MARK SCHROEDER

REASONS AND AGENT-NEUTRALITY

ABSTRACT. This paper considers the connection between the three-place relation, *R is a reason for X to do A* and the two-place relation, *R is a reason to do A*. I consider three views on which the former is to be analyzed in terms of the latter. I argue that these views are widely held, and explain the role that they play in motivating interesting substantive ethical theories. But I reject them in favor of a more obvious analysis, which goes the other way around.

1.   INTRODUCTION

Consider the following sentences:

1   The fact that Katie needs help is a reason to help Katie.
2   There is a reason to help Katie.
3   The fact that there will be dancing at the party is a reason for Ronnie to go to the party.
4   There is a reason for Ronnie to go to the party.

Each sentence uses the word "reason" to express some relation, but the relation expressed by each is different. In **1** it expresses a *dyadic* relation between a consideration[1] and an action, in **2** it expresses a monadic property of an action, in **3** it expresses a *triadic* relation between a consideration, an agent, and an action, and in **4** it expresses a dyadic relation between an agent and an action. But it is natural, nevertheless, to think that these senses of "reason" are not unrelated.

It is natural, for example, to take the "there is" in **2** and **4** at face value, as an existential quantifier. After all, **2** is a consequence of **1** and **4** of **3**, in the same way that we would expect if it were. And if someone tells you that there is a reason to help Katie, it is perfectly fair game for you to ask,

"well, what is it?" So the relationship between **1** and **2**, and between **3** and **4**, ought not to be particularly mysterious. I'm going to take the view that this is correct.[2] So in order to understand **2** we have to understand **1**, and in order to understand **4** we have to understand **3**. What I'll be interested in, in this paper, is therefore the relationship between **1** and **3**. How are they related?

A little bit of vocabulary, then. Because they differ with respect to whether they have an agent as one of their relata, I'm going to say that the dyadic relation expressed by **1** is the *agent-neutral* reason relation, and that the triadic relation expressed by **3** is the *agent-relational* relation. So the reason referred to in **1**, because it stands in this relation, counts as *agent-neutral*, and similarly that in **3** counts as *agent-relational*. And again, I will use these same categories to refer to the *sentences* that *ascribe* agent-neutral and agent-relational reasons. So **1** itself is an agent-neutral *ascription*, while **3** is an agent-relational one. The question before us is: what is the relationship between agent-neutral and agent-relational reasons?

This is not, exactly, the same question as that of what the relationship is between agent-*relative* and agent-neutral reasons. According to the official definition, an agent-*relative* reason is a reason that is a reason for some people, but not for everyone, while an agent-*neutral* reason is one that is a reason for everyone. On this official definition, the distinction between agent-relative and agent-neutral reasons is one *among* agent-*relational* reasons. So according to the official usage, a reason cannot be both agent-relative and agent-neutral. Agent-relative reasons, then, are by definition *merely* agent-relational reasons.

2. THE QUANTIFICATION STRATEGY

The official definition takes a stand on the relationship between agent-relational and agent-neutral reasons. It claims that agent-relational reasons are basic, and that agent-neutral reasons arise when something is an agent-relational reason for

everyone. I like this view, and call it the *Quantification Strategy*. According to the Quantification Strategy, the connection between **1** and **3** is a little bit like the connection between **1** and **2** or between **3** and **4**. Except instead of involving an *existential* quantifier, the connection involves a *universal* quantifier. The Quantification Strategy is basically that **1** involves a universal quantification into the agent-place of **5**:

5 The fact that Katie needs help is a reason for *x* to help Katie.

The Quantification Strategy is a very common view, and many philosophers think that it is obvious. I agree with these philosophers that something much like it is the *best* view to take about this question, and it is half of the purpose of this paper to explain why. But I do not agree that it is obvious. I think, in fact, that it is highly controversial. The fact that it is controversial is evidence that it is not obvious, and the fact that a number of interesting philosophical views involve rejecting it is evidence that it is controversial.[3] The other half of the purpose of this paper is to illustrate how interesting views about a variety of other questions can be supported by taking *other* views about the relationship between agent-relational and agent-neutral reasons.

Before I explain what the views *are* which conflict with the Quantification Strategy, however, let me explain why the Quantification Strategy *is* such a good theory about the relationship between agent-relational and agent-neutral reasons. Any adequate answer to this question is going to have to satisfy two central constraints:

Desideratum 1: Necessarily, for all *r* and *a*, if *r* is an agent-neutral reason to do *a*, it follows that for all agents *x*, *r* is an agent-relational reason for *x* to do *a*.

Desideratum 2: It is not the case that for all *r*, *x*, and *a*, if *r* is an agent-relational reason for *x* to do *a*, then *r* is an agent-neutral reason to do *a*.

If Katie needs help, then there is a reason to help Katie. And if there is a reason to help Katie, then it is a reason for me to help Katie and a reason for you to help Katie. Indeed,

it is a reason for anyone to help Katie. That is what Desideratum 1 tells us. Of course, if the two of us like to dance, we *can* sometimes *say* things like "there is a reason to go to the party tonight" even though we know that there is not an agent-relational reason for Bradley to go there. But these cases are not counterexamples to Desideratum 1, because in such cases we don't mean to be asserting the existence of a genuinely agent-neutral reason. Though we don't mention an agent when we make such assertions, our assertions are merely elliptical. The relation they invoke is manifestly agent-relational, and we are merely saying that there is a reason *for us* to go to the party. All Desideratum 2 says, on the other hand, is that some agent-relational reasons, such as Ronnie's, are *merely* agent-relational.

It's easy to see why the Quantification Strategy is such a natural response to these two desiderata. It trivially satisfies a considerable strengthening of Desideratum 1. And it is clearly compatible with Desideratum 2. Of course, someone could accept the Quantification Strategy and then go on to give a bizarre account of the agent-relational reason relation from which it followed that if R is a reason for X to do A, then for all x, R is a reason for x to do A. But such an account would be treating the agent-relational reason relation as if it had no agent place at all. Any account of that relation whatsoever that made the agent-place non-trivial would satisfy Desideratum 2.

