

THOMAS REID ON FREE AGENCY

Timothy O'Connor

Journal of the History of Philosophy, 32 (1994), 605-622

I

Contemporary philosophers have given a fair amount of attention to the epistemological writings of Thomas Reid, owing to the recent shift towards externalist theories of knowledge. As yet, though, Reid's views on the other main topic of concern to him - the nature of human free agency - have provoked far less explicit discussion. This is unfortunate, since Reid's discussions of various aspects of this problem are acute and on the right track towards its proper resolution. Reid's invocation of a concept of agent causation as essential to a satisfactory account of free and responsible action is by no means original to him; it is implicit in the thought of medieval philosophers such as Scotus and (perhaps) Aquinas, and (on some readings) it goes all the way back to Aristotle.¹ Reid's defense of this approach, however, has a special prominence in the history of philosophy in virtue of his explicit discussion of the notion of agent causation, in which he sharply distinguishes it from event causation.

In this paper, I offer an interpretation of the basic features of Reid's theory, taking account of some helpful discussions by a few recent commentators. Clearly, the most penetrating of these is the book-length treatment in Rowe (1991). Though I disagree with Rowe's account on some important points, my discussion owes much to his. I also defend Reid's position against a couple of basic objections to the agency theory, thereby setting the stage for a fuller defense of the viability of Reid's general approach which I have undertaken elsewhere.

II

"By the liberty of a moral agent," Reid writes, "I understand, a power over the determinations of his own will." He proceeds to amplify this claim, which opens the fourth essay of his Essays on the Active Powers of the Human Mind, with the following remarks:

If, in any action, he had power to will what he did, or not to will it, in that action he is free. But if, in every voluntary action, the determination of his will be the necessary consequence of something involuntary in the state of his mind, or of something in his external circumstances, he is not free; he has not what I call the liberty of a moral agent, but is subject to necessity. (p.599)²

Like most eighteenth-century philosophers, Reid works within the general framework of a volitional theory of agency, according to which each action is initiated by (or, in some cases, essentially consists of) a volitional event.³ Reid appears to conceive the nature of volitions generally in much the way contemporary philosophers characterize intentions.⁴ Those which causally initiate behavior are a particular type of volition, which we may think of as the agent's coming to have an intention to act immediately in a certain way.⁵ (The precise accuracy of such identifications will not bear on the following discussion in any significant way.)

Reid makes it clear in numerous places that "power over the determination of one's will" is not to be understood merely negatively, as the absence of a (prior) sufficient causal condition for the volition. For example, consider the following:

I consider the determination of the will as an effect. This effect must have a cause which had power to produce it... If the person was the cause of that determination of his own will, he was free in that action, and it is justly imputed to him, whether it be good or bad. (p.602)

I grant, then, that an effect uncaused is a contradiction, and that an event uncaused is an absurdity. The question that remains is whether a volition, undetermined by motives, is an event uncaused. This I deny. The cause of the volition is the man that willed it.⁶

Reid takes it to be part of our commonsense view of ourselves that we are often the immediate causes of our own volitions, in that we are capable of exerting power to determine how we shall act. Indeed, he claims that the notion of an agent directly bringing about some event (a volition) is the original and primary sense of the term "cause": "In the strict and proper sense, I take an efficient cause to be a being who had power to produce the effect, and exerted that power for that purpose." (p.65) This implies a thesis repugnant to commonsense (and in that respect very uncharacteristic of Reid), viz. that the patterns of regularity that we observe between events (apparently) not under the direct control of any intelligent agent are not, properly speaking, indicative of genuinely causal relations.

In order to avoid being understood as advancing a wildly implausible claim, he acknowledges at once

another meaning of the word cause, which is so well authorized by custom, that we cannot always avoid using it, and I think we may call it the physical sense; as when we say that heat is the cause that turns water into vapor, and cold the cause that freezes it into ice. A cause, in this sense, means only something which, by the laws of nature, the effect always follows. I think natural philosophers, when they pretend to shew the causes of natural phenomena, always use the word in this last sense; and the vulgar in common discourse very often do the same. (p.67)

However, this reply does not dispel the impression that Reid is here parting ways with a commonsense view of physical phenomena. For, as even Hume recognized, it is part of our ordinary conception of "physical causality" (Reid's term for event causation) that there is a "necessary connection" between a cause and its effect, that it is not a wholly arbitrary matter that events involving objects of a certain sort give rise to the effects they do. But, given his severe empiricist scruples, Hume of course thought that we really have no coherent notion corresponding to such terms as "causal power", "necessary connection," etc. Reid's rejection of th

the ordinary, fuller conception of physical causality is not similarly motivated. Rather, he was convinced that the conception of causal power is first grasped in relation to our own experience of activity; in particular, to the control we exercise over our own wills. Our "lax and popular" application of such notions to phenomena external to us, he supposed, was a mere relic of primitive, animistic beliefs.

