Thomas Reid and Some Regress Arguments

by Christopher Yeomans (Gambier)

Abstract: This paper reconstructs Reid's responses to regress arguments against the possibility of free will, highlighting the role played by long-term decisions ("general fixed purposes") in the explanation of paradigmatic free actions on Reid's account. In addition to reconstructing Reid's response to the two versions of the regress argument that he explicitly discusses, I also construct a Reidian response to Galen Strawson's contemporary version of the regress argument. The depth of Reid's position is most apparent in the resources it provides for responding to this sophisticated articulation of a traditional argument against freedom of the will.[1]

Libertarian theories of action face the worry that they entail an infinite regress of choices. In a very basic form the argument can be formulated as follows: The libertarian claims that we can freely determine our choices. However, for this prior determination to be free, it is necessary that it be freely chosen (or freely willed), for if it is not then the whole action sequence springs from a deterministic element. For this next choice to be free, it must be freely chosen, and so an infinite series of choices is generated (either temporally or merely conceptually). But this is impossible, and so one must conclude that we do not, in fact, have such a freedom of the will. This argument is by no means a recent invention. Hobbes puts the point thus: "[T]he will is not voluntary. For a man can no more say he will will, than he will will will, and so make an infinite repetition of the word will, which is absurd, and insignificant."[2] Thus, the regress argument is intended to demonstrate the conceptual impossibility of a libertarian conception of agency.

In this paper I consider the ability of Thomas Reid's libertarian theory to respond to three versions of the regress argument. The first two are explicitly considered by Reid; they derive from Hobbes and common sense, respectively. The third is the contemporary work of Galen

[1] The author would like to thank Paul Hoffman, John Fischer, Larry Wright, Pierre Keller, and Matt Talbert for useful comments on earlier drafts of this paper.

[2] Hobbes 1994, 72 (chap. 12, par. 5).

Archiv f. Gesch. d. Philosophie 88. Bd., S. 54–81 DOI 10.1515/AGPH.2006.003

ISSN 0003-9101

Strawson. In the end I conclude that these arguments have no force against Reid's theory of agency because they rely on specific conceptions of freedom and explanation which Reid does not and need not accept.

1. Hobbes' Version

Reid considers Hobbes' version in connection with his own doctrine of moral liberty.[3] Moral liberty is a power over the determinations of one's will such that one causes the volition one has. One's volition is simply the decision that one makes to do an action or not to do it. Because Reid thinks that all powers are two-way powers (i.e., if I have the power to do X, then I have the ability not to do X as well (35)[4]), the existence of such power is incompatible with causal determinism.[5] Reid calls determinism "necessity", and it is simply defined as the absence of moral liberty (259, 261).[6] Under the hypothesis of necessity, we are determined to will as we will (or to decide as we decide); if we are so determined then our actions are voluntary but not free (260 f).

Reid presents Hobbes' argument as a challenge to the conceivability of free will, and the argument runs as follows: The only valid conception of liberty is acting as we will (i.e., acting voluntarily or possessing freedom of action). Thus, free actions are those that follow from the will. As a result, a free volition is just one that is willed, and which therefore must have another volition which determines it. But this generates an absurd regress of volitions, and so the notion of a free will is inconceivable (263). On Hobbes' view, deliberation is just a temporal series of desires and fears, the last of which is called the will. So in Hobbes' formulation of the regress argument there is both a conceptual point (i.e., that desires and fears cannot follow from the will because they are the will) and a temporal point (i.e., thinking that they could follow from the will commits one to an infinite temporal series of alternating desires and fears each of which are called a 'will').[7] Freedom cannot obtain with respect to volitions themselves, but only to what is determined by them (i.e., the actions of bodies).

Reid thinks that the argument *presumes* that no other conception of freedom is possible in order to *show* that the particular conception of freedom as freedom of the will is impossible. This seems to be an accurate representation of Hobbes' position. Reid's response is to challenge directly the claim that voluntariness is the only con-

---

[3] This argument is certainly not unique to Hobbes. See also Locke 1959, 329 (Book II, chap. 21, §25), and Leibniz 1981, 181f. For a presentation and discussion of Jonathan Edwards' version of the regress argument, see Weinstock 1975, 341f.

[4] Parenthetical references are to Reid 1969a.

[5] For a good general discussion of Reid's conception of power in the context of his contemporaries, see Weinstock 1975 and Rowe 1991, 45–74.

[6] See also Reid 2001, 10f.

[7] Hobbes 1994, 71f. (chap. 12).

ceivable idea of freedom. For if there is *prima facie* evidence that other conceptions of freedom are plausible, then Hobbes' argument is viciously circular. In support of this response, Reid distinguishes three different conceptions of liberty, all of which he takes to be present in common sense. Each conception is distinguished by reference to that to which it is opposed: the first to physical confinement, the second to legal obligation, and the third (moral liberty) to determinism.

Reid thus has a double response to Hobbes' argument. On the one hand, he responds by claiming that Hobbes' identification of freedom with voluntariness is false as a claim about linguistic usage or the contents of our common understanding. Reid argues elsewhere that the existence of an opinion about something – even a false opinion – is conclusive evidence that we have an idea of that thing in the minimal sense required for this aspect of his response (29). Thus, it is too much for Hobbes to claim that the notion of free will is "absurd speech", at least if absurd is taken to mean 'meaningless'.[8] On the other hand, Reid must make good on his own conception of moral liberty so that it compares favorably with Hobbes' voluntariness as a form of action explanation. Only then can Reid make a stronger comparative claim to the conceivability of free will. In this section I will consider only the first aspect of Reid's response, but I take up the adequacy of Reid's own conception of moral liberty in sections 3 and 4.

Reid takes the first notion of liberty to be the one operative in the regress argument. This is freedom on the output side of the will, or freedom of action (i.e., the lack of any constraint interposed between a volition/decision to do X and actually doing X). As Hobbes puts it, "LIBERTY, or FREEDOM, signifieth (properly) the absence of opposition; (by opposition I mean external impediments of motion;) and may be applied no less to irrational, and inanimate creatures than to rational."[9] This liberty is taken away by physical confinement because we can will to leave our prison cell but we cannot actually walk out the door. Thus, in Reid's jargon, this conception of liberty "extends […] to actions consequent to [the will's] determinations" (263).

Hobbes, then, takes the notion of a free will to be nonsense because "when the words *free*, and *liberty* are applied to anything but *bodies*, they are abused; for that which is not subject to motion is not subject to impediment […]".[10] On Hobbes' view, what we call the will is simply the last desire that precedes an action.[11] If the action is contrary to that desire then the agent has been constrained. If the action is consonant with the desire, then the action is free.[12]

Reid agrees with Hobbes that the will cannot be free in this first way (i.e., free of any constraint interposed between the volition and the action), precisely because this

---

[8] In Hobbes 1999b, 16.
[9] Hobbes 1999a, 139 (chap. 21, par. 1).
[10] Hobbes 1999b, 139 (chap. 21, par. 2).
[11] Hobbes 1994, 71 (chap. 12, par. 2).
[12] Reid follows Locke in holding that the inclusion of desires and other passions in the concept of the will is an important error. On Reid's view, passions can be motivations to act, but a distinct and additional decision of the agent is required for the action to occur.

conception of freedom is *inapplicable* to the will. But on Reid's view this means that the will cannot be constrained or unfree in this manner. As Reid puts it, "the will cannot be confined by external force" (264). The inference from the inapplicability of a particular conception of freedom to the will to the conclusion that the will is not free at all is only valid in the absence of other potentially applicable conceptions of freedom. Otherwise, it remains an open question whether the will is free according to an alternative conception. Thus, Hobbes' argument has no force against a theory that can provide an alternative notion of freedom that does apply to the will.

