

30  
THOMAS REID

*Terence Cuneo and Randall Harp*

There is a standard understanding of Thomas Reid's views regarding the will and action according to which Reid defends both a libertarian account of free will and an agent causal theory of action. While the standard understanding of Reid's views is probably correct, puzzles emerge when one digs deeper into what Reid says when he develops these views. In this chapter, we address three interpretive challenges that face those who wish to understand what Reid says about will and action, namely:

**The Action Challenge:** Reid claims that some of our actions are free. But what behaviors belong to the category *action*? Reid seems to say different things on this issue. In some places, he endorses a broad view according to which events, such as throwing baseballs, count as actions. In other places, he embraces a narrow view according to which only certain kinds of mental events count as actions. The issue presents a challenge to Reid's interpreters because how one answers it determines whether Reid's position is subject to familiar Frankfurt-style counterexamples. (See Frankfurt [1969] and Sartorio in the present volume [Chapter 15] for a discussion of these examples.)

**The Motive Challenge:** Reid claims that actions are ordinarily taken on the basis of motives. But Reid also endorses two accounts of motives that appear incompatible—one of which seems to fit better with his account of free action, the other of which seems to mesh better with broadly common sense views about the motivational roles of mental states such as desires. The fact that Reid works with two different accounts of motives presents a challenge to Reid's interpreters because neither account of motives seems to fit well with Reid's commitments.

**The Volition Challenge:** Reid claims that the agent is an efficient cause (that is, agent cause) of her actions and that actions consist in the exertion of active power. But what is the relation between being an efficient cause (i.e., a substance) and the exertion of active power? Reid never addresses this question, but it presents a challenge to Reid's interpreters, for without an account of this relation, Reid's agent causal view is at best obscure and at worst, incoherent (if, for example, every exertion of active power requires a further exertion of active power).

We won't attempt to resolve these interpretive issues in what follows. Instead, our aims are to make these interpretive issues explicit, explain why Reid's commitments generate them, and introduce the reader to the available options for resolving them.

THOMAS REID

### The Action Challenge

Let us say that every mental or bodily change in the world produced by an agent (or some part of an agent) is the actualization of a *behavior-state*. While the actualization of a behavior-state is an event, it needn't be an action. Your digesting this morning's breakfast, for example, is the actualization of a behavior-state. In some sense, it is due to you (or some part of you), but it is not an action that you perform. In contrast, your deliberating about whether to visit the Taj Mahal is both the actualization of a behavior-state and an action that you perform, as the deliberation is due to you in a way that the digestion of your breakfast is not. But what explains the difference between these two cases? Why does the actualization of a behavior-state count as an agent's action in some cases but not in others? Reid has a very simple answer to these questions: the actualization of a behavior-state counts as your action, says Reid, just in case you are its efficient cause (EAP I.i: 13).

Reid's answer to what we might call the *action problem* presupposes a distinction between 'efficient' and 'physical' causality (see EAP I.vii: 41; COR: 178, for this terminology). Reid adopts the notion of an efficient cause from thinkers such as Samuel Clarke (1998/1705). An efficient cause is what we today would call an *agent cause*, and an agent cause is something endowed with what Reid calls *active power*. Since you are an agent, Reid holds that you are an agent cause and, hence, endowed with active power. It is important to emphasize that, while Reid holds that everything with active power must have a will, he distinguishes active power from will, writing that the "effect produced, and the will to produce it, are things different from active power, but we can have no conception of it [i.e., active power], but by its relation to them" (EAP I.v: 32). So, in Reid's view, active power is more fundamental than will, and is necessarily exerted when we will. "By the *liberty* of a moral agent, I understand, a power over the determination of his own will" (EAP IV.i: 267).

Reid adopts the notion of a physical cause from Hume (see Hume 1999/1748: sect. 7). Roughly, a physical cause is an event of some type that is connected by a law of nature to an event of some other type. When you throw a baseball at a window, for example, your throwing the baseball is a physical cause of the window's shattering, since events of the first type (i.e., throwing a baseball at a window) are connected by a law of nature to events of another type (i.e., a targeted window's shattering upon impact).