But, pace Reid, our practice of ascribing causal powers to inanimate objects is no mere manner of speaking enshrined by custom. As Madden (1983) notes, in criticizing Reid on just this point:

We say that the atmosphere has the power to crush a tin can that has no air inside, the sea the power to crush the submarine that goes too deep, a stick of dynamite the power to explode when detonated, electric current the power to heat a resistance coil when wired in, and so on. We correspondingly say that the atmosphere, sea, stick of dynamite, and the electric current are powerful particulars, or agents, while the can, submarine, and resistant coil are passive particulars, or patients. The former make something happen upon certain releasing conditions, whereas something happens to the latter; the former does something whereas something is done to the latter. (p.329)

Madden goes on to note what may be part of Reid's motivation for supposing this distinction to be verbal only:

However, while such powerful particulars cause something to happen, they do not initiate events; they require a releasing condition such as pumping air out of the can, the submarine's diving too deep, and wiring in a resistance coil. The releasing conditions are part of the cause and hence the powerful particulars are caused to act; they do not initiate an action. On Reid's view, hence, they would not be agents. (p.329)

On Reid's behalf, it should be acknowledged that most of us do suppose there to be an important distinction between the sort of causal powers that responsible adult persons exhibit in their ordinary behavior and that of inanimate natural objects such as a detonated stick of dynamite. The philosophically unreflective person will make this distinction imprecisely, noting that it seems to be true of human behavior, but not of the operation of other kinds of "powerful particulars", that how we act is to some extent up to us, and subject to the influence of reasons. Reid maintains - correctly, in my view - that this general understanding of human agency implies the falsity of the thesis of universal causal determinism and also implies that such actions must involve a unique form of causality.

But that persons sometimes function as causes of certain phenomena in a way unique to their kind is not a reason to suppose that other sorts of objects do not function as causes in any manner whatever. Reid, however, generally professed to be unable to form a conception of causal activity that was itself necessitated to occur:

Were it not that the terms cause and agent have lost their proper meaning, in the crowd of meanings that have been given them, we should immediately perceive a contradiction in the terms necessary cause, and necessary agent. (p.607)

I am not able to form any distinct conception of active power but such as I find in myself... But, if there is anything in an unthinking inanimate being that can be called a active power, I know not what it is, and cannot reason about it. (Letter to Lord Kames, p.59)

Reid did not conclude from these reflections that it was an arbitrary, unexplained fact that nature does operate in a lawful manner. Rather, he deemed self-evident the principle that "of every event there must be a cause, that had power sufficient to produce it, and that exerted that power for the purpose" (p.625). From this principle, together with his basically Humean

n understanding of the regularity of natural phenomena⁷, he concluded that all phenomena that are not the immediate causal consequence of the exercise of active power by human free agents⁸ are most reasonably ascribed to the direct activity of a divine agent.

Surprisingly, though, in a very late letter to James Gregory (written after publication of the Essays on the Active Powers), Reid is less certain about his official view that the possession of genuine causal powers by inanimate objects is unintelligible:

You deny that of every change there must be an efficient cause, in my sense - that is, an intelligent agent, who by his power and will effected the change. But I think you grant that, when the change is not effected by such an agent, it must have a physical cause - that is, it must be the necessary consequence of the nature and previous state of things unintelligent and inactive.

I admit that, for anything I know to the contrary, there may be such a nature and state of things which have no proper activity, as that certain events or changes must necessarily follow. I admit that, in such a case, that which is antecedent may be called the physical cause, and what is necessarily consequent, may be called the effect of that cause.

I likewise admit, laws of nature may be called (as they commonly are called) physical causes - in a sense indeed somewhat different from the former - because laws of nature effect nothing, but as far as they are put to execution, either by some agent, or by some physical cause; they being, however, our ne plus ultra in natural philosophy, which professes to shew us the causes of natural things, and being, both in ancient and modern times, called causes, they have by prescription acquired a right to that name.

I think also, and I believe you agree with me, that every physical cause must be the work of some agent or efficient cause. Thus, that a body put in motion continues to move till it be stopped, is an effect which, for what I know, may be owing to an in

herent property in matter; if this be so, this property of matter is the physical cause of the continuance of the motion; but the ultimate efficient cause is the Being who gave this property to matter.

If we suppose this continuance of motion to be an arbitrary appointment of the Deity, and call that appointment a law of nature and a physical cause; such a law of nature requires a Being who has not only enacted the law, but provided the means of its being executed, either by some physical cause, or by some agent acting by his order. If we agree in these things, I see not wherein we differ, but in words. (pp.73-4, emphasis added)

Reid appears to concede here that the view that ordinary physical objects possess causal powers (and are thereby causally efficacious), in virtue of the properties of the matter of which they are composed, could well be correct. While he does not go so far as to endorse this view (but instead professes ignorance on the matter), this would be a marked change from his earlier claims in the Essays and throughout the rest of his correspondence that the very notion is unintelligible.⁹

I have been discussing Reid's views on the nature of 'physical' causation for the following reasons: First, if one follows Reid in adopting an occasionalistic view¹⁰ of physical (or, to put it in contemporary jargon, 'event') causation, the now common charge that Reid's agent causation is obscure and mysterious has less persuasive force, since there is no putatively clearer variant with which it is being compared. To Reid, the notion of an inanimate object necessarily exercising a causal power is suspect, whereas we all have a well-understood conception of our own unnecessitated causal activity.¹¹ Secondly, his non-realist (i.e., Humean) conception of event causation essentially removes what otherwise would be an important problem, viz., explaining how two fundamentally different forms of causation are integrated within the natural world. But for those who do not follow Reid in his denial of genuine causal effic

acy of the event-causal type, the job of defending and (at least partly) explicating the notion of agent causation is more complicated, and so must extend beyond Reid's treatment.

