly, the moral beliefs and interests of the person, rather than the moral properties of the moral subjects, do the causal work in bringing about the different effects that the moral properties are supposed to have on the person. I will fine-tune the moral cases in specific ways and for purposes that I will try to make clear, but these differences will not affect the basic strategy of apportioning causal responsibility with disregard to historical properties.

3. A Sample Case

Suppose that, at a time T, every person present in a particular society is barely conscious and dying from disease as a result of overwork and malnutrition. We can imagine the same state of affairs arising from different histories and thereby possessing different historical properties. In one history, at an earlier time T-n, the society consists of a large group of struggling farm workers and a small affluent group of farm owners. The small group made voluntary contracts with the struggling farmers. These contracts consisted of little pay in exchange for long hours of work. Soon all the farmers were overworked to the point of severe illness and the depletion of

all natural resources. Although the small affluent group had enough medicine to heal everyone and enough seeds to plant new crops, they had made no agreements to do any such thing. After their castle accidentally burns down, they decide to leave the workers for dead and move on to strike up a similar deal in another society.

In a second, alternative history, at the same earlier time T-n, by majority rule a society of mostly impoverished farmers voted to distribute what little resources they had in a Rawlsian fashion. Although not everyone agreed with this arrangement, the minority complied after considering the societal punishment for noncompliance with majority votes. These egalitarian farmers shared every resource available and tried to follow proper agricultural techniques. However, their storage sheds burned down and other misfortunes such as unfavorable weather eventually led to malnutrition and disease. Despite their cooperation and hard work, at T everyone lies barely conscious and dying.

Suppose that these two states of affairs do not differ in any physical respect. The sick bodies, for instance, lie in the same positions. The affluent group had built their castle in the same area where the egalitarians had built their storage sheds. Each fire left the ground below it in the same condition. And so on. Suppose further that in each of these states of affairs, infection and high fever has caused the same thorough brain damage, such that there are no differing memories or inactive brain states.

Although these two states of affairs do not differ in any physical respect, many people may believe that they differ in their moral properties.[6] If we change the history in the first society to include torture, slave trade, and physically coerced work in place of libertarian agreements and transfers, almost everyone would believe that the states of affairs at T differ in their moral properties.

4. Analysis of the Sample Case

Recall that, on the realist view under question, parts of the world have certain moral properties independently of whether anyone believes that they do, and these moral properties cause some people to do certain things. For instance, suppose that, at T-n, a Rawlsian walks into the inegalitarian farming society with the first history. She feels and expresses moral outrage at the injustice that she sees in the society. Later, at T, she returns to check on the farmers. She feels even more upset about the moral badness that she sees in the society. As a result, she begins a search for the small group of affluent farmers to express moral condemnation to them in person.

According to Miller, Sturgeon, Railton, and Brink, in cases like this the Rawlsian's responses to the conditions in the first

society were caused by the existence of injustice in it, and an explanation of her responses includes reference to the existence of this moral property as a cause.[7] However, suppose further that, at both T-n and T, the same Rawlsian walked into the more egalitarian farming society with the second history. She feels and does very different things in connection with the justice and later tragedy that she sees in the society, such as first praising and later paying sad tribute to the farmers' just and valiant efforts. According to these moral realists, presumably, this instance of justice is relevant as a cause in the explanation of the Rawlsian's behavior.[8]

When the Rawlsian enters the two societies at T, however, the only things with different causal powers in or regarding the two societies are the Rawlsian's different mental or brain states. In each case the Rawlsian has different memories and beliefs about what events led up to the societal conditions that she sees. More importantly, she has both different beliefs about the morality of those conditions and different desires regarding what she wants to do about them.

To be sure, the physical features of the two societies make the same causal contributions to the Rawlsian's sense organs. Although these indistinguishable societal features may be causally responsible for the Rawlsian's different responses to them, it is the differences in the Rawlsian's psychology (or neurology, if you prefer) that make the different causal contributions to, and hence are causally responsible for, the differences in her responses.