There is an immediate difficulty, however. Reid seems to be arguing that one can have freedom of the will without freedom of action, for physical confinement which takes away freedom of action does not take away freedom of the will. On Reid's view, one's ability to determine freely the content of one's will is unconstrained by the existence of limitations on one's ability to actualize the object of one's volition (in the prisoner case this object is 'walking out the door'). This is problematic given that Reid thinks that we can only will in cases such that the willed action is in our power – or at least in cases such that we think that the willed action is in our power (58). Reid gives the example of a man who is affected with palsy in his sleep (and is therefore unable to speak). When he wakes up, he first tries to speak, not knowing that it is no longer in his power. But once he realizes that he does not have the ability to speak, he cannot will to speak, but can at most will to try to speak (62). But in the case of physical confinement, walking out the door is not within the prisoner's power – and here I envision a case in which the prisoner knows with certainty that escape by his own efforts is impossible. If it is not within the prisoner's power to walk out, it seems that the prisoner's will is constrained in the sense that it is not within the prisoner's power to will to walk out once he knows that the door is locked.

In defense of Reid, it is important to keep the details of his notion of moral liberty in mind. In particular, moral liberty is opposed to the necessitation of the will by something other than the agent. Thus, the relevant sense of 'constraint' or 'confinement' of the will is the will's necessitation one way or the other, and not the sense of restriction of possible objects of the will. So in the case of physical confinement, the prisoner's freedom of will would be constrained (in fact eliminated) only if her captors were to make her will not to walk out the door.

On Reid's view, the will is just the power to decide in situations where our decision matters, i.e., to decide for or against actions which depend on our decisions (57). But an action that is the result of external constraint does not depend on my decision; there is neither the need nor even the opportunity to decide anything in such contexts. The prisoner's will is not necessitated or forced to take a certain content; rather, the situation offers no content for the will.[13] Rogers Albritton makes the same point as follows:

---

[13] This is not to say that the prisoner has no will. If I am stuck in a cell, I can still undertake the voluntary act of turning my attention to spiritual matters, or deliberating as to the proper form of my revenge on my captors if and when I do get out. Reid (76–89) thinks that deliberation and attention are voluntary acts.

> Of course, chains and the like can in a certain sense undermine apparent choices. But it isn't the freedom of those choices that chains can abolish, it's the *choices*, in an objective sense […]. Stuck in the balcony, am I? Well, then, alternatives between which I might have chosen have gone glimmering. *Both* of them have. They are out of logical range, so to speak again. The thought, as the auditorium empties, "Well, I'll just stay here in the balcony, then, *that's* what I'll do" isn't even an expression of *amor fati*. It's just fatuous.[14]

This is really quite plausible, as we do not normally think it a restriction of our freedom of will that we cannot jump to Mars or transform ourselves into silicon-based life forms. Freedom to choose is freedom to choose from among available options. But Reid need not go as far as Albritton, who maintains that we have perfect freedom of the will. It might still be true that our freedom of will can be constrained by *internal* forces such as madness or depression, as Reid himself suggests (262). With this potential stumbling block removed, we can now return to Reid's enumeration of the three conceptions of freedom.

The second notion of freedom is political liberty, which is opposed to legal obligation. On Reid's view this conception of freedom does apply to the will. Reid argues that laws are addressed to beings that are taken to have the ability to obey or transgress the law, and thus that legal obligation presupposes an ability to choose for or against the law. Now one might think that a person could follow or act contrary to the law even if his or her will was necessitated to choose either the former or the latter, but Reid argues that this would not count as obedience or transgression. Rather, one only obeys when one freely wills to follow the law, and the same free willing is required for transgression (264). On Reid's view, the very notion of legal obligation presupposes a free agent who can be punished for transgression of the law in virtue of his or her ability to have chosen to obey the law. Now this is clearly different from Hobbes' conception of the nature of legal obligation. Since Hobbes himself assumes that we have only freedom of action and not freedom of will, his conception of legal obligation cannot depend on the ability to freely will to follow or disobey the law. If Reid's view were a correct articulation of common sense reflections on the subject of laws, then this would be evidence for the claim that Hobbes' identification of freedom with voluntariness is false as a claim about common usage. However, it is quite difficult to adjudicate such a dispute. Fortunately, not much should turn on this point in isolation.

Reid's view of the force of law is only coherent to the extent that the conception of freedom of the will that he offers is conceivable. This is the third conception that Reid distinguishes. Reid's conception of moral liberty is explicitly opposed to determinism. Reid frames his conception of moral liberty in terms of his own notion of causality:

---

[14] Albritton 1985, 246.

> 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. The first is as easily conceived as the last. 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. But, if another being was the cause of this determination, either by producing it immediately, or by means and instruments under his direction, then the determination is the act and deed of that being, and is solely imputable to him. (265)

Because Reid's conception of power involves the ability to do otherwise, a power over the will must involve the ability to will otherwise. This is the ability to form a range of different volitions (or make a variety of different decisions) in the same circumstances. Reid equates the power over the determinations of the will with causal control over the will such that one could have produced another volition, had one so exerted one's power of moral liberty. In our contemporary jargon, both alternative possibilities and control are implied in Reid's conception of moral liberty. Now whether this account is ultimately consistent or sufficient is an open question, which I address in the last two sections of the paper. It seems, however, that there is nothing incoherent about it. That is all that Reid needs to demonstrate the circularity of Hobbes' argument. While Reid's conception of causation may strike us as strange, it is neither inconceivable nor *prima facie* inconsistent. To his credit, Reid does not simply assume that this is the only proper conception of cause. Rather, he argues that the naturalistic conception of causation is in fact a derivative application of a primary conception of causation that refers to agents' practical effectiveness in the world (38–43). While Reid does not himself think of common sense in a historical perspective, it is interesting to note that there may be a disparity between our own common-sense intuitions and those of Reid's contemporaries.[15] Deciding what to make of such discrepancies is a difficult matter. We will return to this difficulty when discussing Strawson's own appeals to contemporary common sense.

At this point, Reid has met the minimal demands of the first aspect of his response to Hobbes' version of the regress objection insofar as Hobbes presumes the inconceivability of any notion of freedom other than voluntariness (i.e., freedom of action). The important point is that Hobbes' argument does not demonstrate the inconceivability of that conception (since, in fact, it presumes that inconceivability). Rather it shifts the burden of proof to the libertarian's effort to provide an alternative explanation of free action. That is, it shows that the libertarian cannot rely on voluntariness as a conception of freedom of the will but must provide another conception. In Reid's case, responding to such a burden is going to require a more detailed elucidation of his doctrine of moral liberty. Before proceeding to this elucidation, however, it will be useful to consider a second version of the regress argument to which Reid explicitly responds.

---

[15] For an interesting discussion of this issue, see Rowe 1991, 51f.

2. Common-Sense Version

Because Reid thinks that common sense is an important source of evidence in philosophical investigations, he is obliged to consider a second version of the argument. Common sense, he thinks, holds "That nothing is in our power but what depends upon the will […]"(265). One might conclude from this that we do not have power over the will itself, on pain of the Hobbesian regress. For if the will were to be within our power only in virtue of being controlled by the will, then each free volition would require another prior to it, *ad infinitum* (whether this priority is temporal or merely conceptual). This problem is made even more significant by the fact that Reid himself makes statements quite similar to this common-sense axiom, e.g., "For as all our power is directed by our will, we can form no conception of power, properly so-called, that is not under the direction of will" (309f.). Reid responds by arguing that when we say that some concrete action is in our power, we necessarily imply that the will is in our power because the will is required for such actions. If the will were not in our power then neither could the action be in our power because a necessary means would not be in our power.

Reid argues that this common saying describes only those actions that are "external and visible effects" (265). As for these actions, Reid agrees that those that do not depend on our will are not in our power. Thus, the motion of the planets is not in our power, because it proceeds regardless of how we 'decide' about the matter. Those actions that are in our power must be voluntary, i.e., they begin with a volition (264). Reid then argues that this actually proves that the will is in our power, for otherwise the end (the action in its perceptible manifestation) is held to be in our power, but the necessary means (the volition) is not. Now this is a good argument if one assumes that 'two-way power' is the relevant sense of 'power' here, for it seems that one does not have the ability to produce different perceptible effects if the will is itself determined to choose one way by something external to the agent. There may be good reasons for denying that 'two-way power' is the relevant sense of 'power' here; I assume a compatibilist would take this position. An adequate discussion of this issue would easily fill another paper. Here I simply want to make explicit that Reid's argument relies on this premise.