III

Even within the context of Reid's broader views on the nature of causation in the natural world, however, we can go some way toward removing apparent barriers to accepting the coherence of his notion of agent causation. In this section, I consider a pair of objections, only one of which was explicitly addressed by Reid. But it may be helpful first to follow Rowe (1991) in giving a simple set of necessary and sufficient conditions (accepted by Reid) for the truth of a proposition of the form, X caused event e:

1. X is a substance that had power to bring about e.
2. X exerted its power to bring about e.
3. X had the power to refrain from bringing about e.¹²

These conditions, of course, do not even approach a reductive analysis of Reid's concept of agent causation, something which Reid takes to be impossible. (The same is true concerning the notion of event causation, on non-Humean views.) They do, however, highlight the following two features: (a) the very concept of an agent's causing an event internal to himself implies that he must have the power in those very circumstances not only to bring about that event but also to refrain from doing so. (Actually, as we've seen, this dual capacity is implicit in the very notion of active power for Reid, and so he would regard (3) as redundant.) (b) one cannot specify circumstances under which an instance of agent causation takes place without stating that X did in fact "exert" X's active power. (This is given by (2).) This may seem uninformative or odd, but it is a direct consequence of the fact that (in contrast to necessitating or even probabilistic event causation¹³) the power to bring about some event is quite o

ften unexercised, even though all the conditions necessary for its exercise obtain.¹⁴ (I will discuss the notion of "exertion of causal power" a bit further on.)

Furthermore, we should note that the satisfaction of these conditions is possible only for beings having attributes of a more basic sort:

I should have noticed that I am not able to form a conception how power, in the strict sense, can be exerted without will; nor can there be will without some degree of understanding. Therefore, nothing can be an efficient cause, in the proper sense, but an intelligent being. (p.65)

To be capable of both producing or not producing some event in a given set of circumstances, one must be capable of forming an appetitive attitude towards either of these options, and doing so seems to presuppose some ability, even if quite limited, to conceptualize that alternative. To the conditions of having will and a degree of understanding, Reid (in the Essays) adds the condition of "some degree of practical judgment or reason":

We may perhaps be able to conceive a being endowed with power over the determinations of his will, without any light in his mind to direct that power to some end. But such power would be given in vain. No exercise of it could be either blamed or approved. As nature gives no power in vain, I see no ground to ascribe a power over the determinations of the will to any being who has no judgment to apply it to the direction of his conduct, no discernment of what he ought or ought not to do. (p.600)

Finally, Reid notes that it is also necessary that one believe that the production of the event is within one's power, and depends upon one's will. How could I form an intention to slam-dunk a basketball into a ten-foot high net, for example, when I am firmly convinced that I

am unable to do so? I might intend to try, but this is obviously different from intending to do so.¹⁵

Having noted these basic features of Reid's view, I now turn to two objections to his account, and the notion of agent causation it presupposes. The first, which Reid directly addressed (and which, one gathers, he was on various occasions confronted with), purports to show that Reid is implicitly committed to an infinite regress of volitions or decisions corresponding to every free action. Reid represents the argument thus:

'Liberty,' they say, 'consists only in a power to act as we will; and it is impossible to conceive in any being a greater liberty than this.... To say that we have power to will such an action, is to say, that we may will it, if we will. This supposes the will to be determined by a prior will; and, for the same reason, that will must be determined by a will prior to it, and so on in an infinite series of wills, which is absurd.' (p.501)

Reid correctly dismisses this objection as resting on a misunderstanding of his conception of free agency. (And it really is astonishing that objections of this sort are so often repeated, in view of the fact that it is hard to find a libertarian theorist who accepts the proposed analysis of power over one's will.¹⁶) We needn't have performed another, prior act of will in order to have determined the action-initiating volition. We simply exert active power (a conception of which we form through its effects) in so determining it - i.e., we determine the will directly. The exertion of active power is not itself a type of volition.

Reid imagines that his objector will continue to press him by noting that it is a common place that 'nothing is in our power but what depends upon the will,' "and therefore the will itself cannot be in our power.":

I answer, that this is a fallacy arising from taking a common saying in a sense which it never was intended to convey, and in a sense contrary to what it necessarily implies.

In common life, when men speak of what is, or is not, in a man's power, they attend only to the external and visible effects, which only can be perceived, and which only can affect them. Of these, it is true, that nothing is in a man's power, but what depends upon his will, and this is all that is meant by this common saying.

But this is so far from excluding his will from being in his power, that it necessarily implies it. For to say that what depends upon the will is in a man's power, but the will is not in his power, is to say that the end is in his power, but the means necessary to that end are not in his power, which is a contradiction.

In many propositions which we express universally, there is an exception necessarily implied, and therefore always understood. Thus when we say that all things depend upon God, God himself is necessarily excepted. In like manner, when we say, that all that is in our power depends upon the will, the will itself is necessarily excepted; for if the will be not, nothing else can be in our power. (p.602)

The plausibility of this reply depends upon the cogency of standard incompatibilist arguments that attempt to show that one's having it in one's power either to do or to refrain from some action implies that one's action is not the deterministic outcome of some prior state of affairs which was not in one's power. Hence, if my action is a causally determined consequence of my volition or choice so to act, and that volition or choice was not itself within my power to determine, then neither was (the determination of) the action within my power.

I will not enter into a discussion of such arguments here.¹⁷ If a demonstration of the incompatibilist's thesis can be given, as I believe it can, then it seems justifiable to conclude with Reid that we do presuppose in everyday life that the determinations of our will are within our power. Otherwise, we would be involved in a contradiction fairly close to the surface of

our conceptual scheme, which seems at best unlikely. I am not asserting dogmatically that there simply cannot be incoherencies lurking within the basic metaphysical framework suggested by the dicta of common sense. For all I've just said, it might be the case that any attempt to make fully explicit the metaphysical commitments underlying prereflective thought concerning personhood, responsibility, and so on is bound to fail. But it ought not to be so blatant as to involve our making general assertions of fundamental importance which can be seen to be contradictory on the basis of a bit of simple conceptual analysis.