On the moral realist view, presumably, the different properties of injustice and justice at T supervene on the different historical properties of the (physical features of the) two societies. A moral realist might also claim that these different moral properties are causally responsible for the Rawlsian's responses at T even though the differences in the Rawlsian's psychology are causally responsible for the differences in her responses. One reason to question the plausibility of this claim is that the historical properties of the two societies at T do not do any causal work in bringing about the Rawlsian's responses.[9] If they don't, then it is not clear how any moral properties that supervene on them could do any causal work in bringing about the responses either.[10]

On the other hand, the features of the two societies at T-n on which injustice and justice are also supposed to supervene, such as the different economic distributions and social relations, do differ in their causal powers. Recall that the more equal economic distribution in the second society is fair or just according to a Rawlsian theory, and the less equal distribution in the other society is not. However, it's not clear why we should believe that these different moral properties make any causal contributions to, or are causally responsible for, the

Rawlsian's responses at this time (T-n). Why should the Rawlsian's psychology and the physical features of the societies be causally responsible for the Rawlsian's responses at T but not at T-n? Why should the properties of injustice and justice be causally efficacious at T-n but not at T? At what point in the history of these two societies would the causal efficacy of the moral properties change, and how could it do so? We can raise the same questions about how the different moral properties could make the different causal contributions to, and hence be causally responsible for, the differences in the Rawlsian's responses at T-n.

5. Extrapolation from the Sample Case

In principle, we can apply this argument to any case in which a moral property of something outside of a person's body at a time T is supposed to cause that person to do something. What we need to do is imagine a change in the historical properties of the subject of the moral property at a time T-n in such a way that (1) we would want to say that this subject has a different moral property at T-n and the same different moral property at T, and (2) according to moral realism, the different moral properties have the same kinds of different effects on the person at both times. If historical properties are causally inefficacious, then the only candidates for causal responsibility for the differences in the effects are the people or their psychology, and the only candidates for causal responsibility for the effects themselves are the nonhistorical properties of the moral subjects.[11]

In practice, however, we probably cannot apply this argument to the vast majority of moral cases. Although many cases of moral properties appear to supervene on historical properties, we usually need to see changes in more than historical properties before we see changes in moral properties. When we change the physical properties at T-n that are required for different moral properties, we will also be changing the physical properties at T. For instance, the moral status of keeping people in prison depends on their current psychology as well as their past actions, and prisoners ordinarily do not have identical psychologies or physically indistinguishable brains.

Nonetheless, notice that the argument does not depend on a Rawlsian conception of justice or morality. In the case of states of affairs and the two farming societies, for example, at T-n the more equal economic distribution in the second society is unfair or unjust according to a libertarian or Nozickian theory, and the less equal economic distribution in the first society is not. So the moral properties that a Nozickian would see in the two societies at T-n and T would be reversed from those that a Rawlsian would see, and, we can imagine, her reactions to them at both times would differ in correspondingly reversed ways.

A crucial question here is whether this strategy points to the causal inefficacy of moral properties in general rather than just to that of the particular moral properties in the cases to which it is applied. The answer depends on whether the causal efficacy, or even the moral properties themselves, alleged in such particular cases differs in some metaphysically significant way from the causal efficacy alleged in the many cases for which the strategy will not work. Miller's sample cases include civil rights violations in seventeenth century England (1979, 254–55), slavery (1985, 527), and Nazi involvement in Hungary (1992, 46–54). Sturgeon's cases include slavery (1985, 245–46; 1991, 28–30), someone harming a cat (1986, 70–75), and poverty (1986, 75; 1991, 36 n. 16). Railton's more general case involves the societal discounting of the interests of certain groups and departures from an impartial social perspective with equal concern for everyone affected (1986, 189–94). Finally, Brink's cases include apartheid in South Africa (1989, 187–92).

Should it make a difference whether enslaved, impoverished, exploited, or discounted people have been pushed to severe illness and the brink of death? If anything, the features of my sample case would otherwise seem to strengthen the case for both the existence of the moral properties in them and the alleged accompanying power to cause people to do certain things, such as opposing slavery, deciding not to lend support or prestige to a particular government, and forming moral beliefs about individuals, actions, and distributions. The sorts of features that appear to lead moral realists to postulate the presence of moral properties and their effects on people are not removed or altered by the details required for imagining cases in which the above argument works.