The Quantification Strategy, moreover, also has a number of other virtues. Most notably, it lets us treat all agent-neutral reason *ascriptions* on a par. I've just noted that we can sometimes say "there is a reason to go to the party" and only mean that there is a reason for *us* to go there. But if that is right, then we can think of the Quantification Strategy as simply a special case of this more general phenomenon. In general, we can say that:

Generalized QS: R is a reason to do A $\equiv_{def}$ R is an agent-relative reason for all of [us] to do A.

In some contexts − especially moral contexts − like those in which we can say things like "the fact that Katie needs help is a reason to help her," the context can determine that the scope of who counts as one of "us" must include everyone − even every possible agent. And then when we talk in this way we will be talking about agent-neutral reasons. But in others, like that in which we are discussing our plans for this evening, only some agents − we ourselves − are salient in the context, and Bradley is not.

So the word "reason" and the function of the "there is a reason to" construction isn't simply ambiguous across those cases, even though in one we manage to ascribe the agent-neutral reason relation and in the other we do not. The required difference arises simply as a result of how we restrict the scope of who counts as one of *us*. And this seems like an attractive result, over and above the required desiderata which any theory has to satisfy. It suggests, I think, that we are on the right track.

3. RATIONALISM AND THE INVERSION STRATEGY

Yet as natural as the Quantification Strategy seems, and as obvious as many find it, it is equally obvious that there are very interesting and important views in moral philosophy which are committed to rejecting it. Consider a version of the age-old question of moral philosophy, "why should I be moral?": "why is there a *reason* for me to be moral?" According to a traditional Rationalist kind of response to this question, the answer is, "because *there is* a reason to be moral." Christine Korsgaard tells us[4] that this answer amounts to mere foot-stomping: "Having discovered that obligation cannot exist unless there are actions which it is necessary to do, the realist concludes that there are such actions." But we can do better for the Rationalist than this. After all, the Rationalist is not so silly as to say that there is a reason for you to be moral because there is a reason *for you* to be moral. What he says is that there is a reason for you to be moral because *there is a reason to be moral*. So on

the face of it, he is not stomping his foot at all, but suggesting that your agent-*relational* reason to be moral is explained by an agent-*neutral* reason to be moral.

This answer to the problem can seem hard to credit, if we accept the Quantification Strategy. For if we accept the Quantification Strategy, then we think that there being an agent-neutral reason to be moral simply *consists* in there being an agent-relational reason for each person to be moral. So we *do* think that the Rationalist answer sounds merely like foot-stomping. It sounds a bit like saying that you have a reason to be moral because you *and Alice both* have reasons to be moral. But if it sounds like that to us, that is only because we are convinced by the Quantification Strategy. The Rationalist, I'm suggesting, if we are to take his answer to the question, "why is there a reason for me to be moral?" seriously, has to be understood as holding that agent-neutral reasons explain agent-relational reasons, rather than con-versely.[5] Since Rationalism of this kind is a possible view, gives us a very interesting answer to one of the central tradi-tional questions of moral theory, and is committed to deny-ing the Quantification Strategy, the Quantification Strategy can't be uncontroversial.

Rationalists of this stripe are committed to *inverting* the order of explanation of the Quantification Strategy. They hold that the agent-relational reason relation needs to be understood in terms of the agent-neutral one, rather than conversely. Let us call all views of this kind versions of the *Inversion Strategy*. The Rationalist needs a particular version of the Inversion Strategy, which I will explain in sections 8, 9 and 10. But first, let's look at two more versions of the Inver-sion Strategy. Each of these is useful to another important moral theory, just as the Standard Model is useful to this kind of Rationalism.

4.   KANTIANS AND THE POINT OF VIEW THEORY

Rationalism of this kind is perhaps the view most obviously committed to rejecting the Quantification Strategy. But it is not

the only such view. Consider an updated version of Kant's argument that any rational agent is required to abide by the Categorical Imperative. According to Kant, this commitment must be derived from the bare concept of rational agency, which involves only the idea of acting according to laws or, as an updated Kantian might put it, *reasons*. To be an agent, you have to see yourself as acting for reasons – considerations which are, in fact, reasons to act. But if you have to see yourself as acting in accordance with reasons in favor of one action or another, then we can apply Desideratum 1, which tells us that if there is a reason in favor of some action, it must be a reason for *each* person to perform that action. So it follows from the Kantian argument that in order to be an agent, you have to see yourself as acting for reasons *that are reasons for everyone, and not just you* – i.e., you have to be able to conceive your actions as in accordance with *universal* reasons, or, substituting back into Kant's language, universal *laws*. And that is just the Universal Law formulation of the Categorical Imperative, or as close to it as I can find an argument.

But there is one particularly worrisome step in this argument. And I don't mean the step which claims that in order to act for reasons, you have to see yourself as acting for things which you take to be reasons, although I find that also questionable. I mean the assumption that when you take yourself to be acting for a reason, it must be an agent-neutral reason that you take yourself to be acting for. Why couldn't you take yourself to be acting for an agent-relational reason? One might think that if you are Ronnie, and you are clear-headed, you can go to the party for the reason that there will be dancing there, even though you are perfectly well aware that it isn't a reason for Bradley to go there. At most what you have to do, is to recognize that it is a reason *for you* to go there. Nothing more, it seems, could be required by the mere concept of rational agency.

But it's not hard, I think, to supplement the Kantian argument with a view about agent-relational reasons which rules this possibility out – a view which, however, conflicts with

the Quantification Strategy. And the view is one which it is hard not to find tempting to attribute to a variety of Kantian thinkers.[6] It is what I call the *Point of View* theory. According to the Point of View theory, the "for Ronnie" in "there is a reason for Ronnie to go to the party" works like a propositional operator on the proposition that there is a reason to go to the party. So saying that there is a reason for Ronnie to go to the party is like saying that *from Ronnie's point of view* it is the case that there is a reason to go to the party. This conflicts with the Quantificational Strategy because the Quantificational Strategy accounted for the agent-neutral reason relation in terms of the agent-relational one. But the Point of View theory accounts for the agent-relational relation in terms of the agent-neutral one.