So I will here suppose that Reid is right in maintaining that we do take "the determinations of our wills" to be ordinarily within our power. But a second difficulty of a rather similar sort might seem to arise, once we turn to consider what is involved in an agent's directly determining or causing his own volition, on Reid's account. At the beginning of this section, I noted (following Rowe) three necessary and sufficient "conditions" for the truth of a statement of the form "X caused event e." They lay down that the agent, X, must have had the power to refrain from as well as to perform e. Furthermore, the agent must have exerted that power to bring about e. ("In order to the production of any effect, there must be in the cause, not only power, but the exertion of that power; for power that is not exerted produces no effect" (p.603).) Rowe (1987) argues that this last condition, together with Reid's repeated assertion that every event must have an efficient cause, gives rise to another sort of infinite regress of mental acts:

Now an exertion of power is itself an event. As such, it too must have a cause. On Reid's view the cause must again be the agent herself. But to have caused this exertion the agent must have had the power to bring it about and must have exerted that power. Each exertion of power is itself an event which the agent can cause only by having the power to cause it and by exerting that power.... The result of this principle, however, is that in order to produce any act of will whatever, the agent must cause an infinite number of exertions. (pp.53-4)

In this article, Rowe proposes to handle the seeming difficulty by simply excising the requirement that there be an exertion of power corresponding to each volition:

Acts of will that are produced by the agent whose acts they are, we shall say, are such that the agent causes them but not by any other act or any exertion of the power she has to produce the acts of will (p.54).

But, besides flying in the face of repeated statements by Reid, this proposal also commits one to saying that an agent had the power to produce her volition, and in fact did so, though it was not true in any sense that she exerted that power.

In his later and fuller treatment of Reid's theory (1991), Rowe instead takes a second look at what might be involved in the "exertion" of active power for Reid. Our interpretation of this notion, Rowe suggests, will determine which of three possible responses to the objection we should adopt. The possibilities are as follows¹⁸:

- (1) Accept and defend the coherence of an infinite regress of exertions of power whenever the agent causes an act of will.
- (2) Abandon the causal principle, and hold that an exertion of power to produce an act of will is an event having no efficient cause.
- (3) Maintain that an exertion of power is not itself an event, and so does not fall under the causal principle.

The first of these is similar to a position defended by a twentieth-century agency theorist, Roderick Chisholm (though he develops this line in terms of an ontology of fact-like states of affairs, not event particulars). Since I am no more inclined than Rowe is to interpret Reid in this way, I will not consider this alternative here.

Rowe believes it is difficult to adjudicate between (2) and (3) on textual grounds, but opts for (2) since he believes it to be more philosophically defensible in the final analysis. Indeed, the resulting account is a position that Rowe himself wishes to maintain. So I will begin by considering this approach.

On the present interpretation of Reid's theory, then, an exertion of power to produce an act of will is a type of event, but one which has no efficient cause (i.e., it isn't caused by a prior event or by the agent himself). However, Rowe argues, the abandonment of the causal principle which Reid took to be self-evident is far less drastic than it might initially appear. For, as we have seen, my having the power to cause some event implies, according to Reid, my having the power not to cause it. ("Power to produce any effect, implies power not to produce it" (p.523).) And my causing that event just is for me to exert the power to do so. Taken together, these statements imply that "it is a conceptual impossibility within Reid's theory for God or any other efficient cause to produce in me an exertion of my active power" (Rowe, p.151). Of course, I might be caused to will to do some action by external manipulation or internal passion (in which case the action would not be free). But the conditions for the occurrence of an exertion of active power preclude there being any cause for this kind of event. (One might wonder whether I could be the cause of this event. But clearly this leads us into the regress, for it will require a prior exertion of my active power to bring it about, and so on ad in finitum. This, one might plausibly claim, is simply an impossibility.)

Rowe urges that this result, the conceptual impossibility of there being a cause of an exertion of active power, has the following important implication: "Reid must be viewed as a special sort of indeterminist, for we must understand him as holding that every event that logically can have a cause does have a cause" (p.155). So we can continue to attribute to him a modified version of the causal principle. It is even more significant philosophically that this distinguishes Reid's position from that of simple indeterminism. While a volition is the sort of event that can be caused by another agent or an event, that is not true regarding an exertion of active power. Rowe takes it to be an advantage of Reid's theory that it

has a built-in answer to the question: Why is the event of the agent's exertion of power uncaused? Simple indeterminism has no answer to the corresponding question of why the act of will is uncaused. (p.155)

Though Rowe's suggestion is in some ways fascinating, it doesn't hold up under close scrutiny. For we are not given an account of how we are to think of such events, these "exertions of active power," beyond the suggestions that they are some sort of change occurring within the agent (p.154), that "by which" a person agent-causes her act of will (p.156-7), and by which she does so "directly" (p.157). What are their intrinsic characteristics by which we may evaluate Rowe's claim that such an event possesses the further, desirable features of being

- "not the sort of event that can be caused" (p.155), being
- "such that if it occurs it is logically necessary that [the corresponding volition] occurs" (p.159), and being
- "intrinsically more truly an agent's own action than is an act of will" (p.155)?