6. Changing the Moral Candidate for Causal Contribution While Holding Constant the Nonmoral Candidates

Moral realists, and common people, often explain a person's behavior by pointing to the presence or instantiation of a particular moral property in that person's environment as a cause. Moral anti-realists explain the same behavior by pointing to the person's beliefs (or psychology, which includes the person's emotions, desires, unconscious beliefs, and so on), and to the physical properties of the person's environment as causes. Moral realists do not deny that the person's beliefs and those physical properties play an important causal role. The disagreement is over whether the moral property in the moral explanans does any causal work that we cannot wholly account for with the nonmoral explanans.

Sturgeon argues that moral properties do this causal work by appealing to a counterfactual test for explanatory relevance.

According to this test, if a moral property of something such as an act (or an economic distribution or a societal state of affairs) were irrelevant in explaining a person's response to that act, then the person would have responded in the same way even if the act did not have that moral property. For instance, suppose that Jane sees Albert torturing a cat, and Jane forms the belief that Albert's torturous act is wrong.[12] Sturgeon says that the actual wrongness of Albert's act caused Jane to believe that it was wrong and that this wrongness is relevant in explaining Jane's belief.

In this case, Sturgeon's counterfactual test says that if the wrongness of Albert's act were irrelevant in explaining Jane's belief about it, then Jane would have formed the same belief even if Albert's act had not been wrong. Sturgeon's argument proceeds with two steps. First, since moral properties supervene on natural properties, in order to imagine a case in which the act has a different moral property, we are supposed to have to imagine a different act, such as Albert petting the cat. Second, if we imagine a different act with a different moral property, then we are supposed to be imagining an act that Jane would not have responded to in the same way. Sturgeon concludes that since Jane would not have responded in the same way to the case in which the moral property was not present as she responded in the case in which it was present, the moral property passes his test for explanatory relevance.

This is a powerful argument. It's not clear what it shows, however, in the many cases in which the moral property is supposed to supervene on the physical properties of the environment that are cited in the nonmoral explanans. In these cases, it's not clear whether the argument shows that the moral property or the physical properties are making the causal difference to the alleged effect of the moral property. In order to make this clear, we need to modify Sturgeon's test so that it requires us to change the moral property while holding constant all of the properties in the nonmoral explanans. The point of Sturgeon's argument seems to be that we cannot do this when the former supervenes on the latter. In these terms, Sturgeon's argument does not show that any moral properties pass the test. Rather, the point seems to be that most moral properties do not pass or fail the test, since we cannot carry it out on them.

We can carry out the test, however, in certain cases in which moral properties are supposed to supervene on historical properties. In these cases, we can imagine that the subject of the moral property has a different moral property without imagining that the subject differs in any of its properties that are cited in the nonmoral explanans. We can carry out the test and change the moral property in this way by changing the historical properties on which it supervenes, since the nonmoral explanans does not include these historical properties.

The question is whether the person would respond differently, not because we have changed the nonmoral candidates for causal contribution but because we have changed the moral candidate. Since the nonmoral explanans includes the person's beliefs, this question requires us to imagine a case in which the person has the same beliefs about the subjects of historical properties but also in which those subjects have different historical properties. This kind of case should not be too difficult to imagine, since many people form false beliefs about historical properties under a variety of circumstances.

Notice that it is assumed in my sample cases that the Rawlsian's beliefs about historical properties are true. In the spirit of Sturgeon's counterfactual test, let us imagine some variations in which the Rawlsian's beliefs are false. Suppose that the two societies that I describe in section 3 occupy different regions of space at the same times. Recall that the spatial regions occupied by the two societies have different physical features at T-n but not at T. On a Rawlsian or a Nozickian view, the states of affairs in these societies have different moral properties at both times. Suppose that a different Rawlsian walks into each society at T-n, and they each return as planned to the respective societies at T. The Rawlsian who visits the egalitarian society responds first at T-n by praising and encouraging and later at T by praising and paying sad tribute to, the society members. The Rawlsian who visits the inegalitarian society responds first at T-n by feeling and expressing moral outrage to the opulent society members and later at T by setting out to express moral condemnation to them in person.