   The Point of View theory is neutral about how to interpret what "points of view" are. But imagine, as suggested by the "point of view" terminology, that a necessary condition on it's being the case from Ronnie's point of view that *p* is that Ronnie *believes* that *p*. This would do wonders for the Kantian argument. Suppose, then, that you believe that there is a reason for you to go to the party, but not an agent-neutral reason to go to the party. This is what would have to be the case in order to create trouble for the Kantian argument. Then you are committed, by this version of the Point of View theory, to the claim that you believe that there is an agent-neutral reason to go to the party, but at the same time you actually believe that there is no agent-neutral reason to go to the party. And that, though not a contradiction outright, is at least a *Moorean* contradiction. So if you are to avoid Moorean contradictions, then you can't consistently act for reasons unless you take the reasons for which you act to be agent-neutral reasons − to be reasons equally for everyone. And that would give us the Formula of Universal Law. This doesn't prove that any Kantian is committed to the Point of View theory − perhaps my reconstruction of the "updated" argument is nothing like the version any Kantian has ever accepted − but it does illustrate how *useful* it would be to a Kantian.

5.   CONSEQUENTIALISM AND THE PROPOSITIONAL THEORY

Kantians and Rationalists are still not the only ones to find denying the Quantification Strategy useful. The simplest alternative to the Quantification Strategy derives from the idea that the reason relation is not a relation between a consideration and an *action*, but between a consideration and a *proposition*. On this view, to say that there is a reason for Ronnie to go to the party is to say that there is a reason in favor of the following *proposition*: that Ronnie goes to the party. According to some who take this view,[7] the "for Ronnie" in "there is a reason for Ronnie to go to the party" does double work. It both qualifies the reason, which is agent-relational, *and* qualifies the proposition the reason is in favor of. So it is like saying that there is a reason for Ronnie in favor of the proposition that Ronnie goes to the party.[8] But according to others, we don't need the "for Ronnie" to do double-work. And this gives them a way of accounting for apparently agent-relational reasons in terms of agent-neutral ones. For there to be a reason for Ronnie to go to the party, on this view, is simply for there to be an agent-neutral reason in favor of the proposition that Ronnie goes to the party. Call this the *Propositional Theory*.

The Propositional Theory is particularly useful for *Consequentialists*. Consider one of the chief differences between Consequentialism and common-sense morality. According to common sense, it can sometimes, at least in principle, happen that you ought not to do something, even if by doing it you can prevent two or more other people from doing exactly the same thing. Such cases are called *agent-centered restrictions*. But according to Consequentialists this cannot happen. Indeed, according to some Consequentialists, it is quite hard to understand how there could be such things as agent-centered restrictions. The Propositional Theory facilitates one straightforward argument that there couldn't be.

So suppose that by doing some wrong action, you can prevent two or more other people from doing exactly that action. Since the action is wrong, there is a reason for you

not to do it. And since by stipulation it is exactly the same action, there are reasons for each of the other two or more people not to do it. And these reasons are equally good. By the Propositional Theory, this means that there is an agent-neutral reason against your doing the action, and an agent-neutral reason against each of the others performing this action, and these reasons are equally good. But since the reasons are agent-neutral, they bind you equally. One weighs against your doing the action, but at least two weigh in favor of it. So the weight of reasons weighs in favor of your doing the wrong action. And so, concludes the argument, this must be what you ought to do.

   This is an elegant argument against the possibility of agent-centered restrictions. But the Propositional Theory was crucial to running the argument. If we take the view that reasons weigh in favor of *actions*, rather than *propositions*, then agent-centered restrictions give us no trouble. In that case, we agree that if the action is *wrong*, then there is a reason for each agent not to do it. So there is a reason *for you* not to do it, and a reason for each of the people who you could prevent, not to do it. And the reasons can all be equally good, but the reasons for the others not to do it do not translate into equally good reasons for you to prevent them from doing it. For they aren't agent-neutral reasons in favor of the *state of affairs* that they not do it, but agent-*relational* reasons *for them* not to do it. It is only the Propositional Theory which makes us conflate these two things. So on this view it makes perfect sense to think that you still ought not to perform the wrong action.

   There may be other and better reasons to be puzzled about agent-centered restrictions; I can't evaluate that here.[9] But it should be clear that the Propositional Theory would be *very useful* to the Consequentialist. For the Propositional Theory definitely validates this argument for the impossibility of agent-centered restrictions. And I find it hard not to think that this argument is something like what philosophers who are skeptical about agent-centered restrictions often have in mind, at least implicitly. So once again we can see that an

issue important to a substantive and interesting moral theory is tied up with taking a view about the connection between agent-relational and agent-neutral reasons which conflicts with the Quantification Strategy.

We now know about three very interesting substantive issues in moral philosophy which are intimately related, to one degree or another, to views about the connection between agent-relational and agent-neutral reasons that conflict with the Quantification Strategy. Such views are all versions of the *Inversion Strategy*. So now let us turn to developing these versions of the Inversion Strategy enough to see whether they are, in fact, viable views. Though I have tried in sections 3, 4, and 5 to argue that they are *substantive* and *interesting*, I will now explain why they are, nevertheless, inadequate, and in ways that are quite independent of their implications discussed in sections 3, 4, and 5.

6.   WHAT'S WRONG WITH THE PROPOSITIONAL THEORY?

So let's take the three versions of the Inversion Strategy in reverse order, from easiest to hardest, and the Propositional Theory first. The Propositional Theory, I think, is clearly the worst of the available views. It relies, for starters, on the false view that reasons are in favor of *propositions* instead of *actions*. And it still fails by the lights of our two desiderata on any answer to our original question, because it has no plausible candidate for what the proposition is that the reason is in favor of in our original paradigmatic agent-neutral reason ascription, "there is a reason to help Katie".

Set aside the question of whether reasons weigh in favor of actions or propositions.[10] Let's just look at how the Propositional Theory is to account for our two desiderata. It's easy to see what the proposition is that the reason is in favor of, in the case of "there is a reason for Ronnie to go to the party." It is the proposition that Ronnie goes to the party. But what is the proposition that the reason is in favor of in the agent-*neutral* ascription, "there is a reason to help Katie"?