Concerning the first of these features, Rowe could simply stipulate that for the proper application of the label "exertion of active power," an event must be not only of a certain (as yet unspecified) intrinsic type, but also uncaused. But this would be mere verbal sleight-of-hand, as is shown by the fact that the simple indeterminist could do the same with respect to uncaused volitions or decisions. So Rowe must have in mind a sort of event whose very nature precludes the possibility of its being caused. But how could this be true of any sort of event? How could it be that not even God could cause a certain type of (causally simple) event¹⁹ to occur, a sort which regularly does in fact occur? It is similarly unclear how such an eve

nt type could bear a logical connection to another sort (volitions), given Rowe's own view that causal necessity is not a species of logical necessity.

Clearly, Rowe must be thinking of exertions as a type of mental event. For if exertions on different occasions bring about volitions of widely varying contents, and are such as to make me responsible for my willing the way I do, then it would seem that they must themselves have an intentional content, of the general form "that I bring about a volition of type X by this very exertion of my power."

A natural reply at this point is that we are unaware of any such intentional attitudes immediately prior to our decisions. Nonetheless, we might be able to persuade ourselves that this is the case if we could thereby solve the problem of how free agency is possible. We may, therefore, focus our challenge to Rowe in this way:

You grant that a volition is not intrinsically responsibility-conferring. That is, an agent could not properly be held responsible for a volition (or a larger action to which it gave rise) that occurred uncaused. But you seem to be committed to supposing that an exertion of active power is intrinsically responsibility-conferring (and say as much in claiming that "any such event has the property that its occurrence is up to the agent" [p.225,n.18]). Simply put, why should one suppose this to be so?

Rowe replies to this (in effect) with the following argument (pp.168-9)²⁰, which he contends shows that it "follows from the very conception of active power" that an exertion of it is up to the agent whose exertion it is (and, hence, something for which he is prima facie responsible). Let us assume that I have active power (on Rowe's construal) with respect to my deciding to steal Mark's basketball. Then:

- (1) It is in my power to cause (and in my power not to cause) my decision to steal the ball. (from the assumption)

- ∴ (2) It is up to me whether I cause my decision to steal the ball.
- (3) I cause that decision if, and only if, I exercise my active power to so decide.
- ∴ (4) It is up to me whether I exercise my active power to decide to steal the ball.

The crucial inference from (1) to (2) is, I believe, unsupported. To see this, we need only reflect on what is meant by (1) on Rowe's construal. If we simply read (1) as a claim couched in ordinary English, (2) would seem to trivially follow from it. But no more is implied by (1) than that it is an open possibility, causally, that an exertion-of-active-power event of a certain specific sort occur, as well as that it not occur, and that if (but only if) it does, a decision to steal Mark's basketball will immediately follow. But it has not been shown that if the appropriate exertion-of-active-power event does occur, I will be responsible for its occurrence. Yet this is what (2) claims (given, as (3) notes, that my causing my decision is just my exerting my active power to so decide). Rowe denies the legitimacy of a similar move made by the "simple indeterminist", who claims that it is enough (in terms of the metaphysics of action) for responsible agency that my volition or decision occurs uncaused. Why, then, does he feel entitled to draw the inference of (2) from (1) in the above argument?

As far as I can tell, the only relevant putative feature of "exertions" that ordinary volitions lack is that of being intrinsically invulnerable to being caused. But no explanation has been given of how this might be, and I am confident that it is a necessary truth that all metaphysically-possible events lack such a property.²¹ Quite apart from this contention, though, Rowe fails to show how such a property would be relevant to the matter of how the occurrence of the event could be up to me. One's mind ought to boggle at the suggestion of there being a simple mental event that is intrinsically invulnerable to being caused and such that no one and nothing is responsible for its occurrence; but the source of bafflement, it seems to me, concerns the possibility of an event simply satisfying the first of these clauses.²²

I conclude, therefore, that if we follow Rowe in taking an exertion of active power to be an uncaused event that obtains immediately prior²³ to a volition or decision, the resulting the

ory fails to account for how we can be responsible for our actions. We have seen that this interpretation of Reid was the second of three routes of escape from an objection that Reid is committed to a regress of exertions of active power corresponding to every volition. From the general perspective of the agency theory, I believe, it was clearly a fundamental mistake not to take the third route and maintain that an agent's "exertion of active power" in directly producing a volition is not to be construed as a prior event. (And, hence, the question of what caused the exertion does not arise.) The whole point of the theory is to escape the dilemma posed by the apparent inadequacy of both deterministic theories and simple indeterminism to account for responsible agency. The way out involves recognizing that agents are responsible for their free actions in virtue of an irreducible causal relation between themselves and their volitions. They do not need to do something further in order to bring about a volition; rather, they do so directly.