Now suppose that both Rawlsians unknowingly get lost after T-n and each winds up entering the other society for the first time at T. They each enter the respective societies with the false beliefs that they had visited them before, that the conditions that they see are a result of the conditions that they saw at T-n, and so on. Presumably, this switch has no effect on the different moral properties of the spatially different societal states of affairs. Apart from the Rawlsians and their brains, the switch has no effect on the physical features or the histories of the different societal states.

We are now ready to see whether the moral properties in these two societies at T pass the modified version of Sturgeon's counterfactual test. We have two cases in which both the moral property in the moral explanans is different and the nonmoral properties in the nonmoral explanans are the same. According to the test, if the first Rawlsian responds in the same way in the case in which she unknowingly gets lost as she responds in the case in which she does not get lost, then the moral property in the case in which she does not get lost is not explanatorily relevant. The same goes for the other Rawlsian.

It's an empirical question how the Rawlsians would respond. If the moral properties are causally inefficacious, then we should not expect their absence to make a difference to the Rawlsians' responses. To be sure, if the moral properties are causally efficacious, then we might expect both the Rawlsian who unknowingly enters the egalitarian society (but thinks that she is entering the Nozickian society) to respond with praise and comradeship, and we might expect the Rawlsian who unknowingly enters the Nozickian society (but thinks that she is entering the egalitarian society) to respond with blame and protest. In the least, if the moral properties are causally efficacious, then we might expect both the Rawlsian who unknowingly enters the egalitarian society to respond with less blame and protest, and we would expect the Rawlsian who unknowingly enters the Nozickian society to respond with less praise and comradeship.

Of course, we can similarly test whether the Rawlsians' beliefs (and psychology) in the nonmoral explanans are explanatorily relevant by looking to see what happens when we hold constant the moral property in the moral explanans and change the Rawlsians' beliefs. Suppose that, shortly after T-n, each of the Rawlsians hears about the other society with the different moral property and decides to visit it for the first time at T. Suppose that the Rawlsians unknowingly get lost and wind up returning at T to the same societies that they left at T-n. They each enter the respective societies with the false beliefs that they had not visited them before, that the conditions that they see are the result of the conditions that they had heard about in the other society around T-n, and so on.

If the moral properties but not the Rawlsians' beliefs are causally inefficacious, then we should expect the Rawlsians' responses (to what they each believe to be a different moral property) to change along with their false historical and moral beliefs, even though the moral properties in the respective societies that they return to remain the same. As in the other scenario, if the moral properties are causally efficacious, then we might expect both the Rawlsian in the Nozickian society to respond with less blame or protest and the Rawlsian in the egalitarian society to respond with less praise and comradeship than they would have responded with if they each had not gotten lost and entered the societies with the moral properties that are supposed to cause the opposite sorts of responses.

7. Conclusion

It's an empirical question what results we would see in these cases. Suppose that the responses would correspond to the moral beliefs rather than the moral properties and that the best explanation of these results is that the moral properties are

causally inefficacious. We can return to the question that we asked in section 5 about the argument in section 4. Do the reasons for believing in the causal inefficacy of the moral properties in these cases apply to other cases in which moral properties do not supervene on different historical properties of physically indistinguishable things? I have four considerations to suggest in favor of generalizing the explanation and argument.

First, as suggested in section 5, it seems as though the same features that lead moral realists to posit the causal efficacy of moral properties in physically distinguishable cases are present in the physically indistinguishable cases. Second, as suggested at the end of section 4, it seems as though the different moral properties of the two societies at T-n that supervene on physically distinguishable properties are causally inefficacious. It's not clear why the nonmoral properties should be causally responsible for the alleged effects of the moral properties at T but not at T-n, and it's not clear how the moral properties could be causally efficacious at T-n but not at T. At what point between T-n and T would the causal efficacy of the moral properties change, and how could it do so? Even if the features that lead moral realists to posit the causal efficacy of moral properties in their preferred cases are not present in the two societies at T, it seems at least more likely that they are present in the two societies at T-n.