Efficient causality and physical causality appear to be very different. Agent causes are not, after all, events related to other events by natural laws; they are substances. And physical causes are not agents; they are events. Surprisingly, though, Reid maintains that they are very closely linked, holding that the only genuine causality in the world consists in the exercise of agent causality. If Reid is right, your deciding to throw a baseball at a window comes about through an exertion of your active power. You are this event's agent cause. But that the ball flies through the air and shatters the window upon impact is, Reid says, due to the exertion of God's active power, for the

physical laws of nature are the rules according to which the Deity commonly acts in his natural government of the world; and, whatever is done according to them, is not done by man, but by God, either immediately, or by instruments under his direction.

(EAP IV.xi: 251)

Reid, then, endorses a version of quasi-occasionalism—'quasi' because Reid allows that *both* human beings and God (and possibly other beings) are agent causes. Although

Reid makes no effort to define the notion of an agent cause, he does specify an array of conditions that something must satisfy to be an agent cause. Among these is that an agent cause has active power, and that active power must come with options: "Power to produce an effect," Reid says, "implies power not to produce it" (EAP I.v: 29). Let us (following Yaffe 2004: Chapter 2) call this condition:

**Power to do Otherwise:** If an agent has the power to act in a certain way, then he also has the power not to act in that way.

Power to do Otherwise affords an even clearer understanding of Reid's answer to the action problem: it tells us that the actualization of any behavior-state of which someone is an agent cause must be such that that person had the power to not actualize that behavior-state when he did. Although this is helpful, we still don't know exactly what sorts of behavior-states, according to Reid, satisfy this condition. Do mental events such as imagining the Taj Mahal as well as bodily movements such as throwing baseballs satisfy Power to do Otherwise and, hence, count as actions?

Reid says different things on this matter. In some places, he acknowledges that there is an ordinary understanding of action that is broad according to which the actualization of a wide array of behavior-states such as throwing baseballs, flagging down cabs, and scratching one's ear count as actions (see EAP III.i.i: 97, IV.i: 198). Call this the *wide understanding of action*. In other places, Reid acknowledges that there is a 'philosophical' understanding of action according to which only the actualization of those behavior-states that are the direct consequence of an exertion of active power—that is, only those behavior-states that are volitions—are actions (see EAP I.i: 13, IV.ii: 203). Call this the *narrow understanding of action*. Sometimes Reid seems to work with the wide (or 'broad') understanding; other times he appears to work with the narrow understanding. Each understanding has its pros and cons.

The virtue of the wide understanding is that it approximates our ordinary understanding of action according to which the actualization of behavior-states such as throwing baseballs and flagging down cabs are actions and, hence, behavior-states for which we can be held accountable. The wide interpretation, however, also has a downside, which is that it exposes Reid's view to counterexamples that are potentially quite damaging.

Reid would have been familiar with some of these counterexamples. Consider, for example, Locke's man-in-a-locked-room scenario, which is a precursor to so-called Frankfurt-style counterexamples. In Locke's presentation of the scenario, a man is locked in a room. Although the man decides to stay in the room, he lacks the power to leave because, unbeknownst to him, the door is locked. This case provides the materials for the following objection to Reid's view:

1. The actualization of a behavior-state B counts as an agent's action at t only if B satisfies Power to do Otherwise at t. (from Reid's answer to the action problem and Power to do Otherwise)
2. The man in the locked room's behavior of staying in the locked room doesn't satisfy Power to do Otherwise: he has the power to stay in the room but lacks the power to refrain from staying in the room. (assumption)
3. So, when the man in the locked room decides to stay in the room, his behavior of staying in the room does not count as his action. (from 1, 2)

4. But his behavior of staying in the room does count as his action. (assumption)
5. So, (1) is false: it's not the case that the actualization of a behavior-state B counts as an agent's action at t only if B satisfies Power to do Otherwise at t. (from 1–4)

Let's now add that, in Reid's view, the actualization of a behavior-state counts as an action if and only if it is a free action (EAP IV.i: 267, IV.iii: 212). Given this assumption, the argument just stated also implies:

6. The man in the locked room's behavior of staying in the locked room is not a free action.

This appears to be a problem for the wide understanding of action.