Reid argues that this version of the regress argument is misled by the universal form of the statement, "That nothing is in our power but what depends upon the will" (266). He thinks that the above argument dem-

onstrates that the will is implicitly exempted from this dependence, and he gives another common saying in which the same kind of implicit exemption is made:

> Thus when we say that all things depend on God, God himself is necessarily excepted. I[n] 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. Every effect must be in the power of its cause. The determination of the will is an effect, and therefore must be in the power of its cause, whether that cause be the agent himself, or some other being. (266)

The analogy with God suggests that the will does not depend on the will, i.e., that the will does not require another act of will to determine it. If this is a plausible interpretation of the common-sense saying at issue, then Reid's response is sufficient to rebut the argument. Unfortunately, it is quite difficult to know what to decide in this connection, especially given that our own common sense may be quite different from that of Reid's eighteenth-century audience.

These responses do not directly help us to understand how to interpret Reid's own claim that all power is directed by the will in such a way as to avoid the regress. One might claim that Reid has simply misspoken here, and that this statement should be rejected because it contradicts his claim that the will does not depend on the will. But this is a disconcerting option given the categorical form of Reid's assertion. That is, the form of Reid's expression leads one to believe that he thinks that this is clearly correct and even analytic to the notion of power: "[…] we can *form no conception of power, properly so-called*, that is not under the direction of will" (309f., emphasis added). If Reid is just flat wrong about this, then that fact seems to indicate a deep problem in one of his crucial concepts. Alternatively, one could apply Reid's second response to the common-sense regress to this statement. One would then interpret the statement as necessarily implying the exception of the will from those things which are under the will's direction. But this option leaves us unable to explain the positive relation between power and the will and the way in which the will could be in our power, and these are elements of Reid's view that need to be elucidated.

Fortunately, there is a third way of interpreting Reid on this point which maintains the force of his claim without leading to the contradiction noted above. To see this, note that the contradiction arises from taking the terms 'directed' and 'direction' in a certain sense, i.e., as indicating the highest-level choosing available (i.e., as indicating 'where the buck stops'). This is the sense in which we might say that a CEO or Board of Directors directs their company. But there is another sense of 'direction' in which instrumentality plays a larger role. In this sense we might say that the power of the company is directed by the middle management, insofar as they unfreely and instrumentally translate the decisions of the Board of Directors into concrete actions. Thus, we might say that all our power is under the direction of the will in the sense that the will is a necessary means to its exercise. It is a necessary means because the decisions of an agent, like the decisions of a Board of Directors, are likely to be rather

general and thus to require interpretation by a particular manner of implementation in order to become actual and effective. For example, if I decide to go to a basketball game, I need to make choices regarding timing, transportation, method of payment, etc. The will is potentially involved not only in deciding *what* to do but *how* to do it. The unity of the agent precludes the kind of division of labor available in the corporate example, so philosophically the concept of will serves the double function of describing decision and the implementation of decisions. Reid identifies the will with this second aspect (the implementation of an agent's power). The agent as a whole is the proper cause of an action (and thus the CEO to the will's middle management). The agent's causal power generates the first general decision (in Reid's jargon, a change in the determination of the will), and then the will proper implements this decision, thus directing the agent's causal power. I admit that Reid himself does not draw this distinction, but it is consistent with Reid's response to the common-sense regress argument, and it does not sever the connection between power and will.

This hermeneutic option has its own consequences. The interpreted statement comes from Reid's first argument for the existence of our power of moral liberty. In this section he argues from our natural belief in this power to the existence of the power. But even as I have interpreted the statement it is clear that it cannot apply to active power in the broadest sense that includes moral liberty (because then we still have the contradiction), but only to the more limited sense in which active power is what makes our volitions effective. Thus, my interpretation leads to the conclusion that Reid is mistakenly arguing from a belief in active power in the narrow sense (i.e., freedom of action) to the existence of moral liberty. However, there is independent evidence for this confusion in Reid's argument.[16] The most obvious two cases are Reid's third and fourth arguments for the necessity of our belief in such liberty. The third argument is that we must believe in our ability to act consonant with our decisions – this looks like a definition of freedom of action – and the fourth argument concerns the special case in which the decision is a promise (305f.).[17] Thus, I think the interpretation preserves as much sense as can be made of Reid's claim, while eliminating the possibility of a regress.[18]

---

[16] This is a deep difficulty, for if our conception of causation comes from our experience of practical effectiveness as agents, and yet that practical effectiveness can be explained in terms of freedom of action alone, then it seems invalid to argue from that effectiveness to moral liberty.

[17] Paul Hoffman pointed out this problem in Reid's argument to me.

[18] There is a third potential form of the regress argument which turns on the notion that the exertion of active power must be an event that requires a cause and thus an exertion of power. This is articulated in Rowe 1991, 147–154. Rowe offers what I take to be the correct solution to this regress, which is to deny that an exertion is an event in the same sense as a bodily movement is (i.e., in the sense of being a change that a substance undergoes). Rowe finds this denial implausible, as does McDermid 1999, 298–300. See also O'Connor 2000, 49.

3. Reid's Theory of Action

Before going on to discuss Strawson's argument, it will be useful to stop briefly to articulate the main elements of Reid's theory of action, and in particular to clarify the relation between moral liberty and the will.[19] To do so will require us to explain the relation between reason and free action, and this will motivate Strawson's argument by showing how and where it has a foothold in Reid's libertarianism. In addition, the general development of Reid's position will give us the resources to provide a Reidian response to Strawson's argument.

The will can take as its object actions that depend on the will, i.e., the production of changes in thought or bodily motion.[20] The object of moral liberty is the production of a change in the determination of the will through the exertion of active power – and because this action does not depend on the will it is, strictly speaking, neither voluntary nor involuntary. Instead, it must be understood to be directly caused by the agent.

The paradigm cases of moral action, on Reid's view, are those in which we display self-government. Self-government is when we make our rational choice (i.e., our best judgment) effective over against the influence of our passions or desires. Self-government requires what Reid calls "fixed purposes".[21] A fixed purpose is a volition which has as its object a future action, and which is usually formed after a period of deliberation (83). A particular fixed purpose has as its object an individual action; once the action is undertaken, the volition leaves the mind (84). Now insofar as particular fixed purposes follow an agent's

---

[19] Since my goal in this paper is to reconstruct Reid's responses (actual and potential) to versions of the regress argument, I only delve into the structure of Reid's conception of agency where it is necessary for this purpose. For general discussions of Reid's conception of agency, see Rowe 1991, McDermid 1999, Weinstock 1975, Woozley 1987, Madden 1982 and Lehrer 1989, 203–287.

[20] Eliminating changes in will from the list on 268.

[21] Although Reid never makes this claim exactly as I have formulated it, it seems clearly implied by his discussion on 85f. This is clearest at two points. First, fixed purposes are used to explain the voluntary self-mastery of virtue in contrast to the involuntary promptings of "natural affections" (85). Second, in one statement of this contrast Reid seems to identify fixed purposes and self-government: "In men who have no fixed rules of conduct, no self-government, the natural temper is variable by numberless accidents" (86). At the very least it is true that paradigm cases of self-government require fixed purposes, but I think the general drift of Reid's discussion strongly suggests the stricter notion that they are required for self-government as such.

deliberation, they are one way in which we govern or set ends for ourselves. But those "who have made the greatest advance in self-government, are governed, in their practice, by *general* fixed purposes" (86, emphasis added).