We want Desideratum 1 to turn out to be true, so consider the candidate that it is the proposition that *everyone* helps Katie. If there is a reason in favor of the proposition that everyone helps Katie, then it is certainly a reason in favor of the proposition that you help Katie. After all, your helping Katie is necessary for it to be the case that everyone helps Katie. So that would explain Desideratum 1. But clearly there can be a reason to help Katie even if there is no reason in favor of the proposition that *everyone* helps Katie. After all, if Katie needs help, she doesn't need *everyone* to help her − she just needs *enough* help. So this view about what proposition the agent-neutral reason is in favor of requires something far too strong.

On the other hand, perhaps the reason to help Katie is a reason in favor of the proposition that *someone* helps Katie. But then why should we think that there is a reason for *each* person to help Katie? In particular, why should we think that there is a reason in favor of the proposition that *you* help Katie? Why must a reason in favor of the proposition that *someone* helps Katie necessarily be a reason in favor of the proposition that *you* help Katie? This question is harder to answer, if we keep in mind just how plausible it is that a reason must count in favor of each of the necessary consequences of the proposition that it is in favor of. If that is right, then if there is a reason in favor of the proposition that Ronnie goes to the party, it must also be a reason in favor of the proposition that someone goes to the party. But then our explanation of your reason to help Katie only works if it *also* follows that Ronnie's reason is also a reason in favor of the proposition that *Bradley* goes to the party. And then the Propositional Theory fails on account of Desideratum 2.

The only way to save the Propositional Theory is to abandon the view that a reason must count in favor of the necessary consequences of the proposition it is in favor of. But the idea that the force of reasons transfers from ends to *necessary* means is one of the bedrocks of thinking about how the force of reasons transfers from one thing to another. So the Propositional Theory looks hopeless for accounting for the two most

obvious central desiderata for any theory about the connection between agent-relational and agent-neutral reasons. Correlatively, Consequentialists should look harder for a better argument that there is something puzzling about agent-centered restrictions, and non-Consequentialists should check Consequentialist reasoning carefully, to see whether it covertly relies on the Propositional Theory − or anything very much like it.

7.   WHAT'S WRONG WITH THE POINT OF VIEW THEORY?

According to the Point of View theory, saying that there is a reason for Ronnie to go to the party is like saying that from Ronnie's *point of view*, there is a reason to go to the party. In principle, the Point of View theory isn't committed to thinking that *points of view* are anything like points of view in any colloquial sense, although the version of the Point of View theory required in order to make my updated Kantian argument work *did* require this. But in principle, all that the Point of View theory is in general committed to holding, is that the "*for Ronnie*" works like a propositional operator on a content that is about agent-*neutral* reasons.

Yet it is relatively easy to rule out a wide range of versions of the Point of View theory, on the grounds of our two desiderata for any viable answer to our original question. For according to Desideratum 1, it must follow from the fact that there is an agent-neutral reason to help Katie that there is an agent-relational reason for you to help Katie. So according to the Point of View theory this is like saying that if there is a reason to help Katie, it must be the case that *from your point of view* there is a reason to help Katie. And that is like saying that agent-neutral reasons are transparent to your point of view. And if that is so, then points of view can't be anything like points of view, in the conventional sense. Reasons certainly aren't transparent to our beliefs − otherwise how could there be so much disagreement about what the reasons are?

Compare again: in order to get the updated Kantian argument to work, I claimed that it had to be a *necessary* condition

on it's being the case that *p* from Ronnie's point of view that Ronnie *believed* that *p*. But Desideratum 2 requires that it be possible that there is a reason for Ronnie to go to the party, even though there is no reason to go to the party. So according to the version of the Point of View theory required in order to make the updated Kantian argument work, it follows that it must be possible that Ronnie sometimes believes that there is a reason to do something which there is no reason to do.

So put together, we have two constraints on how points of view must work, derived from our two desiderata. First, agent-neutral reasons must somehow be transparent to points of view. Yet there must also be *further* reasons from points of view which are not, in fact, actual agent-neutral reasons. These two constraints rule out any ordinary understanding of what points of view are that I know about, including that necessary in order to run the updated Kantian argument. I don't claim that this refutes the Point of View theory in general, but I do claim that it undermines much of the source of its intuitive appeal — the appeal of thinking that when you say that there is a reason for Ronnie to go to the party, you are saying something like that for Ronnie, it is as if there is a reason to go to the party.

The challenge which remains for the still-standing versions of the Point of View theory is not simply to demonstrate *consistency* with our two desiderata on any viable theory. It is to provide an *explanation* of *why* these two desiderata, and particularly the first, are true. The Quantification Strategy isn't merely consistent with the two desiderata. It gave us a good explanation of *why* Desideratum 1 is true. For Desideratum 1 is a consequence of the Quantification Strategy. I've now argued that *some* plausible-sounding versions of the Point of View theory are inconsistent with the two Desiderata, and left open whether other versions are consistent with it. Yet a *good* theory about the relationship between agent-relational and agent-neutral reasons should actually *explain* the most central and obvious entailments connected with these two relations. *If* there is an adequate version of the Point of View theory — a possibility that I don't rule out — it will have to do this for us.

One final unpromising feature of the Point of View theory. If "for Ronnie" is simply a propositional operator, then it should in principle be able to be applied to contents with other sorts of content, not about agent-neutral reasons. For example, it should make sense to say that "there is a tunnel for Ronnie under the English Channel" or that "for Ronnie the Potomac floods in the spring". Whatever account the Point of View theorist gives us about this operator, it has to be one which makes these claims make as much sense at the one that "there is a reason for Ronnie to go to the party", in *addition* to explaining the two desiderata. I'm not optimistic.

8.   THE SUBSUMPTION ACCOUNT

The third version of the Inversion Strategy − the one that is necessary for the Rationalist to run his response to the question, "why is there a reason for me to be moral?" − is more sophisticated than the two I've just set aside in sections 6 and 7, and I think many more people accept it, at least implicitly. It will also require much more care, in order to see what is deeply problematic about it. A first attempt at such an account might go like this: take the agent-neutral reason relation as relatively primitive, and say:

Subsumption #1: R is an agent-relational reason for X to do A $\equiv_{def} \exists b R$ is an agent-neutral reason to b & A-ing is a way for X to do b.