What we earlier referred to as the narrow understanding of action side-steps this objection completely, as it holds that the formation of volitions is the only type of action that the agent performs. If this is right, the correct way to describe the man-in-a-locked-room scenario is to focus on the man's formation of volitions through the exertion of his active power. When the man decides to stay, his decision results from the exertion of his active power. But when he decides this, he could decide otherwise; he could decide not to stay. Of course he would be unable to execute his decision and leave because he would discover that the door to the room in which he's staying is locked. Still, his volition to stay satisfies Power to do Otherwise. Given Reid's further claim that the only types of actions are free actions, his decision to stay is a free action.

But there is a price for accepting the narrow view of action. This account not only radically constricts the range of behavior-states that can count as actions, but also 'internalizes' action, implying that none of the behaviors that we commonly call actions, such as throwing baseballs, could be actions (compare, for example, EAP I.i and I.vii: 36 with IV.i). For a philosopher who prided himself in defending common sense, this would not be a welcome result.

Reid himself never faces the Action Challenge head-on. But he does say some things that suggest a way of blunting its force. When Reid presents the narrow understanding of action, he acknowledges that the actualization of many behavior-states would not count as actions; only the formation of a volition through the exertion of active power would count as an action. Nonetheless, Reid continues, agents can be held morally responsible for the actualization of behavior-states that are not actions in the narrow sense, such as throwing baseballs at windows. Here is what Reid says:

That there is an established harmony between our willing certain motions of our bodies, and the operation of the nerves and muscles which produces 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 [cause], according to the established laws of nature, is hid from us.

Still,

the man who knows that such an event depends upon his will, and who deliberately wills to produce it, in the strictest moral sense, the cause of the event;

and it is justly imputed to him, whatever physical causes may have concurred in its production.

(EAP I.vii: 40)

If Reid is right about this, we should distinguish an agent's being the *cause* of the actualization of a behavior-state from an agent's being the *moral cause* of an actualization of a behavior-state. An agent is the cause of the actualization of a behavior-state B when and only when she is the sole agent cause of the actualization of B (and all events that might be necessary for actualization of B, such as the operations of nerves and muscles that are necessary for the movement of the arm). When an agent is the cause of the actualization of a behavior state, then the actualization of that behavior-state is an action in the narrow sense. In contrast, an agent is a moral cause of the actualization of B when and only when she is an agent cause (though not necessarily the only agent cause) of B and can rightly be held accountable for the actualization of B. When an agent is a moral cause of the actualization of a behavior-state, then the actualization of that behavior-state is an action in the wide sense.

While distinguishing different types of action has the disadvantage of multiplying senses of 'action,' the disadvantage is probably only apparent; Reid is happy to allow that one and the same term expresses a multiplicity of meanings provided that such an allowance is justified. And, in this case, it looks as if it is. Given certain theoretical purposes in which we are attempting to identify, in a philosophical precise way, what counts as an action (and, hence, a free action), we need to operate with the narrow understanding of action. But, given certain practical purposes in which we engage in practices of holding each other accountable for what we do, we need to employ the wide understanding of action. If Reid is right, we need to employ multiple action concepts in order to serve this range of purposes.

### The Motive Challenge

In his chapter "Observations Concerning the Will," Reid writes that "in all determinations of the mind that are of any importance, there must be something in the preceding state of the mind that disposes or inclines us to that determination." The reason is that if

the mind were always in a state of perfect indifference, without any incitement, motive, or reason, to act, or not to act, to act one way rather than another, our active power, having no end to pursue, no rule to direct its exertions, would be given in vain.

(EAP II.i: 51)

A close reading of the *Active Powers* reveals, however, that just as Reid operates with two different understandings of action, he also embraces two different accounts of motives that appear incompatible. The position Reid presents in the early chapter "On the Influence of Incitements and Motives on the Will"—what we'll call the *wide account of motives*—says that motives come in two sorts: some are mental states, while others are ends in the sense of being that for the sake of which an agent acts (EAP II.ii). The view of motives presented in the later chapter "Of the Influence of Motives"—what we'll call the *narrow account of motives*—says that motives come in only one sort: those ends for the sake of which an agent acts.

Let us begin with the wide account of motives. Following Reid's terminology, call those motives that are mental states "incitements" and those that are not "ends." Reid says incitements come in two types. There are, in the first place, what Reid calls the "mechanical" principles of action. The mechanical principles of action include instincts for self-preservation, food, procreation, and the propensity to imitate (cf. EAP III.ii–iii). In addition, there are those incitements which Reid labels the 'animal' principles of action. The animal principles of action include what Reid calls the benevolent affections, such as gratitude and esteem; the malevolent affections, such as resentment and the desire to better someone else in some matter, as well as a multitude of other incitements, such as the desires for power and knowledge.