The third consideration involves a way to run the first two considerations from the present to the future even if we cannot run them from the present to the past. In other words, suppose that we have a moral property that is supposed to have an effect on a person at a time T, and we cannot imagine a change in the historical properties of the subject of the moral property at a time T-n in such a way that (1) we would want to say that this subject has a different moral property at T-n and the same different moral property at T, and (2) according to moral realism, the different moral properties have the same kinds of different effects on the person at both times. Nonetheless, we might still be able to imagine a physically different subject of a different moral property at T in such a way that (1) we would want to say that, at a future time T+n, these two subjects are physically indistinguishable and have the same moral properties that they had at T, and (2) according to moral realism, these different moral properties have the same kinds of different effects on the person at both times. If so, we can apply the claims in section 4 to these cases.

Finally, the fourth consideration involves the modified version of Sturgeon's counterfactual test in section 6. Recall that, on this version, Sturgeon's argument does not really show that any moral properties pass the test. Rather, it shows that most moral properties do not pass or fail the test, since we

cannot carry it out on them. If this is a plausible test, and if we cannot carry it out in most cases, it seems a plausible conjecture that we should pay attention to cases in which we can. I have tried in this paper to introduce one class of these cases. Although there may be others, they seem to constitute a small percentage of moral cases. Nonetheless, it does not seem implausible to suppose that the moral properties in these cases will always fail the test. It's not clear how to balance a large number of moral cases not passing or failing the test with a small number failing it, but skepticism seems to be a reasonable option.[13]

Notes

[1] See Miller 1979 (252–255), 1985 (526–530, 552–554), 1987 (111), 1992 (46–54), Sturgeon 1985 (245–246), 1986 (70–75), 1991 (27–30), 1992 (97–98, 105–110), Railton 1986 (164–165, 171–173, 183–184, 191–193), Boyd 1988 (181–182, 188–189, 205–206, 210–211), and Brink 1989 (187–197).

[2] See Sturgeon 1985 (250 n. 26), 1986 (74–75), 1991 (28–29), 1992 (98, 106), 1995 (171), Miller 1985 (527, 553), Railton 1986 (171–172), and Brink 1989 (187). Although they do not discuss what distinguishes the cases in which they speak of the causal efficacy of moral properties from the cases in which they speak of the explanatory relevance of moral properties, in some passages they seem to use these phrases interchangeably. See Sturgeon 1985 (250 n. 26), 1991 (27–28), 1995 (171), and Brink 1989 (189), summarizing Railton 1986 (191–192). Sturgeon 1995 (171) in particular seems to make clear that we should interpret moral explanations as causal ones, and I will do so in this paper.

[3] For some related examples, see Kim 1990 (68), Enc and Adams 1992 (638), and Antony 1996 (72–73).

[4] I follow the tradition of restricting these claims to phenomena above the atomic level. I imagine that there are reasons in physical theory for this restriction.

[5] For assertion and discussion of this claim, see Shoemaker 1980, Dretske 1990, 1998, Kim 1991, Enc and Adams 1992, and Antony 1996. It does not mean that past intrinsic and contemporaneous relational properties are not explanatorily relevant. Indeed, the indistinguishable features of the paintings have different causes in many of them. Rather, this means that the historical properties in the present of having had or having been affected by those nonhistorical properties in the past do not do any causal work in bringing about the effects on the mirror. Dretske and Kim seem clear on making and separating these points.

[6] Actually, it would be enough for my purposes if the features in each society did not differ in any physical respect that is believed to be morally relevant. Barely conscious people, for example, probably do not need to be lying in the same positions.

[7] Miller (1979, 255; 1992, 60, 63–64), Sturgeon (1985, 245; 1986, 75; 1991: 27–29, 36 n. 16; 1992, 97–98), Railton (1986, 191–192), and Brink (1989, 187) seem to suggest that injustice can cause people to do just or morally good things, such as condemn or oppose the injustice.

Miller (1979, 255), Sturgeon (1986, 70–75), and Railton (1986, 191–192) seem to suggest that injustice or moral badness can cause people to form beliefs about such injustice or badness, and Sturgeon and Brink seem to claim that one person's morally bad character can both cause that person to believe or do something bad (Sturgeon 1985, 244, 249; 1986, 75; 1992, 97–98; Brink 1989, 187), and cause another person to believe that the first person has a bad character (Sturgeon 1985, 234, 243; Brink 1989, 187). In addition, Boyd (1988, 193–195, 209–214) claims that moral properties make causal contributions to the use of moral terms—with increasing accuracy in reference, no less.