The thought behind this attempt is that if we are to account for the agent-relational reason relation in terms of the agent-neutral reason relation, then for each agent-relational reason we need some agent-neutral reason. In order to satisfy Desideratum 1, it had better be the *same* consideration, so we build that in to our account. But in order to satisfy Desideratum 2, we need some sort of leeway between what R is an agent-neutral reason to do, and what it is an agent-relational reason for someone in particular to do. *Ways* of doing something are supposed to do this work for us. In this way, we *subsume* agent-relational reasons under agent-neutral reasons. There is an agent-relational reason for you to do something, when it is

a *way* of doing something that there is an agent-neutral reason to do.

So, for example, there may be an agent-neutral reason to eat healthily. But what it takes to eat healthily may differ from agent to agent. Diabetics and Atkins dieters need to eat in different ways from the rest of us, in order to eat healthily. So it follows from Subsumption #1 there is an agent-relational reason for Diabetics to eat in ways that there is no agent-relational reason for Atkins dieters to eat, and vice-versa. We can think of *ways* talk as a generalization on *means-end* talk, if that's helpful.

This account actually does remarkably well. So long as we assume (as we should) that A-ing is always a way for X to do A, the account satisfies Desideratum 1. Moreover, so long as we assume that it is sometimes the case that A-ing is a way for X to do B, but not a way for Y to B, the view also satisfies Desideratum 2. Yet it still gets something deeply wrong. To see why, we need to look more carefully at how this account will deal with the case of Ronnie and Bradley (Bradley, recall, is the one who doesn't share Ronnie's merely agent-relational reason to go to the party, because he can't stand dancing). It must find some action-type B such that going to the party is a way for Ronnie to do B, but not for Bradley. Let's suppose for the sake of argument that this action is *doing what one likes to do*. The account will assume that there is some $\mathscr{R}$ such that $\mathscr{R}$ is an agent-neutral reason to do what one likes to do. Let pass for now exactly what $\mathscr{R}$ might be. Whatever it is, according to this account, $\mathscr{R}$ is the agent-relational reason for Ronnie to go to the party. Intuitively, however, the reason for Ronnie to go to the party is that there will be dancing there – something that is special to Ronnie's case. Intuitively, Ronnie has some reason to go to the party that is not the same as Brett's reason to practice his bass or Vera's reason to play chess, things that Brett and Vera like to do.

The problem is that though this account vindicates claims of the form, "there is a reason for X to do A," it fails miserably to capture truths of the form, "R is a reason for X to do A". Because "there is a reason for X to do A" transparently

quantifies[11] over reasons, we really are committed to things being reasons. So our theories had better be able to account for those reasons. This point is usually ignored in the literature, but that doesn't make it irrelevant.[12] So this first attempt at the inversion strategy yields unacceptable results.

A second attempt turns out to be less satisfactory, but at least to start us in the right direction. The idea is this: the first account got us into trouble because it made all of the agent-relational reasons that were accounted for using the same agent-neutral reason come out to be the same – and the same as that agent-neutral reason. So we should instead build into our account a way for different agent-relational reasons to derive from the same agent-neutral reason, in a way that will reflect our intuitive judgments about what Ronnie's, Brett's, and Vera's reasons are. The account that gets this part roughly correct says:

Subsumption #2: R is an agent-relational reason for X to do A $\equiv_{def}$ $\exists s$ $\exists b$ s is an agent-neutral reason to do b & A-ing is a way for X to do b & R is an essential part of the explanation of why A-ing is a way for X to do b.

Why is going to the party a way for Ronnie to do what he likes? Well, it's because 1) he likes to dance, and 2) there will be dancing at the party. Each of these facts figures essentially in this explanation, and each is intuitively a reason for him to go to the party. So this account explains what is going on in Ronnie and Bradley's case, as long as we again assume that there is some $\mathscr{R}$ which is the agent-neutral reason to do what one likes.

Unfortunately, however, this account fails by the lights of Desideratum 1. It would seem that nothing figures essentially in the explanation of why helping Katie is a way for X to help Katie – after all, this is trivial. But then it follows that nothing *at all* about agent-relational reasons follows from the fact that there is an agent-neutral reason to help Katie – namely, that she needs help. Nothing turns out to be a reason for anyone to help Katie, in virtue of the existence of some agent-neutral reason to help Katie. So this account not only gets wrong what various agents' agent-relational reasons are,

it also yields the wrong results about when they have agent-relational reasons. In this regard, it is worse than the previous account. It does, however, help to illustrate what motivates the account in the next section, and to show why the Subsumption Account needs to resort to such a view.

9.   THE BEST SUBSUMPTION ACCOUNT

The trouble with the Subsumption Account is that if it is to validate the Rationalist's view, then for every agent-relational reason there must be some agent-neutral reason from which it derives in some way. Desideratum 1 puts pressure on the account to allow that the agent-neutral reason should just be the same consideration as the agent-relational reason that derives from it. But this guarantees the wrong results when we look more closely at Ronnie and Bradley's case. What we need, is a way of allowing that these considerations should be the same in just those cases to which Desideratum 1 applies, but different in all of the cases which will worry us with respect to Desideratum 2. There is a unified way of getting this result, which does considerably better than the first account we considered, and it borrows from each of the previous two accounts. It says:

Subsumption #3: R is an agent-relational reason for X to do A $\equiv_{def}$ $\exists s$ $\exists b$ s is an agent-neutral reason to do b & A-ing is a way for X to do b & R is the conjunction of s with each truth that figures essentially in explaining why A-ing is a way for X to do b.

This account simply takes as X's agent-relational reason to do A the conjunction[13] of all of the things that counted as X's agent-relational reasons to do A according to each of the last two accounts.

If we continue to suppose that nothing figures essentially in explaining why helping Katie is a way for X to help Katie, then this account satisfies Desideratum 1. The fact that Katie needs help is an agent-neutral reason to help her, and we don't have to conjoin anything with it in order to get an agent-relational reason for X to help her. Meanwhile, if we continue to suppose that there is some $\mathscr{R}$ that is an agent-neutral reason to do what

one likes, then the account can also do tolerably well with respect to Ronnie and Bradley's case. $\mathscr{R}$ conjoined with the fact that Ronnie likes dancing and with the fact that there will be dancing at the party is an agent-relative reason for Ronnie to go there.