A general fixed purpose ("GFP") is a volition that has for its object "a course or train of action, intended for some general end, or regulated by some general rule" (84). One example Reid gives is the decision to go into a certain profession (84), but by far the most important examples are our resolutions to act justly, honestly, etc. On Reid's view, to ascribe a particular virtue to a person is just to say that she has formed the corresponding GFP (85f.). GFP's are crucial to self-government because they direct an agent to virtuous action even when her own emotions incline away from such action (86). Furthermore, these purposes require reason. GFP's include either a general end or a general regulative rule, and on Reid's view reason is required for the formation of all abstract and general principles.[22] To transpose Reid's view into the terms of contemporary agency debates, our earlier conclusion that (on Reid's view) a volition cannot take the will as its object means that these general fixed purposes are most plausibly understood not as second-order volitions in the sense that they concern what first-order volition should be effective for a particular action, but rather as first-order volitions in their own right. This is consistent with the way Reid explains the function of GFP's. But they might be understood as second-order volitions in the sense that they serve as the central organizing principles of an agent's character. Other, more local decisions might be made that were directed towards the maintenance of these long-term decisions (e.g., decisions to avoid temptation or to familiarize oneself with moral exemplars as a way of strengthening one's commitment to one's GFP's).

One might think that Reid's notion of a volition extending over time is incoherent. Normally, when we think of a volition we envision a decision to do something immediately, or what Timothy O'Connor calls "immediately executive states of intention to act"[23]. But in Reid's defense we should note that insofar as even momentary acts of will have some temporal duration, there doesn't seem to be any theoretical incoherence in thinking that an act of will might continue for longer. Furthermore, it seems plausible that one might form a volition (i.e., make a decision) to take the dog for a walk as soon as the basketball game is finished. It doesn't seem plausible to say that, when the game is over, another decision is required to take the dog for a walk.

---

[22] Reason is required only for their formation, and not for merely conceiving of them (i.e., representing them). See Reid 1969b, 541f.

[23] O'Connor 2000, 72.

Rather, I act on the basis of my earlier decision. The decision persists in the meantime, although I need not have it consciously before my mind at all times. The ease and rapidity with which I could respond to a question as to what I was doing after the game would be indicators of such persistence. Of course, I might change my mind, but this is no counter-example; rather I simply change my volition before I execute it. Thus the mere existence of temporal distance between the formation of a fixed purpose and its execution does not require another volition. In fact, such a requirement would quickly generate an infinite regress of its own. On Reid's view, the volition-action sequence is not an inherently temporal sequence at all. In contemporary terms we should understand it instead as an analytical distinction regarding the conceptual priority of different elements of agency. This is another reminder that Reid's agent causation is not to be understood in terms of the general model of physical event causation in which there is an intrinsically temporal distinction between cause and effect.

One might still think that the contrast between the generality of a general fixed purpose and the particularity of the actions which it regulates would require a second volition, but Reid does not see things this way. Instead, Reid writes of a GFP as a volition to act according to a rule "when there is opportunity" (85). Thus, what mediates between the general purpose and the particular action is the existence of an opportunity for action, and this is built into the structure of the GFP. Now it might still be the case that the voluntary act of deliberation is required in order to determine whether a given circumstance is, in fact, an opportunity for the execution of a GFP, but once that determination has been made the action follows directly from the GFP without the interposition of a second volition for the particular action. In this sense GFP's are similar to Kantian maxims, at least when understood along the lines suggested by Barbara Herman. On Herman's view maxims are something like "deliberative presumptions". That is, they serve directly as the principles of our will in the straightforward cases, and in difficult cases we must deliberate as to which principle the case at hand falls under.[24]

The GFP's explain how self-government with respect to an individual action works. It is not the case that for each free action one's will is initially undetermined and what ensues is a struggle between reason and the passions to determine the will. Rather, the will is already determined in advance, and the struggle ensues at the level of deliberation, where we are tempted by our passions to "reopen the case" for the GFP and therefore are tempted to change that purpose itself. Thus, Reid writes that "every known transgression of justice demonstrates a change of purpose at least for that time" (85).

From the contemporary perspective of interest in action explanation, the doctrine of GFP's is both a strength and a liability for Reid's view. It is a strength because Reid's insistence on the relative fixity and stability of such purposes gives them an explanatory value not usually possessed by 'reasons' for action. Because our liberty is bound up with self-government, and self-government is interpreted as the mainten-

[24] Herman 1993, esp. chap. 7.

ance of our fixed purposes through changing circumstances, the requirements of stability and consistency of purpose are built into the very conception of free action. As a result, free action can be explained by the conjunction of such resolutions and given opportunities (where the latter can presumably be specified in such a way as to explain the timing of the action). Reid puts the point as follows: "When we know a man's principles, we judge by them, rather than by the degree of his understanding, how he will determine in any point which is connected with them" (87). Reid further claims that for any given agent it is impossible that "the general tenor of his conduct should be contrary" to his GFP's (93). Our distinctively moral character (as opposed to our emotional predispositions or 'affections') is composed of fixed purposes oriented towards the rightness or wrongness of actions. That is, Reid thinks that when we attribute a particular virtue to someone, we are saying that he or she has formed the corresponding fixed purpose (85). To apply this point to contemporary worries regarding action explanation, we should take Reid's conception to mean that character virtues are already to be conceived as decisions grounded in the activity of the subject, so that the determinative relation between virtues and one's actions could in principle be fully deterministic without endangering the freedom of the agent. That is, while from the event-causal perspective the agent as such is dispersed into different events whose causal relations threaten the freedom of one of those events (e.g., a choice or an action), from the agent-causal perspective these distinctions are internal to the agent's nature taken as a whole.

But this doctrine is a liability because it seems only to push the explanatory question back one level. Now we want to know about the formation of these GFP's and whether they were freely formed. After all, moral liberty on Reid's view is the ability to freely determine the contents of one's will. This question is more complicated because moral liberty not only involves active power, but also reason. Now one might think that moral liberty is equivalent to active power, and that it is therefore in virtue of our possession of the latter alone that we are free.[25] Reid, however, argues that reason or practical judgment is required for moral liberty, i.e., for the robust conception of freedom required to ground moral responsibility (259f.). Reid even goes so far as to claim that moral liberty is the conjunction of reason and active power (301).[26]

To see why Reid makes this claim, it is useful to analyze his conception of moral liberty in terms of the two requirements for a libertarian conception of freedom as articulated in the contemporary debate, i.e.,

---

[25] For an interpretation that makes this equivalence, see Stecker 1992, 197–199.

[26] Weinstock 1975, 340, 343, sees this connection of reason with free will as Reid's most important contribution in the field of agency, and thinks that such a connection is sufficient to respond to the Hobbesian regress. Although I agree that this connection is important, it also provides a basis for the third version of the regress argument which I consider in section 4.

alternative possibilities and control. Active power guarantees alternative possibilities because all powers are two-way powers and such powers are clearly incompatible with determinism. Control involves the ability to select effectively between alternative possibilities in such a way that the volition selected represents the agent in some important respect.

Active power provides for the agent's effectiveness in choosing in a limited sense, because it is the quality of the agent in virtue of which the agent actually *makes* his decision (268). But the choices which are thereby made effective are made into 'choices' in the robust sense by the practical judgment or reason of the agent, for only this faculty can determine reasons for preferring one choice to another (259f.). The choices, then, are the way in which the agent sets ends, and these may be the ends supported by judgment or those supported by desires. In paradigmatic free action these ends find their expression in patterns of actions organized by a long-term decision to act in a certain way (a virtue interpreted as a GFP). Action explanation implicitly involves relating an act to such a decision, and thus placing it in a certain pattern organized by a choice that is at least in principle responsive to reasons.[27] This view avoids the problem of the mysteriousness of the agent-cause at the expense of introducing reason as a necessary element in free action. This then suggests Strawson's version of the regress argument, which focuses on the problem of the source of our volitions and the motivations which are involved in our decisions.