As for ends, these come in two different kinds. Call ends of the first sort *agent-reasons*. To acquaint yourself with this concept, think of some case in which you perform some action and the consideration in light of which you performed that action—that feature of your situation that, in your estimation, made the action worth performing or recommended its performance. This feature is an agent-reason. Reid says agent-reasons are the "rational principles" of action, of which there are two: first, one's "good on the whole" and, second, one's "duty" (EAP III.i). Reid calls these the rational principles of action because they "have no existence in beings not endowed with reason, and, in all their exertions, require, not only intention and will, but judgment or reason" (EAP III.iii.i: 152). Call ends of the second sort *end-states*. These are those states that an agent endeavors to actualize through the exertion of his active power. (One can think of end-states as a subset of behavior states: they are behavior-states that an agent can aim to actualize directly through the exertion of her active power). Examples of end-states would be your kicking a ball into the left-hand corner of the goal or pruning an apple tree.

Ends of both sorts play a role in action. Agents act to bring about end-states in light of agent-reasons. Somewhat more exactly, Reid's view is that, cases of indifferent action aside, an agent S acts for the sake of an end if and only if there is some end-state E such that S actualizes E through the exertion of his active power in light of some agent-reason R.

What we are calling the narrow account of motives is less capacious than the wide account. For, according to the narrow account, motives are necessarily ends. Whatever role incitements play in action, according to this view, it is not that of being a motive.

Why would Reid have offered these conflicting accounts of motives, ultimately pairing the narrow account of motives with the narrow understanding of action? Here is a hypothesis: if the wide account of motives were correct, incitements would be physical causes of why we act. Remember, though, that Reid commits himself to the claim that all causality is agent causality. It follows that were a desire a physical cause of why you act, this physical cause would be the effect of the exertion of someone else's active power—presumably, God's—on you. This poses a threat to Reid's account of action according to which when you act—and, hence, act freely—that action consists not in the exertion of some other agent's active power but in *your* active power. This last claim is captured in the thesis that (following Yaffe 2004) we'll call:

**Efficient-Causal Exclusivity:** Every event that has an agent cause has one, and only one, agent cause.

Given Efficient-Causal Exclusivity, we can state the Motive Challenge as follows:

- (1\*) If the narrow account of motives is true, then incitements have no motivational role to play in the production of action. For, if they did, actions would have too many causes.
- (2\*) If the wide account of motives is true, then the actualizations of behavior-states that Reid maintains are paradigmatic actions are not actions at all, as they would have too many causes.
- (3\*) So, if either the narrow or the wide account of motives is true, then either incitements have no motivational role to play in the production of action, or the actualizations of behavior-states that Reid maintains are paradigmatic actions are not actions at all.

Like the Action Challenge, Reid never explicitly addresses the Motive Challenge. Nonetheless, Reid might have resources within his thought that allow him to address it. In Reid's view, an agent can be more or less responsible for an action, depending to what degree that action is the product of passion (EAP I.vii: 40). Although Reid does not develop this thought, it suggests that Reid was (at times) comfortable with the idea that actions can have multiple causes. And if actions can have multiple causes, then it would be possible for the actualization of a behavior-state to count as an agent's action even though it had multiple causes. Presumably, the test for whether the actualization of a behavior-state counts as an agent's action would be whether the agent could still rightly be held accountable (to some degree) for the action. Or to state the same point using the terminology introduced in the last section: the test for whether the actualization of a behavior-state counts as an agent's action is whether the agent is a moral cause of the actualization of that behavior-state. Of course, this maneuver would require Reid to reject Efficient-Causal Exclusivity. But, if he did, Reid could endorse the wide view of motives, provided that an agent is at least a partial cause of the actions he performs.