[8] Brink (1989, 187) and Sturgeon (1991, 29; 1992, 98, 105–106) seem to claim that social justice causes social stability, and Miller (1987, 111–113; 1992, 46, 49–54, 63–64), Brink (1989, 187), and Sturgeon (1992, 97–98) seem to claim that a person's morally good character can cause that person to do something morally good, such as oppose or condemn an injustice. Since Railton (1986, 191–192), Brink (1989, 187), and Sturgeon (1991, 27–30, 36 n. 16) also seem to speak of opposition to or condemnation of injustice as a case of social injustice causing social instability, it seems a plausible conjecture that they would speak of praising or paying tribute to just people or just arrangements as a case of either social justice causing the sorts of things that ordinarily contribute to social stability, or what moral realists would otherwise regard as an effect of justice. To be sure, Sturgeon 1992 (98, 122 n. 5) cites as an example of what he has in mind by justice causing stability John Rawls's early conception of a stable society, according to which sentiments of concern that support just arrangements play a significant role. In addition, notice that Sturgeon (1992, 112 n. 10, 115 n. 24) registers agreement with Boyd (1988). Boyd appears both (1) to use an empirical and social account of nonmoral human goodness to define analytically moral goodness in terms of a social conception of the maximization of nonmoral human goodness (1988, 203, 207, 209, 224–225), and (2) to claim that moral goodness (or badness) normally causes people to sympathize with moral subjects and act accordingly in the same perceptual way, I take it, that edible food normally causes people to see and smell edible things, and eat accordingly (214–215).

[9] As stated in note 5, for assertion and discussion of this claim, see Shoemaker 1980, Dretske 1990, 1998, Kim 1991, Enc and Adams 1992, and Antony 1996. Dretske (1990, 6–8) makes precisely this claim about cases in which value supervenes on historical properties. See also Dretske 1998, 516–523.

[10] Notice a disanalogy that obtains between the example of the societies and the example of the paintings from section 2. On the one hand, the physical features of the societies at T-n are causally responsible for the memories and beliefs that the Rawlsian has at T, and these memories and beliefs are substantially causally responsible for the Rawlsian's responses at T. On the other hand, the physical features of the paintings at earlier times are not causally responsible for the physical features of the mirror at later times. Nonetheless, if we accept the assumption about the causal inefficacy of historical properties, we can still say that the causal efficacy of the intrinsic properties at T-n does not confer causal efficacy on the historical properties at T (of having had those intrinsic properties at T-n). Cf. note 5.

[11] Perhaps the nonhistorical properties of the moral subjects, such as the physical features of the two societies at T and T-n, are only causally responsible for the effects that they have on the people's sense organs, and that the people, or some aspect of their bodies or minds, are causally responsible both for their responses, and for the differences in their responses, to the moral subjects. On the other hand, Miller (1978, 403–414; 1987, 116–118) has argued that correctly identifying certain people and their conscious psychology as the actual causes of social phenomena leaves those phenomena unexplained when they would have occurred anyway in the absence of those particular causes. Miller (1987, 86–98) also argues that standards of explanatory adequacy for causal descriptions are different for different fields, times, and practical interests. Even on this broad account of explanation, however, the actual causal factors singled out for description must be causally sufficient to bring about the explanandum (1987, 87, 91) even if they do not constitute an adequate explanans (1978, 402–404, 412–413). The moral properties in my sample case do not meet this criterion. In addition, I imagine that most moral anti-realists would be glad both to expand their explanans of the alleged effects of moral properties on human behavior to include such things as unconscious motives and social functions, and to assert that other people would behave in similar ways under similar circumstances. These inclusions and assertions seem to point to explanatory facts about human psychology and social relations, not, as Miller (1985, 527–528, 553–554; 1987, 111–113; 1992, 54–60) suggests, to facts about moral properties. To be sure, notice that Miller gives unconscious factors more weight in his account of social explanation (1978, 393–395, 397–398, 411–412; 1987, 116) than in his account of moral detection (1992, 60–76, esp. 63–65, 69). Presumably, he sees no imbalance here because (1) he believes that the methodological motivation for rejecting moral explanations does not apply to social explanations, and (2) he takes acceptance of social explanations to preclude rejection of moral explanations. One reason not to be moved by (1) that I will discuss in section 6 is that we may have empirical, rather than methodological, reasons for rejecting moral explanations. One reason not to be moved by (2) is that we may be able to construe social explanations as one kind or feature of nonmoral explanations that include in their nonmoral explanans unconscious motives, social circumstances, and the intrinsic properties of past social circum-stances (as opposed to the historical properties of present ones) as candidates for causal contribution to the alleged effects of moral properties.