4. Strawson's Version

Galen Strawson's formulation of the regress argument is an objection to the notion of self-determination, and can be summarized as follows:

(1) The paradigm case of free action is rational action.
(2) Rational action is "a function of, or determined by, how one is, mentally speaking."
(3) To be truly responsible for an act one must be responsible for one's mental state at the time.

[27] The notion of 'action explanation' is to some degree foreign to Reid's exposition, since the explanation of individual actions is not attempted in the way a contemporary reader might expect, nor is a procedure explicitly suggested for such explanation. However, the connections between Reid and contemporary concerns regarding agency can be made clearer if one can see what action explanation comes to if one accepts Reid's conception of agency.

(4) To be responsible for one's mental state requires that one has "consciously and explicitly chosen to be the way one is, mentally speaking."
(5) But one can only be said to have chosen one's mental state on the basis of principles [i.e., the choice is a function of those principles].
(6) To be responsible for this choice requires that one be responsible for these principles, which requires that one has consciously and explicitly chosen them.
(7) This choice requires principles which must themselves be chosen, etc.
(8) Thus, "True self-determination is logically impossible because it requires the actual completion of an infinite regress of choices of principles of choice."[28]

As we have just seen, Reid agrees with (1) in a certain sense of 'rational action'. That is, Reid thinks that a 'choice' in the robust sense requires reasons for thinking that the choice was good or bad. On that much Reid and Strawson agree.[29] They disagree, however, on how to understand the explanatory role of reasons. As articulated in (2), Strawson thinks that the only possible model by which the explanatory force of reasons can be understood is that of causal force or "full determination". The only alternative is randomness.[30] This is most succinctly captured in the claim that an action is a *function* of one's reasons. (From now on, I will refer to this as the "functionalist" model of action explanation.) This model requires that reasons – understood as mental states (beliefs and desires) – have full metaphysical standing so that they can be capable of standing in causal relations.[31] Thus there are two main

[28] This version is drawn from Strawson 1995, 13–31. The same argument can be found in Strawson 1986, 28–30.

[29] The discussion in Strawson 1986 seems to profit from an ambiguity in the notion of rational action. In making the claim that such actions must be free the implicit appeal is to rational action as thoughtful and reasoned action (see *ibid.*, 32f. and 40f. for examples of this claim). But in the argument itself reasons are understood simply as mental states (*ibid.*, 34). In this sense all explicable intentional or even conscious actions seem rational. So there is no intuitive appeal to the notion that all rational actions are free in the sense of 'rational action' required by the argument.

[30] Strawson 1986, 25.

[31] Strawson steps back from the claim that the relations involved must be causal to the claim that they must be fully determinative. But neither 'cause' nor 'full determination' receives any determinate content in Strawson's exposition. In fact, 'full determination' is simply defined in terms of the satisfaction of the conditions of rational explanation. As a result it is difficult to know precisely what the distinction between the two notions is. I take it that Strawson wants to argue that his argument does not depend on any particular theory of causation. Since I do not take issue with Strawson based on a particular theory of causation, from here on I will simply refer to the relation Strawson posits between reasons and actions as causal.

features of Strawson's functionalist model of action explanation: (1) the action is a determinate function of one's reasons; and (2) such reasons must be fully real in order to play this determining role. Strawson proceeds to apply this model at each stage of the regress and to each point at which the libertarian might insert indeterminism in the causal stream of action. Because this functionalist model is the crucial premise in the argument, we can dispense with consideration of the application of the argument and focus on the model itself. If Reid can consistently and plausibly reject the model, then particular applications of it are unimportant.[32] In addressing a similar sort of argument, Reid attacks both features of this model.

On Strawson's view, an agent *a*'s reasons are "real things, desires and beliefs, things that he actually has. Or, in an alternative idiom, they are real states of *a*, desire- and belief-states, states that he is actually 'in'."[33] To begin with, Reid explicitly denies that reasons ('motives' in Reid's jargon) can be fully real: "We cannot, without absurdity, suppose a motive, either to act, or to be acted upon; it is equally incapable of action and of passion; because it is not a thing that exists, but a thing that is conceived; it is what the schoolmen called an *ens rationis*" (283). Furthermore, even if motives were to exist in this way, they would not be agents and would thus be incapable of being causes, on Reid's view (285). In the scholastic division of *entia rationis* into negations, privations, and relations, a motive could most plausibly be considered a kind of relation between an action and either a value or state of affairs that might result from an action. It is doubtful, however, that Reid has anything so metaphysically complicated in mind. Instead, I think the force of the reference is to emphasize the dependence of motives on the activity of the agent. I take it that his notion of motives as *entia rationis* is part of the agent-causal picture in which reasons are not independent events or entities with their own proper identity conditions, but are rather analytically distinguishable only as aspects of an agent's activity. They are not independent states or even real properties of the subject, but are rather mental creations of a subject who manipulates them in certain ways to comprehend the stakes of action. In interpreting Reid's view, one need not think that they are somehow in a separate, Platonic realm of existence. Reid's view is in fact stronger, denying them inde-

[32] Strawson discusses a Leibnizian view that is similar to Reid's (Strawson 1986, 52–56). Because his criticism of that view depends on the functionalist model, I have omitted specific discussion of that criticism.

[33] Strawson 1986, 34.

pendent existence in any form. In the scholastic tradition, *entia rationis* were beings that lacked real affirmative being but acquired another sense of being in virtue of their positing by the intellect of a subject (e.g., blindness as the privation of sight). In the case of agency we are primarily dealing with the first person perspective of agents whose own mental activity grounds the *entia rationis* in question, but we might also predicate a reason with respect to some other agents even if they were not aware of that reason and thus no mental state corresponded to it (e.g., we could say that an agent should have considered some reason that was not considered, or had good though unknown reasons for doing the opposite of the action performed). In both cases, reasons cannot be understood as real properties, states, or even events.[34]

Strawson argues for his view of the metaphysical status of reasons by appeal to the notion that functionalist explanations of action can be "simply true", and that this is only possible if reasons are "real as roses"[35]. Strawson glosses 'simply true' as the claim that simply true explanations really do state the reasons for which the action was performed.[36] An explanation is 'simply true' in the relevant sense if it corresponds directly to some real (non-linguistic) relation between reasons and actions. This is why reasons must have the same metaphysical standing as actions.[37] At another point, in trying to clarify the notion of full determination which is defined in terms of simply true explanations, he claims that when an explanation gives full determination that explanation is "all there is to it" and "the whole truth"[38]. He also says that for a rational action "it is correct to say *sans phrase* that it was performed for a certain reason"[39]. Thus, I take it that 'simply true' means something like 'corresponds to reality without qualification or restriction'. The problem is that on this interpretation it is difficult to make sense of his acknowledgment that such explanations operate only at a certain (but unspecified) level of abstraction. (No such explanation is complete in the sense of explaining such details as the trajectory of the physical movement involved in actions.[40]) This process of abstraction would seem to be an explicit qualification on the truth of such ex-

[34] For a brief review of *entia rationis*, see Klima 1993.
[35] Strawson 1986, 33.
[36] Strawson 1986, 33.
[37] Strawson 1986, 35.
[38] Strawson 1986, 38.
[39] Strawson 1986, 37.
[40] Strawson 1986, 38.

planations. Strawson must mean that such explanations fully explain certain *aspects* of the action.