There is a way to develop this point. Drawing from Reid's own remarks, let us distinguish two types of agency, which we can call 'moral agency' and 'non-moral agency.' Now take some creature that actualizes some behavior-state B. If that agent's actualization of B is wholly determined by incitements that are operative within that creature, then it is a non-moral agent with respect to the actualization of B. If that creature's actualization of B is not wholly determined by incitements that are operative within it, then it is a moral agent with respect to the actualization of B. According to Reid, creatures such as dogs are probably non-moral agents, since (in all likelihood) their behavior is wholly determined by their strongest incitements (EAP II.ii: 53). Non-moral agents are not moral causes of their behavior, however, and so they do not perform actions in either the narrow or the wide sense (although Reid is happy to allow that we use the term 'action' to describe the behavior of non-moral agents; see EAP III.1.1).

Persons are more complex. In some cases, people actualize behavior-states strictly from desire or instinct. When they do, their agency is comparable to that of a dog; they are non-moral agents. In other cases, however, people can actualize behavior states through the exertion of active power. When they do, they are moral agents and can perform actions in both the narrow and the wide senses.

## The Volition Challenge

To this point, we've canvassed two interpretive challenges to Reid's views on will and action—what we've called the Action Challenge and the Motive Challenge. Both challenges exhibit a similar structure: they charge that Reid must surrender either some element of our broadly common sense understanding of agency or some element of his account of action. Since Reid holds that all (moral) actions are free actions, these challenges are ones that Reid must address if he is to vindicate his account of free action.

In this last section, we address a challenge that exhibits a different structure. This challenge begins with the observation that Reid's account of (moral) action—and, hence, free action—has a quadripartite structure. Any instance of free action includes (1) an agent; (2) the exertion of that agent's active power; (3) the formation of a volition; and (4) the effects of the operation of the volition. Reid's interpreters have puzzled over the relations that these four elements of free action bear to one another. The puzzle does not concern the relation between (3) and (4); the operation of one's volition is bound to its effects by natural laws. If you will yourself to throw a baseball, for example, that willing is bound together with the ball's flying through the air by natural laws of various sorts. Rather, the puzzle concerns the relations of (1), (2), and (3). In fact, some have worried that any plausible attempt to spell out the relations of these three elements involves Reid's position in a vicious infinite regress.

To appreciate why, first consider the relation between (1) and (3). Reid claims that agents are the cause of the determination of their wills (EAP I.v). But agents determine their wills through the exertion of their active power; it is (2) that connects (1) with (3). But now we face a puzzle about (2). Suppose (2)—the exertion of active power—is an event. Suppose, also, that every event must have a cause, as Reid claims (EAP IV.ii: 202, 203). In Reid's view, the cause of an event must be an efficient cause. And to be an efficient cause of some event, something must exert its active power to bring about that event. But if so, an agent cannot be the efficient cause of the exertion of her active power, for that would require there to be a volition V that brings about that exertion of her active power. But V is an event. So, V could bring about that exertion of her agent power only if there is some further volition V' that brings about V. And so on, ad infinitum.

Reid's interpreters have suggested different solutions to this concern, which we call the *Volition Challenge*. In *Thomas Reid on Freedom and Morality*, William Rowe proposes an answer that he presents not as Reid's own but on Reid's behalf. The answer consists in denying that the exertion of active power is an event and, so, does not require a cause. Rowe's thought is that, under a charitable interpretation, Reid works with a narrow understanding of events according to which events are changes that things undergo. Since the exertion of active power is not a change that a thing undergoes, it is not something that requires a cause (see Rowe 1991: 146–61).

In *Manifest Activity*, Gideon Yaffe takes a different approach. Yaffe writes:

If exertion is understood as no different from effort, or trying, then there is no regress, and the relation between the agent and his exertion of power—in the only case, namely, that of failure, in which the exertion is anything at all—is the relation of efficient causation. If an agent succeeds in acting, then his exertion is no action; it is not something he did or, what's the same, something

he efficiently caused. However, if he fails, then his exertion is his act; he is the efficient cause of it. Since, when he fails, he successfully exerted himself, his exertion to exert, in this case, is not some further action of which he is the cause. Exertion, in the case of success, is not action, and so is not something for which an efficient cause must be identified.

(Yaffe 2004: 155–6)

Yaffe claims that in cases of unsuccessful action, as when an agent tries to throw a baseball but fails (perhaps because of the sudden onset of paralysis) there is nonetheless a successful action performed, namely, the exertion. But when there is a successful action, the exertion to perform that action is ‘absorbed’ into the action. Since the successful action in this case is an exertion, there is no need to appeal to some further exertion that explains the exertion.