Of course, what we wanted was that each of the latter was an agent-relational reason for Ronnie to go there, and not merely that their conjunction with the unknown $\mathscr{R}$ would be. So even on this account, our ordinary ascriptions of reasons are not literally true. It improves on the first inversion account by allowing that Ronnie and Brett and Vera all have different agent-relational reasons, but it doesn't quite make the fact that there will be dancing at the party turn out to literally be a reason for Ronnie to go there. All it yields is that this is one of the conjuncts in a reason for Ronnie to go there, or part of one of Ronnie's reasons to go there.

But this may be good enough. True, sometimes we can say that the fact that there will be dancing at the party is a reason for Ronnie to go to the party, and sometimes we can say that the fact that Ronnie likes dancing is a reason for him to go to the party. But when Ronnie is deciding to go to the party, it seems, he shouldn't add the weights of these two reasons together – they seem to count in the same way towards his going to the party. So perhaps there really is some sense in which they are really the same reason.[14] The third version of the inversion strategy gets this intuition right. According to it, each ascription functions to ascribe literally the same reason, surface appearance to the contrary.

There's a familiar idea from the literature which explains why we can use ascriptions like "the fact that there will be dancing at the party is a reason for Ronnie to go to the party" and "the fact that Ronnie likes dancing is a reason for Ronnie to go to the party" in order to ascribe the same reason.[15] It is that contextual conversational pragmatics[16] determine which information about someone's reason is needed in the conversation. If we already know that Ronnie likes dancing, then it makes sense to say that the fact that there will be

dancing at the party is a reason for him to go there. Even though this isn't literally true, we have everything that we need in order to construct what Ronnie's complete reason really is, so the rest isn't worth mentioning. It may also be true that we can *infer* that this fact would only be citable as a reason for him to go there, if he liked dancing. In that case, though we do not know it already, we can infer that Ronnie likes dancing, and so this part of his reason again turns out not to be worth mentioning. Likewise, when only the fact that Ronnie likes dancing is mentioned, we either already know, or are able to infer, that there must be dancing at the party. That is why we don't have to mention them in ordinary-language reason ascriptions.

So as far as I can see, the Subsumption Account is going to have to be committed to some such contextualist pragmatics, in order to make sense of our actual reason ascriptions. This isn't a bad thing. As I noted, it helps to make sense of the idea that the fact that there will be dancing at the party and the fact that Ronnie likes dancing shouldn't count separately, for him, in favor of going there. But I've now argued that the Subsumption Account is committed to more than the claim that these are parts of Ronnie's reason. It is committed to the claim that the mysterious $\mathscr{R}$, which is the agent-neutral reason to do what one likes, is also part of Ronnie's reason. Not only that, but in order for its conversational pragmatics to work, it is committed to holding that it is so obvious what $\mathscr{R}$ is, that it is not worth mentioning. And it is this mysterious agent-neutral reason that turns out to make the most trouble for inversion views, and for this last and most promising inversion view, in particular.

10.   A MYSTERIOUS REASON

So far, we've been assuming that inversion strategies can appeal to the existence of some $\mathscr{R}$ that is an agent-neutral

reason to do what one likes, in order to account for merely
agent-relational reasons that arise from likes and dislikes.
Unfortunately, making good on this assumption is not easy.
It's hard to see what single feature of the world it could be
that counts, for each person, in favor of that person's doing
what she likes.[17]

It's not enough to assure ourselves that it's obvious that
"everyone has a reason to do what she likes." For from

Transparent: $\forall x \forall a$ If x likes to do a then $\exists r$ (r is a reason for x to do a).

it certainly doesn't follow that

Opaque: $\forall x \exists r$ r is a reason for x to do what x likes to do.

let alone that

Needed: $\exists r \forall x$ r is a reason for x to what x likes to do.

Transparent says that when each person likes to do some-
thing, there is a reason for her to do *it*. Opaque says that
each person has a general reason to do the following action
(if there is such a thing): do what she likes. Needed says the
same as Opaque, but asserts that this is an agent-neutral rea-
son, the same for everyone. Transparent gives us an obvious
reading of "everyone has a reason to do what she likes" that
is obviously true.[18]

But the truth of Transparent is compatible with there being
no *general* agent-relational reason, even just for Ronnie, to
do whatever he likes. It certainly does not entail the truth of
the Opaque reading of the intuitive, "everyone has a reason
to do what she likes." Moreover, even if it did, Opaque is
compatible with each person having different agent-relational
reasons, so that there is no single consideration to *be the*
agent-neutral reason to do what one likes. Yet this is what
the Subsumption Account requires. So let me illustrate:

The truth of the matter is that the fact that there will be
dancing at the party might be an agent-relational reason for
Ronnie to go there – and going there, after all, is doing what
he likes. The fact that practicing his bass will improve his
chances of becoming a rock star is a reason for Brett to

practice his bass – and improving his chances of becoming a rock star, after all, is what he likes. And so on. The fact that Plato is a philosopher is a reason for Brett to study Plato – after all, Brett likes to study philosophers. Brett's reason to study Plato and his reason to practice his bass are both reasons that he has in virtue of what he likes, but they aren't the same reason. So it would follow that everyone has an agent-relational reason to do each of those things which happen to be among the things that she likes to do, but not that there is any general agent neutral reason to do what one likes. This makes sense of the intuitively obvious claim without appeal to an agent-neutral reason such as $\mathscr{R}$. The claim *is* perfectly true – *on* the *transparent* reading, the one on which the description, "what she likes," takes *wide* scope over the reason-ascription.

Now there *may* be a *sense* in which, when Ronnie has a reason to go to the party because he likes to dance, and Brett has a reason to practice his bass because he likes getting better at playing his bass, Ronnie and Brett have the same reason. If there is, it is the same weak sort of sense in which, when Ronnie believes that he is in New York, and Brett believes that *he* is in New York, they believe the same thing. There is even a sense in which Ronnie and Brett have their reasons *because* they are the "same" in this way.

But this isn't enough for the Subsumption Account to work. Even the Quantification Strategy can allow that Ronnie and Brett have these reasons because there is some true generalization of the form,

True Generalization: $\forall x \forall a \forall b$ If x likes to do b and a-ing is a way for x to do b, the fact that a-ing is a way for x to do b is an agent-relational reason for x to do a.