Thus the view seems to be that one can isolate a certain subset of antecedents and qualities of the action without significant reference to that which lies outside of that subset, and without reference to the practice within which such isolation has significance. Only if this is possible could such explanations be simply true in the sense required to motivate us to endorse Strawson's view of the nature of reasons at the expense of Reid's view. Here we must depart from Reid's text to defend his view, and reply to Strawson that such isolation is not possible. The reason that the relevant level of abstraction is so intuitive to us is that we are used to picking out *morally* relevant features of actions and antecedent states. First, this suggests that explanations which operate at this level of abstraction are not 'simply true' in an unqualified sense but are rather crucially parasitic on (i.e., qualified by) our practices of moral evaluation. That is, such explanations correspond or refer to real relations only in virtue of their involvement in such practices (i.e., the explanation does not correspond directly to those relations, but only through the mediation of our moral practices). Put another way, the reference between the explanatory or discursive connection (the 'because') and the supposed real causal connection is not based on a direct identity between the connections themselves. Rather, that identity between the connections is only provided by a practice that gives us a way of interpreting that discursive 'because'. But because such an interpretation is in fact constitutive of the explanation, it counts as an internal qualification or condition on the adequacy of that explanation. This means that the explanation is logically dependent on the practice that provides that interpretation, and is therefore constrained by the conditions provided by practice.

Only if one has specified and limited one's requirements for that explanation can one say that such explanations are complete or "the whole truth". For example, if an agent's desire to save a child's life and belief that picking a struggling child out of a pond will save her life are to constitute a full explanation of the agent's going into the pond, then it must be specified that we are not interested in the way in which the agent enters the pond or pick up the child, nor in the history of the desire or belief, nor in any other beliefs or desires that the agent had related to the situation, nor in how such beliefs and desires could determine an action. Without such explicit limitation on the kind of information we require from the explanation, it would be impossible to consider it a full explanation. For any given action the number of ques-

tions that could be asked, and corresponding explanations that could be given, is enormous. A psychologist, a physicist, and a moral philosopher would be interested in very different information.

One accepts a Strawsonian explanation as fully sufficient only within certain practices of moral evaluation. From a Reidian perspective, the procedure for interpretation that the practice of moral evaluation provides is precisely the connection of volitions and actions only against the background of an agent's responsibility (after all, Reid's preferred term for our freedom is 'moral liberty'). Such explanations then use beliefs and desires to articulate the moral content of the action as that which is liable for judgment. The validity or adequacy of this form of explanation is therefore dependent on the validity or adequacy of our evaluative practices. If those practices are unjustified (i.e., if their adequacy cannot be demonstrated), they should be rejected. But because those practices are constitutive of the kind of explanations we are considering, this means rejecting this form of explanation as inadequate and unjustified. Because the justifiability of the evaluative practice is just what Strawson denies in claiming that responsibility-entailing freedom is impossible, his position therefore entails the inadequacy of the explanations as well. His position is therefore self-contradictory in the sense that it necessarily presupposes what it appears to reject.

Second, one might argue that the explanations in question do not depend on the validity or justifiability of the evaluative practice, but only on its existence. That is, one might take Strawson to be exploiting a feature of that practice in order to show the incoherence of the practice as a whole. But even if one argues that the sufficiency of such explanations is not linked to the validity of our practices of moral evaluation, it at least involves reference to questions that arise in the context of moral evaluation. Thus it will never be true that such an explanation will be true without qualification or absolutely true.[41] Such explanations explain the moral content of the action, and the notion of moral content only has meaning within a system or practice of moral evaluation. If the sufficiency of such explanations is parasitic on our practice of moral evaluation, we should give these explanations an interpretation that is consistent with the practice. Thus, if Strawson's conception of expla-

[41] Strawson 1986, 36f., responds to a similar concern as to the relativity of an explanation to an audience by limiting 'rational explanation' to that subset of the whole audience-relevant explanation that is not so relative. However this terminological shift produces an empty conception if the very level of explanation is context-relative.

nation is inconsistent with that practice, the former should be discarded and not the latter.

Now one might think that Strawson's position could be softened to allow for this, while maintaining that one's reasons really do explain one's actions. That is, one might think that a 'simply true' explanation might only require that the agent acted for certain reasons and that the agent actually had those reasons. But notice that this is to back away from the conception of full determination by reasons that Strawson posits, because now no causal relation is required between the reasons and the action. This means that action is no longer a function only of one's reasons, and so premise (2) in Strawson's argument must be discarded. But this is the crucial premise in the argument, and so the argument as constructed fails. Instead of being stuck with the alternative between full determination and randomness, we now have the possibility of another form of determination. There is no reason in principle why Reid's conception of agency could not provide this new form of determination, i.e., this interpretation of what it means to act for a reason. To put this point in a way that connects it with the previous discussion of moral practice, now instead of a two-term relation between reasons and action, we have a three-term relation in which reason and action are connected by our moral practice. That is, their relation is not one of direct determination but rather a relation mediated by that practice as providing an interpretation of the 'for' in 'for reasons'. If one wanted to maintain a correspondence between elements of the explanation as such and elements of the real actions that are to be explained, one might look for a third thing in the agent which corresponded to our moral practice. A plausible suggestion would be that this third thing is just our moral liberty (i.e., our internal moral practice).

To return to Reid, the non-determinative role he assigns to reasons seems more consistent with our practice of moral evaluation than the full determination posited by Strawson. If such explanations respond to questions concerning how we should understand an action in a moral sense, they do not obviously call for a causal or determinative answer. In moral evaluation we are primarily concerned with how the agent understood the action. This helps us to understand what we are evaluating when we make a moral judgment. In this context, an agent's reasons really do give us a good explanation of the action, just not in the sense of 'explanation' that Strawson accepts. Thus, there are good reasons to deny Strawson's founding assumption that action explanations can be 'simply true', and with it the claim that reasons must be fully real and determinative.

To be fair, Reid's justification for his own position on this matter is also undeveloped. He seems content to shift the burden of proof back to his opponent. He argues that in conceiving of reasons as causes one conceives of the agent as a passive substance that changes in ways that are proportional to the amount of force impressed on it by its reasons. This contradicts what Reid takes to be the common-sense notion of the agent as a cause, which Reid thinks is the only clear notion we can form of efficacy.

Although one could pursue Reid's arguments for his conception of causality to defend him on this point, I think that this would be a distraction, for Strawson is right that the kind of worry he is pressing against libertarian views does not depend on any particular causal conception of reasons (see above, note 31). Even if Reid is granted his conception of causation, so long as reasons over which we have no control are necessary and sufficient conditions for action, one still cannot explain how free action is possible.[42] Against this sort of view Reid claims that some free actions require no motive, and thus that reasons are neither necessary nor sufficient conditions for action.

Reid thinks that there are many free actions that take place without any determining (i.e., necessary and sufficient) motive. His primary examples of this include situations in which one has the choice between two equally sufficient means for obtaining an end. One chooses without having any reason to prefer one to the other (285). Reid acknowledges that there can never be a moral evaluation at stake in such actions, and thus that they are 'free' in a thinner sense. This would be what freedom amounted to if it were equivalent to active power, which falls short of what we want out of a notion of freedom. Nonetheless Reid thinks that the existence of such situations is indirectly relevant to the question of free rational action insofar as it shows that we must have the capacity to determine our will independently of motivations. Reid's claim is that such actions have and require no motive, and thus that motives are not a *necessary* condition for action.

But one might think that even in these circumstances one has reason to choose one course of action just because the choice has to be made.[43] We pay with one particular coin just because we want to buy the product and go home. Reid's response has to be that this case at least shows that reasons are not *sufficient* conditions, because a free choice of the agent still must be brought in to explain why we used *this coin as opposed to that one.* No functionalist reasoning from motivations to action will be able to provide a fully contrastive rational explanation for such actions, i.e., no such explanation will explain why I use this coin instead of that one. One may be able to provide a reason why I choose without a reason (e.g., shortness of time), but this does not provide the kind of contrastive explanation that is demanded of the libertarian. Yet such actions are imputable to the agent in a basic sense – they are her free actions – even if no moral judgment is involved (285f.). Essentially, Reid is arguing that there are actual, free actions that cannot be given a functionalist explanation because the reasons for them do not play the requisite role. Yet such actions seem perfectly intelligible, rational, and explicable. Again, this sense of 'free' is certainly not the full

[42] This seems implied by his remarks at Strawson 1986, 38f.
[43] I owe this objection to Matt Talbert.

concept of freedom, but actions which are free in this thin sense are evidence for the existence of precisely that capacity for undetermined choice which the determinist or skeptic rejects in the full notion of freedom.