Rowe recognized this in his earlier article, as we saw. He supposed, however, that this required abandoning the claim that an agent exerts his power to produce a volition. This was, in a way, properly motivated insofar as one supposes that "X exerted his power" can only mean that another, prior event occurs which (event-)causes the volition. But as Rowe himself conceded, it would be remarkable that Reid should fail to see the ensuing difficulty if this is how he was thinking of the matter.²⁴ An alternative construal identifies the "exertion of active power" with (the obtaining of) a causal relation between agent and volition.²⁵

Direct confirmation of this may be found in the following. In Reid's usage, "agency" and "efficiency" are synonymous with "exertion of active power."²⁶ If this is borne in mind, it will be seen that Reid explicitly explicates "exertion of active power" as a causal relation between the agent and his volition:

If it be so that the conception of an efficient cause enters into the mind, only from the early conviction we have that we are the efficient of our own voluntary actions, which I think is most probable, the notion of efficiency will be reduced to this, that

it is a relation between the cause and the effect, similar to that which is between us and our voluntary actions. This is surely the most distinct notion, and, I think, the only notion we can form of real efficiency. (pp.524-5, emphasis added)

Equally strong (though perhaps less readily apparent) support may be found in a passage from a letter to James Gregory. Commenting on a manuscript written by the latter, Reid writes:

[Y]ou speak of our having a consciousness of independent activity. I think this cannot be said with strict propriety. It is only the operations of our own mind that we are conscious of. Activity is not an operation of mind; it is a power to act. We are conscious of our volitions, but not of the cause of them.

I think, indeed, that we have an early and a natural conviction that we have power to will this or that... But I think this conviction is not properly called consciousness. (pp. 82-3, emphasis added)

If exertions of active power, or instances of activity in the strict sense (for Reid), were intentional events prior to and productive of volitions, then we should be (often) directly conscious of them, just as we are of our volitions themselves. But Reid clearly thinks this is not the case; while we firmly believe ourselves to be "active" in his sense, we cannot be said to have direct conscious awareness of this fact in any given instance. This is a (somewhat) more plausible position to take, if such activity involves no more than a form of causal relation (the only form, according to Reid) between the agent and each of his (free) volitions.²⁷ In general, he would maintain, we perceive the relata of causal relations but not the relations themselves. (Compare the remarks near the end of his career that I noted earlier, in which he seemed to countenance the possibility of genuine causal relations between events in nature. It is clear that at Reid there thought that if there were such relations, we certainly don't perceive them.)

Finally, when we take "exertion of active power" to signify a causal relation between agent and volition, we can make sense of Reid's claims that (a) an exertion of active power can not be caused, and (b) there is a conceptual connection between an exertion of active power and its corresponding effect (a volition).²⁸ As regards (a), a causal relation is not the sort of entity that one can (directly and independently) cause to obtain. (Suppose it were. Then consider an instance of "ordinary" event causation, A's causing B, where A and B are events. If there is a cause of the relation's holding between A and B, presumably it is A itself.²⁹ But if A causes this relation ("r") between A and B, then it follows that there is a further causal relation ("r*") between A and r. But if there is a cause of r, then surely there is one of r* as well. And so we should be led to the absurdity of positing an infinite number of existent relations corresponding to each causal relation that holds between events in the world. If, then, we reject the demand for a cause of the relation holding between two events as unwarranted, we may likewise conclude that no cause is to be sought for agent-causal relations.) And as for (b), the obtaining of a relation between two relata presupposes that each of the relata exists or obtains. (This is true whether the relation in question be event- or agent-causal.)

So, what seemed inexplicable on the event construal of "exertions of active power" can be seen to follow immediately, once we understand Reid to be referring to an (irreducible) type of causal relation. My interpretation not only has strong textual support, but it also provides the broad outlines of a coherent picture of the metaphysics of free agency. Obviously, a number of important objections have not been considered here. One such question that has likely occurred to the reader is the following: "Even if I grant that the notion of agent A's bearing an irreducible causal relation to his volition, V, is intelligible, what are we to say of the event A's causing V? Does this event have a cause? If so, doesn't the regress continue to loom? And if not, given your own (and Reid's) rejection of simple indeterminism, how can you maintain that the agent is nonetheless responsible for it?" Consideration of this and other contemporary objections to the agency theory would take us beyond the texts of Reid. Accordingly, I've reserved this task for another occasion.³⁰

Indiana University

Bibliography

- Donagan, Alan (1979). "Chisholm's Theory of Agency", in E. Sosa, ed., Essays on the Philosophy of Roderick Chisholm. Amsterdam: Rodopi.
- (1987). Choice: The Essential Element in Human Action. London: Routledge, Kegan, & Paul.
- Dretske, Fred (1988). Explaining Behavior. Cambridge, MA: MIT Press.
- Gallie, R. (1989). Thomas Reid and 'The Way of Ideas'. Dordrecht: Kluwer Academic Publishers.
- Ginet, Carl (1980). "The Conditional Analysis of Freedom", in P. van Inwagen, ed., Time and Cause. Dordrecht: D.Reidel, 171-86.
- Madden, E.H. (1983). "Commonsense and Agency Theory", Review of Metaphysics 36 (Dec), 319-41.
- O'Connor, Timothy (1993). "On the Transfer of Necessity", Nous 27, 204-18.
- (1995). "Agent Causation", in Timothy O'Connor, ed., Agents, Causes, and Events: Essays on Indeterminism and Free Will. (New York, Oxford University Press), forthcoming.

- Reid, Thomas (1895). The Works of Thomas Reid, D.D., 8th ed. Ed. Sir William Hamilton. Edinburgh. [Reprint. With an introduction by H.M. Bracken. Hildesheim: Georg Olms Verlag, 1983]
- Rowe, William (1987). "Two Concepts of Freedom", Presidential Address, Proceedings and Addresses of the APA 61, 43-64.
- (1991). Thomas Reid on Freedom and Morality. Ithaca: Cornell University Press.
- Van Inwagen, Peter (1983). An Essay on Free Will. Oxford: University Press.

¹ Rowe (1987) has noted that this concept also plays a role in the less-developed theories of free agency put forward by some of Reid's eighteenth-century contemporaries, including Samuel Clarke and Edmund Law.