And even the Quantification Strategy can allow that this is *why* the fact that Ronnie likes to dance is a reason for him to go to the party. For this might be part of the account of what it takes to have an *agent-relational* reason to do something. But that doesn't mean that the Quantification Strategy is appealing to a further, *agent-neutral*, reason in explaining

Ronnie and Brett's agent-relational reasons. It's just allowing some true generalizations about what agent-relational reasons they have.

No matter what $\mathscr{R}$ turns out to be, it is going to have to be a truth, a fact about the world, in the way in which the fact that Katie needs help is a fact about the world, and the fact that there will be dancing at the party is a fact about the world. And it has to be the sort of thing that intuitively counts, for each and every person, in favor of her doing what she likes. I simply have no idea what this fact is supposed to be. That doesn't mean that there's no such thing; it just means that the Subsumption Account owes us something significant, in order to make good on its promise to deal adequately with the case of Ronnie and Bradley.[19]

The going version of the Subsumption Account, however, is committed to more than this. For it is committed to some kind of contextualist pragmatics, in order to explain why it's not necessary for us to mention *all* of Ronnie's reason, in order to properly ascribe something as a reason for him to go to the party. Now all of the stories like this that I know of work by explaining why listeners in the context are able to infer the rest of the information about what Ronnie's reason is, without needing to be told. And this works quite well, so long as we only suppose that the fact that Ronnie likes to dance and the fact that there will be dancing at the party make it into Ronnie's agent-relational reason to go to the party. But in fact, in order to solve the problems with the second inversion account, we had to make sure that $\mathscr{R}$, the agent-neutral reason to do what one likes, *also* makes it into his agent-relational reason, and in the same way as these other two facts. But according to the contextualist pragmatics, that should mean not only that there really is some $\mathscr{R}$ that is a reason to do what one likes, but that it should be so obvious that we never need to mention it. And that seems to be a fairly strong commitment. It's certainly not obvious to *me* what $\mathscr{R}$ is.

Still, suppose that it's obvious to everyone that there is *some* $\mathscr{R}$ such that $\mathscr{R}$ is an agent-neutral reason to do what one

likes. Would that be enough to get the pragmatic contextualist account to work? Perhaps. The explanation couldn't be that the reason we don't have to mention $\mathscr{R}$, however, is that we already know what it is. But that leaves room for other kinds of explanation. Still, I think the considerations in this section show that we need to be careful about whether, when it seems obvious that everyone has a reason to do what she likes, this is really what is being claimed. I truly do not find it obvious that there is such an $\mathscr{R}$, and neither, I take it, should anyone else who accepts the Quantification Strategy. The real problem with the Subsumption Account, therefore, is that the contextualist pragmatics it needs seems to require that the Subsumption Account itself be uncontroversial. But nothing about the Quantification Strategy requires that the Quantification Strategy be uncontroversial. It just requires that it be true. And that is an undeniable theoretical advantage.

11.   CONCLUDING REMARKS

In this paper, I've characterized the distinction between agent-neutral and merely agent-relational reasons in a theory-neutral way which, unlike the official definition of agent-neutral reasons, does not presuppose the controversial Quantification Strategy. I've defended a generalized version of the Quantification Strategy, and shown why it suits the criteria for an account of the relationship between agent-relational and agent-neutral reasons so well. And I've developed, and dismissed, three competing accounts – versions of the Inversion Strategy. But I've argued that the question is one of substance, not one to which to take answers for granted. It is all too often the taking for granted of answers to questions like this one, which underwrites deep disagreements about a variety of substantive issues in moral theory. And when that turns out to be the case, there is nothing for it but to investigate the adequacy of these views not by the lights of their controversial implications for ethical theory, but in terms of

the most obvious constraints on the question to which they are supposed to provide answers. Sometimes, and I think in this case this is true, the *interestingness* of these views is not enough to overcome their theoretical inadequacies.[20]

NOTES

[1] "Consideration" is the usual fudge-word for the kind of thing which can be a reason. See, for example, the use of Scanlon (1998). But there are good reasons to think that properly speaking, considerations are something like facts or true propositions.

[2] This has been prominently, although not exactly explicitly, disputed. See Nagel's account of reasons in *The Possibility of Altruism*. Nagel appeals to claims of the form, "there is a reason for Ronnie to go to the party" in his account of claims of the form, "*R* is a reason for Ronnie to go to the party." So although he doesn't say so explicitly, he is obviously committed to holding that "there is" in "there is a reason for Ronnie to go to the party" is not an existential quantifier.

[3] And the fact that it is controversial shows that we need a more theory-neutral way of distinguishing agent-relative and agent-neutral reasons than the official definition. I've just given such a theory-neutral way of drawing the distinction.

[4] Korsgaard (1996, 34). Korsgaard goes on to say what she thinks is problematic about this kind of Rationalism, which she called "realism": "In a similar way, if someone falls into doubt about whether obligations exist, it doesn't help to say, 'ah, but indeed they do. They are *real* things'" (38). And again: "And that is the problem with realism: it refuses to answer the normative question. It is a way of saying that it can't be done" (39).

[5] If this version of the Rationalist's answer to the question "why should I be moral" doesn't sound familiar to you, try the version in which the questioner asks, "why ought I to be moral?" and the Rationalist responds that being moral is *required* or *right*. Here, again, he appears to be using an agent-neutral normative concept in order to explain an agent-relational one. And again, many would differ. Many would think that what it is for an action to be right is just for it to be the case that everyone ought to do it.

[6] Korsgaard, most obviously. See, in particular, Korsgaard (1997). It is also tempting to read Scanlon this way, when he insists that a reason is a consideration which "counts in favor" of some attitude. Scanlon must not accept the account outlined in section 5, since he thinks that reasons count in favor of attitudes. But he also must not accept the Quantification Strategy, because then he would have to say that a reason merely counts in favor of an action, *relative* to some agent. See Scanlon (1998).

⁷ For example, according to John Broome (2004, 28−31).

⁸ I explain some of the hard problems for the analogous view about "ought" in Schroeder (2004).