Strawson's response to this general line of argument is to invoke his own definition of rational action as action that is fully determined by reasons. Thus, he argues that the cases in question are essentially different from rational action precisely in virtue of their lack of rational explicability by reference to reasons. But this requires a conception of rational action that Reid explicitly rejects. That is, on Strawson's view rational action is to be explained under the description of being determined by the agent's reasons, or as a function of those reasons. Any action that cannot be given such an explanation is, by definition, not rational. On Strawson's view the same external action – for example paying with the particular coin – can be determined by reasons or not, but one has a fundamentally different explicandum in each case. That is, what has to be explained is that the agent acted for certain reasons. But it is unclear that one must accept this as the basic description under which the action is to be explained; Reid certainly does not, and it seems to strain common usage. On Reid's view one tells the same basic causal or determinative story for all actions, rational or otherwise. Reasons have an explanatory role to play when one asks about the moral content of an action – we will want to know, for instance, whether the agent acted against or in accord with her best judgment – but we need not interpret such reasons causally.

In a sense we might say that, from Reid's perspective, Strawson is running together answers to two different questions. One is, 'What caused the action?' Another is, 'How should we evaluate the action in a moral sense?'[44] Reid seems to think that imputability is involved in the answer to the first question. That is, an action is by its very nature imputable to its cause, so that the answer to the first question is *ipso facto* an imputation of the action to the agent who is the cause (whether this be the proximal agent or someone exerting an irresistible influence on that agent). But one does not thereby have an answer to the second question. An answer to the second question would involve the reasoning-process of the agent (or lack thereof) in addition to facts about the agent's history (education, etc.) and the nature of the action itself (i.e., whether it was in accord with or contrary to morality). For most instances of action evaluation the two interests behind the two ques-

[44] One can see these two questions as two different ways of interpreting Strawson's general question, "Why was the action performed?" (Strawson 1986, 33).

tions are inseparable. We ask 'Who did that?' when we have some moral interest in an action, and we ask 'Was it the right thing to do?' when we know who did what. When Reid discusses freedom and causation he usually has both questions in mind, but the cases considered here and the distinction between active power and moral liberty are reminders that the two interests are in principle distinguishable.

Here again we are at an impasse. Even if one adopts Strawson's view on action description, however, it still relies on his conception of rational explanation from reasons to action. This is a conception of explanation that Reid rejects.

Strawson's basic conception of action explanation is best articulated in the following passage:

> When we act, at a given time, the way we act is, *in some quite straightforward sense*, a function of the way we then are, in respect of character and motivation. We act as we act *because* of how we then are in respect of character and motivation. However one understands the 'because' in this statement there can be no serious dispute about its truth.[45]

Given that the process of decision and action is understood deterministically, the problem that faces the libertarian is explaining how the indeterministic elements in the process could have come about and how they could contribute to the freedom of the action.[46] The difficulty with this conception is that Strawson does not give an explanation of what "in some quite straightforward sense" means. I take it to be correlated with the way in which rational explanations can be simply true, and thus that the same difficulties apply. The important thing is to do justice to the intuition motivating Strawson's claim. This intuition, I take it, is that for rational action there must be some connection between reasons and action. A model of the explanatory role of reasons in rational action is therefore a model of this connection.

The model offered by Strawson seems to be drawn from physical causation. Reasons must be fully real in order to enter into equally real determinative or causal relations with actions that are real events. This would then motivate the view that volition or choice is to be understood in terms of conscious and explicit choice, since the different elements articulated by the explanation would be independent states or events and thus inherently perceivable as such by the agent. Despite its intuitive plausibility this model is at best underdeveloped. How is it pos-

---

[45] Strawson 1986, 312.
[46] Strawson 1986, 32.

sible that beliefs and desires could have this kind of effectiveness in the world? What explains the fact that they often do not have such effects? I am not suggesting that it would be impossible to attempt to answer such questions, but rather that the way in which actions can be determined by reasons is far from straightforward. The burden or proof shifts to the skeptic about free will to show that one must accept the view that all free choice must be understood as conscious and explicit choice. Until then it seems perfectly plausible for Reid to deny this claim, and thus to deny premises (4) and (6), and thereby block the regress.

A further complication of Strawson's view is that the explanation of even a physical system is never 'simply true' in the relevant sense. Even on a relatively realist view such explanations must be qualified by background constraints on measurement and observation, *ceteris paribus* clauses, and other factors that may not even be fully specifiable. Furthermore, it would be more than tendentious to claim that human action just is a physical system and nothing else. Reid correctly sees that this question-begging move is tempting precisely because the best arguments for the principle of sufficient reason (and therefore determinism) come from examples of mechanical systems (328). Of course, Reid denies that human action can be understood in mechanistic terms. This is, in fact, the crux of the dispute between the determinist and the libertarian.

Now Strawson would clearly reply that it is not enough to poke holes in his conception of explanation and account of the connection between reasons and actions – rather one must show how a positive libertarian account is possible. But important damage has already been done to Strawson's case against libertarianism. It is now clear that other models of the explanatory role of reasons may be offered; the functionalist model need not be the only model of rational explanation. And Reid implicitly gives us such a model in likening the influence of reasons to 'persuasion', 'advice', and 'exhortation' (283). At another point he compares reasons to lawyers arguing a case before a judge (288). Thus the suggestion seems to be that we understand free action on the model of social interaction, or at least on a kind of juridical model in which motives are like arguments.

Here is the place to address the lingering worry that even this conception of agency (and, implicitly, of action explanation) provides the basis for another skeptical attack on free will. In those cases where there are motives for action which fully explain the action, isn't it still true that we have no choice in the action? The response to this worry must insist on the more complicated, three-term structure of agency and action explanation that was discussed earlier and is part of this juridical model. The connection between reasons and action is the agent's *taking* of the reasons as justifying the action and understanding her action under the description offered by such justification. That is, ultimately the connection between reasons and actions is the

agent's free decision to act *on the basis of those reasons.* To put this in the terms I used earlier, the agent's moral freedom mediates between reasons and action in the order of the real just as our practice of moral evaluation mediates between reasons and action in the order of explanation. On this model, motives are satisfactory explanations only because we take the agent's causal power of rational decision for granted, and so it is elided in the superficial expression of the explanation. Such explanations are therefore compatible with alternative possibilities because they do not eliminate the possibility of different explanations had the agent decided differently. If the agent had chosen to act on different motives, those motives would also be explanatory in the same sense as the first. Crucially, they are explanatory because they allow us to understand how the agent conceived of the action, and not because they cite elements which necessitate or compel the action.

Of course, it may be true that after reflection by the agent one course of action stands out as the best, and which therefore has a kind of necessity from a subjective point of view. But to put the matter in a Kantian form, this is the necessity that the action *ought* to happen, not that it *will* happen. Furthermore, the phenomenon of weakness of will is good evidence that we are not necessitated or compelled to act on our best rational judgment. Although it may be true that akratic action is irrational, our capacity to so act is what makes us praiseworthy for acting rationally, on Reid's view. I admit that in order to respond to contemporary concerns I am developing Reid's view beyond what is explicitly given in his writings, but I think it is obvious from what Reid does say – especially regarding motives as *entia rationis* – that any action explanation must take this non-functionalist form. The contemporary determinist or naturalist will ask for a different kind of explanatory scheme, but given the usefulness of the broadly Reidian scheme I have been outlining here, the burden is clearly on the determinist to justify such a scheme. It will not do to insist, as Strawson does, that his functionalist scheme is the only one available.

One further point to make is the importance of not conflating explanations that silently presuppose our moral liberty with those that bypass it altogether (e.g., by citing mental illness or irresistible desires). The latter undermine free agency in particular instances because they show that our capacity for moral liberty was prevented from functioning properly and was therefore not part of the sequence of action production. The former make that capacity so central that is silently presupposed.