Each suggestion is intriguing but not entirely satisfactory. (It may be that no response to the Volition Challenge is). Rowe’s suggestion restricts the meaning of the word ‘event’ in a way that has little intuitive or textual support. (Rowe readily admits the latter is true). And Yaffe’s proposal that exertions or tryings disappear or are absorbed into actions also receives little textual support. Moreover, the proposal is puzzling. Why would tryings disappear or be absorbed into actions when we act? (Generally speaking, when some thing meets some success condition—such as a proposition’s meeting the condition *true*—that thing does not disappear). It makes sense, then, to continue to look for other ways of addressing the Volitional Challenge.

One promising alternative, which meshes with Reid’s texts, is to distinguish the *efficient causal relation* from *being an efficient cause*. As its name indicates, the efficient causal relation is a relation that an agent bears to an exercise of will or the effect of the exercise of will. For this relation to hold, it must be that (1) there is an agent endowed with active power; (2) that power is exerted; and (3) the exertion satisfies Power to do Otherwise. In contrast, an efficient cause is a person or agent. Reid repeatedly makes this point, stating: “That which produces a change by the exertion of its power, we call the *cause*” (EAP I.i: 13). And: “I consider the determination of the will as an effect. This effect must have a cause which had power to produce it; and the cause must be either the person himself, whose will it is, or some other being” (EAP IV.i: 201).

Now consider Reid’s claim that every event or change has an efficient cause. This can be understood as stating either:

- (a) Necessarily, every event E is such that there is an efficient cause to which E bears the efficient causal relation.

or:

- (b) Necessarily, every event E has an efficient cause; there is some agent that brings about E.

If Reid’s claim that every event has a cause states (a), then the Volition Challenge looms. But if it states (b), then there is no such challenge, for (b) does not imply (a). Tellingly, Reid nowhere states that his claim ought to be read as (a), although his interpreters often read him as being committed to it (see, for example, Rowe 1991: 147; Yaffe 2004: 154).

Now consider the exertion of active power, which we can assume is an event or change. Reid’s view tells us that this event or change must have an efficient cause. If we understand Reid simply to commit himself to (b), then what he says is compatible with there being nothing that bears the agent causal relation to that event. And if that is so, no regress need get started. All that needs to be the case is for there to be an efficient cause in whom that exertion is related in the right way; the exertion must *belong to* or be *grounded in* the agent. So, our suggestion is that when Reid says that every event or change has an efficient cause, he commits himself only to the claim that that event or change must belong to or be grounded in some agent. For this to be true, there need be nothing that an agent ‘does’ to be the efficient cause of some exertion of active power. Rather, to say it again, it must simply be the case that that exertion belongs to or is grounded in that agent. To state the view we are attributing to Reid paradoxically, exertions of active power are uncaused (they do not bear the efficient causal relation to some agent) but have causes (they are grounded in some agent). Or to put the point less paradoxically, exertions of active power do not have causes but nonetheless are ‘grounded exertions.’

What relation must an agent bear to exertions of her active power for her to be their cause? We’ve suggested that these exertions must belong to or be grounded in her person. Under this interpretation of Reid’s view, the following claims are true: agents possess active power. When agents act, they do so through the exertion of active power; the exertion belongs to or is grounded in the agent. But agents do *not* exert active power in order to act; the exertion of active power is not a thing that agents (strictly speaking) do. One might be tempted to ask: But *how* does an agent act through the exertion of active power? Under the interpretation we’re exploring, Reid’s view must be that, although the question might seem intelligible, it is not.

### Acknowledgments

Thanks to Kevin Timpe for his comments on a draft of this chapter.