² Throughout this essay, references to Reid's writings will be to the 1967 edition of his works. Apart from statements drawn from his private correspondence, the quotations are taken from Reid's Essays on the Active Powers.

³ For an example of a passage in which Reid clearly commits himself to a volitional account of this sort, see note 8 below.

⁴ Evidence for this includes the fact that Reid recognizes a purpose whose content concerns a time beyond that of the immediate future as "strictly and properly a determination of the will, no less than a determination to do it instantly" (p.539). (See

also p.542, where he also includes as a type of volition life-long resolutions or purposes as to a general course of conduct.)

⁵ I am thereby assimilating Reid's action-starting type of "volition" to Donagan's (1987) "determinate intention" or "choice".

⁶ From a letter to Dr. James Gregory, in Reid (1967), p.88. (The next two quotations from Reid that follow are also taken from letters to Dr. Gregory.)

⁷ Passages that strongly indicate that Reid followed Hume concerning the nature of "physical" causality include the following: "We see an established order in the succession of natural events, but we see not the bond that connects them together" (p.522). "What D. Hume says of causes, in general, is very just when applied to physical causes, that a constant conjunction with the effect is essential to such causes, and implied in the very conception of them" (Letter to Dr. Gregory, p.67). "I can see nothing in a physical cause but a constant conjunction with the effect. Mr. Crombie calls it a necessary connection; but this no man sees in physical causes..." (Letter to Dr. Gregory, p.88).

⁸ The only events in nature that clearly fall into this category are free volitions: For "[w]e perceive not any necessary connection between the volition and exertion on our part, and the motion of our body that follows them....That there is an established harmony between our willing certain motions of our bodies, and the operation of the nerves and muscles which produce those motions, is a fact known by experience. This volition is an act of the mind. But whether this act of the mind have any physical effect upon the

nerves and muscles, or whether it be only an occasion of their being acted upon by some other efficient, according to the established laws of nature, is hid from us." (p.528)

⁹ There are two phrases in the above passage, however, that might give one pause about my general reading of it. First, Reid speaks of the divine Being as giving "this [efficacious] property to matter." This might suggest that an object's causal powers could be "added onto" it, so to speak, so that it would be a matter for divine choice whether this very object would be capable of producing a particular effect in certain circumstances or not, or even, perhaps, which such effect it might produce. On such a view, it would not be part of the essential nature of an object of a certain structure and composition that it have the powers it has, but it would be instead a completely contingent fact about it. I confess to having some difficulty in understanding this sort of view, which posits a real (non-epistemic) form of necessity in the causal relations between events, yet denies that it is simply a species of metaphysical (or "broadly logical") necessity. (Admittedly, such a view is rather popular among contemporary philosophers.) An alternative is to interpret Reid's insistence that "the ultimate efficient cause is the Being who gave this property to matter" as simply highlighting the evident fact that it is not up to such an inanimate object that it is found in certain circumstances and necessarily produces the effect it does, but is, on Reid's view, a product of the divine creation and ordering of the physical world. That the object produced such an effect in those circumstances owes its explanation, ultimately, to the

divine creative choice. We may similarly read the other questionable phrase - "If we suppose this continuance of motion to be an arbitrary appointment of the Deity" - as stressing the fact that it is not a necessary feature of the world that it be governed by the causal laws operative in it, since it might have been constituted by very different sorts of elementary particles, having different fundamental properties. God was presumably free to choose which general pattern of physical phenomena would be realized in the world, and would simply create matter appropriate to this end ("the means of its being executed").

¹⁰ The view that is occasionalistic is the one given in Reid's published writings. The suggestion that he entertains in his late correspondence with Dr. Gregory (viz., that God imbues matter with causal properties) would be a departure from occasionalism, but as we've seen, Reid does not go so far as to accept this view.

¹¹ As Donagan (1979) comments, contemporary philosophers who allege that Reid's agent-causality is an obscure, illegitimate attempt to extend the notion of causation beyond the well-understood parameters of event-causality

uncannily reverse the objections Reid made in the eighteenth century against event-causation. Reid pointed out that the necessity possessed by the laws of Newtonian physics was so mysterious that Newton himself indignantly denied that he thought it causal, whereas for centuries the model of causality had been that of the relation of an agent to his actions, and

above all, that of the divine agent to his creative actions....In terms of this, the original notion of cause, Reid complained that the post-Newtonian concept of event-causation was improper and obscure. (p.218)

¹² Rowe (1991), p.49. Compare Reid, p.65: " In the strict and proper sense, I take an efficient cause [of some event] to be a being who had power to produce the effect, and exerted that power for that purpose...Power to produce an effect, supposes power not to produce it; otherwise it is not power but necessity, which is incompatible with power taken in a strict sense."

¹³ In an instance of probabilistic event causality, the power to bring about one of a range of effects will always be exercised, though which effect is brought about is undetermined.

¹⁴ "It cannot be said that there is a constant conjunction between a proper cause and the effect; for, though the effect cannot be, without power to produce it, yet that power may be, without being exerted, and power which is not exerted produces no effect" (p.66). As we will see, on Rowe's interpretation of what is meant by condition (2), it turns out to be a more informative condition than I've suggested. As I will try to show, though, it raises insoluble problems of its own and is, I think, a misunderstanding of Reid and the agency theory in general.