This point that the agent's moral freedom mediates between reasons and action in the order of the real just as our practice of moral evaluation mediates between reasons and action in the order of explanation may seem to be incompatible with the distinction between the thin freedom (active power alone) involved in unmotivated or insufficiently motivated action, and the thick freedom that combines active power with rational judgment with respect to actions that have potentially decisive reasons for or against them. In these cases of the thin freedom of under-motivated action, there would appear to be no moral evaluation to mediate between reasons and action in the order of explanation, since the choice at issue does not involve a moral contrast. But the case of action from insufficient or counterbalancing reasons is just a special case of action in which one uses the interpretation provided by moral evaluative practice even when there is no moral difference that attaches to the distinctions

between possible actions. One is responsible for the action, but that responsibility has no moral import because the action lacks moral significance. This juridical understanding of the connection between reasons and action seems basic to Reid's view of action, and seems to be independent of any particularly social aspects of his model of this connection. While this social or juridical model lacks the computable definiteness of the functionalist model, it has the comparative advantage of being drawn from a sphere which seems, at least pre-reflectively, to lie closer to the nature of action than basic physical systems. Furthermore, it is a type of activity in which we have a highly developed competence and in which the normativity of reasons is quite explicit. Thus it seems likely that such a model could be quite informative.[47]

This is clearly not the place to develop such a model, and it must be said that Reid's articulation of it is woefully inadequate. In fact, he thinks that the influence one can have over another is ultimately inscrutable:

> That a man may have great influence upon the voluntary determinations of other men, by means of education, example and persuasion, is a fact which must be granted, whether we adopt the system of liberty or necessity. How far such determinations ought to be imputed to the person who applied those means, how far to the person influenced by them, we know not, but God knows, and will judge righteously. (302)

But we need not follow Reid's homiletic instincts here. There is no shortage of intelligent and informative philosophical work on the nature of public argument and decision-making. And, unlike functions, this form of explanation may have a feature that is characteristic of most forms of action explanation, which is that there are better and worse explanations within each genre. For example, some purposes seem more closely related to certain actions and therefore offer better teleological explanations.[48] This gives the social/juridical conception a richness lacking in the functionalist model.[49]

Before going on to address potential future developments of Reid's theory in light of these arguments, it is important to note (for the record) that Reid is not content merely to shift the burden of proof back onto the skeptic. In addition to the objections discussed above, Reid makes the following further objections to determinism.

---

[47] I should note that this model does not seem to be intrinsically tied to the metaphysical conception of agent causation. Presumably, one could provide another interpretation of agent causation, and one could accept the juridical model without accepting agent causation.

[48] This model will not be able to offer a fully contrastive explanation in cases such as that in which there are no motives at all (or equal motives on either side). I think Reid accepts this failure as part of a healthy realism about the limits of our explanatory abilities.

[49] It might be thought that this talk of models of explanation makes our knowledge of actions non-objective, but I do not think this is accurate. The use of models means that our knowledge is not absolute or unqualified, but such knowledge would still be objective in a normative sense (i.e., some explanations and some models will be better than others and therefore binding on all rational inquirers).

First, it is unable to account for the existence of actions for which there is either no motive or equally good reasons on either side (285f.). Reid presses much the same point when he argues that there are some cases in which there appear to be motives only opposing the action. These are examples of "willfulness, caprice or obstinacy" (286f.). If motives determined actions these would be inexplicable. This relies on Reid's sense of the term 'motive' to be sure, for the determinist will likely reply that there is *some* mental state that determines the action, even if it does not look like a reason to the agent. Most importantly, the determinist must provide some non-circular explanation for the presence and determinative force of motives in such cases. The natural and common way of doing this is in terms of the strength of motives, and Reid finds this notion vacuous. If it is not to be uninformative (i.e., if it is not to be the case that the only reason one infers that X was the strongest motive is because the action was consonant with X and not Y) there must be independent criteria for the strength of motives. However, on Reid's view such criteria can only be derived from our experience of struggling to determine our will in the conflict between rational principles and our desires. But such conflict is only conceivable on the basis of our power to determine the will, and thus presupposes what was to be denied (288–290). Finally, Reid notes that, at best, reasoning from motives is probabilistic. One therefore cannot conclude that actions are determined by motives, but only that there is a kind of pattern on which we can normally rely. Such a pattern is perfectly consistent with freedom of the will (291f.).

Conclusion: Development of Reid's Position

The regress argument derives its force from the fact that it seems to show that, at the crucial moment, the libertarian has nothing further to say and that the basic structure of libertarian theory allows for no development of the conception of the crucial element (the agent-cause). Although Reid himself thought no further development of his theory was possible, we need not accept his judgment on this issue. Two related areas of further development are suggested by the preceding discussion.

First, it would be interesting to probe Reid's conception of the effort involved in determining one's will as a general framework for understanding the influence of motives. Here the social/juridical model would help to articulate the features of deliberation and decision in a way that did not reduce these to a decision procedure along functionalist lines.

Second, Reid sees the GFP's as formed around the time of our transition from childhood to adulthood, and he seems to suggest that a certain amount of education and support is required for the development of the rational capacities that are required for truly free action (299, 290, 87). As this suggests that agents who did not have that education

and support might not be responsible for their actions, it would be important to understand how education and example function both to develop our rational powers, and to provide the proper opportunities for the formation of our GFP's.

These suggestions for future development, and the nature of the models by which the theory can be developed, are made clearer by the confrontation of Reid's theory with the regress argument. Far from being disastrous, this confrontation instead suggests an informative research program for the potential development of libertarian theory.

Albritton, R. 1985. "Freedom of the Will and Freedom of Action". *Proceedings and Addresses of the American Philosophical Association* 59: 239–251.

Herman, B. 1993. *The Practice of Moral Judgment*. Cambridge, MA.

Hobbes, T. 1994. *The Elements of Law Natural and Politic and Human Nature*. Ed. J. C. A. Gaskin. New York.

–. 1999a. *Leviathan*. Ed. J. C. A. Gaskin. New York.

–. 1999b. "Of Liberty and Necessity". In *Hobbes and Bramhall on Liberty and Necessity*. Ed. V. Chappell. New York: 15–42.

Klima, G. 1993. "The Changing Role of *Entia Rationis* in Mediaeval Semantics and Ontology". *Synthese* 96: 25–58.

Lehrer, K. 1989. *Thomas Reid*. New York.

Leibniz, G. W. 1981. *New Essays on Human Understanding*. Ed. P. Remnant / J. Bennett. New York.

Locke, J. 1959. *An Essay Concerning Human Understanding*. Ed. A. C. Fraser. New York.

Madden, E. H. 1982. "Commonsense and Agency Theory". *Review of Metaphysics* 36: 319–341.

McDermid, D. 1999. "Thomas Reid on Moral Liberty and Common Sense". *British Journal for the History of Philosophy* 7: 275–303.

O'Connor, T. 2000. *Persons and Causes*. New York.

Reid, T. 1969a. *Essays on the Active Powers of the Human Mind*. Ed. B. Brody. Cambridge, MA.

–. 1969b. *Essays on the Intellectual Powers of Man*. Ed. B. Brody. Cambridge, MA.

–. 2001. "Of Power". Ed. J. Haldane. *The Philosophical Quarterly* 51 (no. 202): 3–12.

Rowe, W. 1991. *Thomas Reid on Freedom and Morality*. Ithaca.

Stecker, R. 1992. "Thomas Reid's Philosophy of Action". Philosophical Studies 66: 197–208.

Strawson, G. 1986. *Freedom and Belief*. Oxford.

–. 1995. "Libertarianism, Action, and Self-Determination". In *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. Ed. T. O'Connor. New York.

Weinstock, J. 1975. "Reid's Definition of Freedom". *Journal of the History of Philosophy* 13: 335–345.

Woozley, A. D. 1987. "Reid on Moral Liberty". *Monist* 70: 442–452.