⁹ Well, okay, but just one pass. There are certainly other *reasons*, but I'm skeptical about whether they are better. For example, the argument might turn on the assumption of some kind of *teleological* explanation of the reasons involved. Such arguments often do work this way − the proponent tries to convince us that if the action is wrong for you and each of the others to do, then that must be because there is something *bad* about any of you doing that action. But surely this kind of teleological assumption should be more obviously controversial in this context. Deontology is often characterized as precisely the view which rejects this kind of teleological explanation of reasons in terms of value or disvalue. So since the teleological argument has more obviously controversial premises, I take it that it is not a better motivation for the Consequentialist's view. See Scheffler (1982), Kagan (1989), and Smith (2003).

¹⁰ I've addressed the question for the case of "ought" in Schroeder (2004), and hope to take up further arguments for the view on a future occasion.

¹¹ On the face of it, Thomas Nagel denies this (see section 1). But it's not as though he denies that there are things that are reasons − on the contrary, he commits to an account of what sort of things they are. Much argument would be required to establish that we shouldn't think that when there is a reason to do something, something is the reason to do it.

¹² Ignoring what someone's reasons *are* to do something *can* be a way of simplifying discussion of some related philosophical questions. But some of the issues that doing so glosses are important, and this is one. We can't simply ignore it all of the time.

¹³ Another version of this same account might not appeal to *conjunction*. It might make the reason out to be the fusion of each of the other facts, or something like that. It's all the same, for what I have to say.

¹⁴ I don't myself hold this view; I'm merely articulating why things don't look so bad for Subsumption #3. I've articulated my contrasting view in chapter 2 of *Slaves of the Passions*, forthcoming from Oxford University Press.

¹⁵ See, for example, Davidson (1980). Davidson isn't talking in that paper about normative reasons − the kind of reasons in which we're interested − but rather about *motivating* reasons. Still, he didn't distinguish clearly, and neither did many who followed him. As a result, it's not uncommon to see philosophers supposing without argument that someone's normative reason is really the entire sufficient condition for it to be the case that she has a reason. This is what Thomas Nagel says, for example, in *The Possibility of Altruism* (1970). But it's worth noting that pragmatic contextualism about reason ascriptions is also compatible

with, though not a commitment of, the Quantification Strategy. The Quantification Strategy, however, *can* make the different reason ascriptions all come out to be literally true.

[16]   Indeed, the idea is so familiar that some philosophers have corrected me, when I have said things like, "intuitively, the fact that there will be dancing at the party is a reason for Ronnie to go there," and told me that this is obviously only part of Ronnie's reason. These philosophers find the contextualist pragmatics so obvious that they've ceased to see the benefit of an account that makes ordinary ascriptions literally true.

[17]   Actually, it's also hard to characterize what the action could be that it is a reason to do, such that every action that needs to be explained by this reason really gets subsumed by it, *and* leads to a satisfactory explanation of the resulting reason. See Schroeder (forthcoming a) for details. Here, I'm only concerned with what $\mathscr{R}$ is supposed to be, because in order to get the conversational pragmatics to work, the Subsumption View has to assume not only that there *is* such an $\mathscr{R}$, but that we all *know* what it is − that it is so obvious as not to be worth mentioning.

[18]   Here I bracket the question of whether the connection between desires and reasons holds only under some further restriction. The important claim I am making is not that Transparent is obviously true when read unrestrictedly, but rather that it is *at best* obvious that Transparent is true. The point is simply that Opaque and Needed are much stronger.

[19]   Some possible candidates which I find unpromising: $\mathscr{R}$ is the fact that sometimes doing what one likes is part of the *good life* for human beings. Or perhaps $\mathscr{R}$ is the fact that other things being equal it is better that more people to do what they like more of the time. If we say the former, we can't say that what makes the good life good is that there are reasons to prefer it, because then we're back in the water when it comes to explaining why this is the good life, looking for another agent-neutral reason to prefer a life in which one sometimes does what one likes. Likewise, if we say the latter, we can't say that what makes one state of affairs better than another is that there are sufficient reasons to prefer it, because then we're back in the water when it comes to explaining why such a state of affairs really would be better. Moreover, if we take the second line, we now have a puzzle about why Ronnie's reason to go to the party is any better than his reason to make sure that Bradley stays away. He can bring about just as much of people doing what they like by getting Bradley to stay away, after all, and on this account, his reason to do what he likes is just that it's better when more people do what they like.

REFERENCES

Broome, J. (2004): Reasons, in Wallace, Pettit, Scheffler and Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* Oxford University Press, Oxford.

Davidson, D. (1980, 1963): 'Actions, Reasons, and Causes', Reprinted in his *Essays on Actions and Events* (pp. 3−20), Oxford University Press, Oxford.

Kagan, S. (1989): *The Limits of Morality*, Oxford University Press, Oxford.

Korsgaard, C. (1996): *The Sources of Normativity*, Cambridge University Press, Cambridge.

Korsgaard, C. (1997): The Normativity of Instrumental Reason, in Cullity and Gaut (eds.), *Ethics and Practical Reason* (pp. 251−244), Oxford University Press, Oxford.

Nagel, T. (1970): *The Possibility of Altruism*, Princeton University Press, Princeton.

Scanlon, T.M. (1998): *What We Owe to Each Other*, Harvard University Press, Cambridge.

Scheffler, S. (1982): *The Rejection of Consequentialism*, Oxford University Press, Oxford.

Schroeder, M. (2004): 'The Scope of Instrumental Reason', *Philosophical Perspectives* 18 (Ethics), 337−364.

Schroeder, M. (forthcoming a): 'The Humean Theory of Reasons', in Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Vol. 2.

Schroeder, M. (forthcoming b): *Slaves of the Passions*, forthcoming from Oxford University Press.

Smith, M. (2003): Neutral and Relative Value After Moore, *Ethics* 113(3), 576−598.

University of Maryland College Park
Department of Philosophy
Skinner Building
College Park, MD 20742 USA
E-mail: mschroed@umd.edu

Present Address
University of Southern California
School of Philosophy
Mudd Hall
3709 Trousdale Pkwy
Los Angeles, CA 90089 USA
E-mail: maschroe@usc.edu