¹⁵ And we can form an intention ("will") to perform some action that is not actually in our power, though we believe that it is. Reid gives the example of a man, struck with palsy in his sleep that renders him incapable of speech, who attempts to speak in

ignorance of what has happened. As soon as he realizes his incapacity, though, "he ceases to make the effort" (p.532).

¹⁶ The inclination to suppose that by power over any event, including volitions, one can only mean "a power to act as we will" is so strong for some that even a present-day commentator on Reid raises the very objection given in the text, despite Reid's explicit repudiation of it as resting upon a confusion! Gallie (1989) writes: "...if I am the efficient cause of the determinations of my will then the determinations of my will depend upon my will. But then if whatever depends upon my will comes into being through a determination of my will then no determination of my will can come into being, it would seem, unless by way of another determination of my will." (p.138)

¹⁷ I refer the reader to van Inwagen (1983) and Ginet (1980) for carefully laid out modal versions of the argument. For a discussion of the critical responses their work has provoked, see my "On the Transfer of Necessity".

¹⁸ I have adapted these from Rowe's discussion on pp.148-49.

¹⁹ Where the type is characterized wholly in terms of intrinsic features. Obviously (if incompatibilism is correct), God could not cause an instance of the type "freely choosing to \emptyset ", but this is because the type has stipulations regarding its manner of production built into it. Also, I add the qualification "causally simple" because it may well be that events having a certain sort of internal causal complexity could not have sufficient causal

conditions. This would be true, I believe, of an event of the form (agent) S causes e, an event which is not wholly distinct from and prior to the event e that is brought about, but which is, rather, the bringing-about-of-e (which includes e as a component). I will not go into this matter here, as I believe that this is clearly not what Rowe has in mind.

²⁰ I have changed the example.

²¹ Again, this claim may not hold without exception for events of a causally complex nature.

²² Before introducing the argument in the text above, Rowe suggests (without further elaboration) that the following two considerations together make it "at least doubtful that my exercise of active power need be capricious or fortuitous" (p.166):

- a) "[b]y virtue of its very nature it cannot have a cause."
- b) "it is only by virtue of it that I as agent can cause and be responsible for any action or change."

I have already remarked that, without further explanation of the intrinsic character of such an event, the consideration posed by (a) seems irrelevant. Regarding (b), it seems that whether we are responsible for our own actions (including internal changes, such as volitions) depends on whether we are responsible for the "exertions" that are said to give rise to such actions. (I have no doubt that Rowe as an incompatibilist agrees about this.) To note, in effect, that it seems doubtful that we are not responsible for our

exertions of active power, given our responsibility for the actions to which they give rise, can only be question-begging in a context where the possibility of responsibility for such "exertions" is being questioned.

²³ Although I believe Rowe is thinking of an exertion of active power as occurring just prior to the corresponding volition, I know of no passage in which he explicitly says so.

²⁴ Consider the fact that he often refers to both his unrestricted causal principle and exertions of active power within a single sentence. An example is the following remark (also noted by Rowe in this connection) in his assessment of Leibniz's principle of sufficient reason (as applied to human action):

If the meaning of the question be, was there a cause of the action? Undoubtedly there was: of every event there must be a cause, that had power sufficient to produce it, and that exerted that power for the purpose. (p.625)

(Another is his comment that "every change must be caused by some exertion, or by the cessation of some exertion of power" (p.515).

²⁵ In a footnote to (1987), Rowe considers the possibility that an "exertion of active power" might not be an event after all, for Reid, though he thought it unclear as regards Reid's texts. Even here, though, he fails to identify this notion with the agent-causal relation, and so leaves it unclear what an exertion of active power

would amount to on such a reading. He takes a similar view as regards Reid's intent in (1991), though he has come to regard the basic claim of an irreducible causal relation, which underlies such a reading, as philosophically untenable. (See, for example, his remarks on pp.156-57.)

²⁶ See the June 14, 1785 letter to Dr. Gregory (p.65) and p.523 of the Active Powers.

²⁷ Nonetheless, I believe that Reid is mistaken in this claim: contra Reid, the agency theorist should maintain that we are directly aware of our own activity in causally producing intentions to act.

²⁸ "...it is a contradiction to say, that the cause has power to produce the effect, and exerts that power, and yet the effect is not produced" (p.268).

²⁹ In Explaining Behavior (Ch.2), Fred Dretske suggests that there are straightforward cases of an event A's causing B having a cause. A simple example is my pressing your doorbell (A) resulting in the ringing of the bell (B). Had the button not been wired to the bell in the appropriate way, (B) would not have obtained. The electrician who installed the device, it seems, was in some sense causally responsible for A's causing B, even though he did not in any way initiate this event. In Dretske's terminology, the electrician's activity was the structuring cause of A's causing B (by setting up the circumstances in such a way that A, should it occur, would result in B), though not its triggering cause (the cause of A, my pushing the button).

We should observe that in offering this analysis, Dretske is not taking the cause event (A) to be a set of sufficient circumstances, but rather (in accordance with ordinary usage) the most salient event alone, such that this event causes B in such and such circumstances (necessary for B's occurring). But, of course, the agency theorist is following the same practice in denominating an agent, S, as the cause of some event, e, since the agent can exercise such a causal power only given various necessary conditions. Nonetheless, whatever was responsible for these further necessary circumstances would not seem to be a structuring cause (in Dretske's sense) of S's causing e, since they do not make it necessary that e occurs, given that the agent was in some particular state just prior to the occurrence of e. (We might say instead that they were necessary structuring conditions or factors.)

³⁰ See my forthcoming "Agent Causation".