### Bibliography

- Clarke, S. (1998/1705) *A Demonstration of the Being and Attributes of God and Other Writings*, E. Vailati (ed.). Cambridge: Cambridge University Press.
- Frankfurt, H. (1969) “Alternate Possibilities and Moral Responsibility,” *Journal of Philosophy* 66: 829–39.
- Harris, J. (2007) *Of Liberty and Necessity*. Oxford: Oxford University Press.
- Hoffman, P. (2006) “Thomas Reid’s Notion of Exertion,” *Journal of the History of Philosophy* 44: 431–47.
- Hume, D. (1999/1748) *An Enquiry Concerning Human Understanding*, T.L. Beauchamp (ed.). Oxford: Oxford University Press.
- Reid, T. (2002) *The Correspondence of Thomas Reid*, P. Wood (ed.). Edinburgh: Edinburgh University Press (abbreviated as COR).
- Reid, T. (2010/1788) *Essays on the Active Powers of Man*, K. Haakonssen and J. Harris (eds). Edinburgh: Edinburgh University Press (abbreviated as EAP).
- Rowe, W. (1991) *Thomas Reid on Freedom and Morality*. Ithaca: Cornell University Press.
- Rowe, W. (2004) “Thomas Reid’s Theory of Freedom and Responsibility,” in T. Cuneo and R. van Woudenberg (eds), *The Cambridge Companion to Thomas Reid*. Cambridge: Cambridge University Press, pp. 222–42.
- Yaffe, G. (2004) *Manifest Activity*. Oxford: Oxford University Press.

### Further Reading

- Cuneo, T. (2011) “A Puzzle Regarding Reid’s Theory of Motives,” *British Journal for the History of Philosophy* 19: 963–82. (This paper explores Reid’s answer to what we refer to as the Motive Challenge.)

- Harris, J. (2007) *Of Liberty and Necessity*. Oxford: Oxford University Press. (This book is an instructive historical survey of how early moderns thought of the freedom of the will, which presents a variant of what we call above the 'standard interpretation'.)
- Rowe, W. (1991) *Thomas Reid on Freedom and Morality*. Ithaca: Cornell University Press. (This is the first book-length discussion and defense of Reid's position, which also defends a variant of the standard interpretation.)
- Van Cleve, J. (2015) *Problems from Reid*. Oxford: Oxford University Press. (Chapters 14 and 15 are a highly instructive discussion of Reid's views regarding will and action.)
- Yaffe, G. (2004) *Manifest Activity*. Oxford: Oxford University Press. (This book provides a historically sensitive and sophisticated discussion of Reid's views regarding the will and action, also defending the standard interpretation.)

### Related Topics

Agent Causation  
 Non-Causal Libertarianism  
 Frankfurt-Style Examples  
 Gottfried Leibniz  
 David Hume

## 31

# IMMANUEL KANT

*Benjamin Vilhauer*

### Introduction

Kant's most distinctive ideas about free will are the ones he develops with his "critical turn" towards the doctrine he calls "transcendental idealism," which begins with his *Critique of Pure Reason* in 1781. Kant's most striking claim is that it is possible for human beings to have an incompatibilistic sort of free will even if determinism is true. Scholars have been intrigued and confounded by this claim ever since, and there is still little consensus on how it should be understood. Isn't this claim a straightforward contradiction? Was Kant really offering a compatibilist view in a confused way? The controversy is in part due to the fact that his views on free will are tied in tightly with the theory of transcendental idealism and its distinction between phenomena (things as they appear) and noumena (things as they are in themselves), which has divided scholars from the moment Kant first advanced it. I will begin with a brief explanation of transcendental idealism and its connection to free will, and go on to provide a brief overview of three important interpretations of Kant's free will theory, which I will call compatibilism, deflationism, and libertarianism. I will go on to discuss libertarianism in more detail, because it is the most textually accurate interpretation.

### Transcendental Idealism

The core ideas in Kant's doctrine of transcendental idealism include (i) the "transcendental distinction" between *noumena*, or things as they are in themselves, and *phenomena*, or appearances, that is, things as they appear; (ii) the claim that phenomena are as they are not just because of the way noumena are, but because our own mental activity is involved in making them as they are; (iii) the claim that the mental activity involved in making phenomena as they are is necessary for us to have any experience of objects at all, and that we therefore cannot have experience of noumena; and (iv) the conclusion that since we cannot have experience of noumena, there are limits on what we can know about noumena. On this abstract characterization, transcendental idealism can sound anodyne, in the sense that most philosophers would probably acknowledge that we cannot stand apart from whatever mental activity is necessary for the world to appear to us, in order to know it independently of this activity. But Kant holds that the mental activity involved in making phenomena the way they are is quite radical. Kant argues that we have synthetic *a priori* knowledge about phenomena which is only possible if our minds contribute space and time to reality—that is, if the very fact of the spatial and temporal extension of phenomena is due to our own mental activity.