[12] This example comes from Harman 1977 (4–5, 7–9) and 1986 (59–64). For his discussion of it, see Sturgeon 1985 (232–233, 246–250), 1986 (70–75).

[13] Thanks to Lara Denis, James Klagge, David Robb, and P. Kyle Stanford for comments on earlier versions of this paper. Thanks to Joel Kupperman, Thomas Bontly, John Troyer, Gunnar Bjornsson, and members of the University of Connecticut Philosophy Department for comments on more recent versions. Thanks also to Helen Gladish Adam Potthast.

References

Antony, L. 1996. Equal rights for swamp-persons. *Mind & Language* 11:70–75.

Boyd, R. 1988. How to be a moral realist. In Sayre-McCord 1988, 181–228.

Brink, D. O. 1989. *Moral realism and the foundations of ethics*. Cambridge: Cambridge University Press.

Dretske, F. 1990. Does meaning matter? In *Information, Semantics, & Epistemology*, ed. E. Villanueva. Cambridge: Basil Blackwell, 5–17.

———. 1998. Action and autonomy. In *Cosmos of science*, ed. J. Norton and J. Earman. Pittsburgh: University of Pittsburgh Press, 515–529.

Enc, B., and F. Adams. 1992. Functions and goal directedness. *Philosophy of Science* 59:635–654.

Harman, G. 1977. The nature of morality. New York: Oxford University Press.

Harman, G. 1986. Moral explanations of natural facts—Can moral claims be tested against moral reality? *Southern Journal of Philosophy*, Supplement 24:57–68.

Kim, J. 1991. Dretske on how reasons explain behavior. In *Dretske and His Critics*, ed. B. Mclaughlin. Cambridge: Basil Blackwell, 52–72.

Mackie, J. L. 1973. Truth, probability, and paradox. Oxford University Press.

Miller, R. W. 1978. Methodological individualism and social explanation. *Philosophy of Science* 45:387–414.

———. 1979. Reason and commitment in the social sciences. *Philosophy and Public Affairs* 8:241–66.

———. 1985. Ways of moral learning. *Philosophical Review* 94:507–56.

———. 1987. *Fact and method*. Princeton: Princeton University Press.

———. 1992. *Moral differences*. Princeton: Princeton University Press.

Prior, E., R. Pargetter, and F. Jackson. 1982. Three theses about dispositions. *American Philosophical Quarterly* 19:251–259.

Railton, P. 1986. Moral realism. *Philosophical Review* 95:163–207.

Sayre-McCord, G., ed. 1988. *Essays on moral realism*. Ithaca, NY: Cornell University Press.

Shoemaker, S. 1980. Causality and properties. In *Time and cause*, ed. Peter van Inwagen. Dordrecht: D. Reidel, 109–135.

Sturgeon, N. 1985. Moral explanations. Reprinted in Sayre-McCord 1988, 229–255.

———. 1986. Harman on moral explanations of natural facts. *Southern Journal of Philosophy*, Supplement 24:69–78.

———. 1991. Contents and causes: A reply to Blackburn. *Philosophical Studies* 61:19–37.

———. 1992. Nonmoral explanations. In *Philosophical Perspectives*, 6, Ethics, 1992. Atascadero, CA: Ridgeview Publishing Co.

———. 1995. Evil and explanation. *Canadian Journal of Philosophy*, Supplement 21:155–86.

Tooley, M. 1972. Armstrong's proof of the realist account of dispositional properties. *Australasian Journal of Philosophy* 50:283–87.