

REASONS EXPLANATION AND AGENT CONTROL:  
IN SEARCH OF AN INTEGRATED ACCOUNT

Timothy O'Connor and John Churchill

Perhaps the central challenge for indeterministic (“libertarian”) accounts of human freedom is one of *integration*: squaring one’s understanding of an agent’s control over his own free action with a plausible account of how such actions are properly explained by the reasons the agent had for so acting. Two types of account predominate.<sup>1</sup> One is centered on the notion of agent causation. The other holds that a free action is the (event) causal, but nondeterministic outcome of antecedent factors including the states of the agent’s having reasons for so acting. Many philosophers judge that typical agent causal accounts of freedom improperly sacrifice the possibility of rational explanation of the action for the sake of securing control, while others judge that the reverse shortcoming plagues typical event causal accounts. (Of course, many philosophers make both these judgments.) After briefly rehearsing the reasons for these verdicts on the two traditional strategies, we undertake an extended examination of Randolph Clarke’s recent attempt to meet the challenge by proposing an original, “integrated agent-causal” account of human free action. We argue that Clarke’s account fails. In a final section, we sketch a more promising route to integration.

---

<sup>1</sup> We shall not consider here a third type, on which control is an intrinsic, noncausal feature of free choices or actions. For versions of this view, see Goetz, 1988, Ginet, 1990, McCann 1998, and Pink, 2004.

## I The Dilemma of Control vs. Explanation

The central intuition animating indeterministic approaches to freedom is that what we do freely is “up to us,” something that we initiate in a way that is not prefigured by the pattern of events leading up to the action. (We may allow that in some cases, my choosing of an action was free and yet determined by immediately antecedent events provided the fact that it was so determined was itself something over which I had control at some previous time. We shall ignore this kind of ‘derivative’ freedom in what follows and focus instead on cases in which an agent is ‘directly’ free.) It is thus an intuition that, when acting freely, we exercise an autonomous control over the selection of the action.

According to one view, this control is an ontologically basic kind of activity, a purposive, or 'agent', causality, distinct from the more ubiquitous mechanistic or 'event' causality. Taking a feature as a primitive is a reliable way to ensure that one's overall theory really does allow for the feature instead of offering a pale substitute in the manner of various implausible reductionisms. And the longstanding difficulty of giving a plausible analysis of our pretheoretical notion of autonomous control suggests that the gambit of primitive posit is not simply absurd. However, it may come at a very high price elsewhere in the theory. If this exercising of direct control over my own (free) action is itself uncaused, as agent causationists generally assume, then it can easily seem to itself be something that 'just happens.'<sup>2</sup> It is a 'fresh start' in the universe, uninfluenced by what

---

<sup>2</sup> Richard Taylor allowed that an agent-causal event could be caused, but held that it must be uncaused if it is to be free (1966, 109-115). More recently, Ned Markosian has

has gone before. What 'just happens' in this way is thus inexplicable—and in particular, it is inexplicable by the agent's reasons at the time for contemplating that course, whatever they may have been. If that is correct, we may be tempted to echo Leibniz's sentiment that "what is asserted is impossible, but if it came to pass it would be harmful."

This kind of consideration has been prominent in leading contemporary philosophers to reconsider the 'primitive, uncaused capacity' route to explicating human freedom. By putting the desideratum of explicability by the agent's reasons at the forefront of concern, we may be tempted to suppose that even free actions are caused by the appropriate reasons-bearing states of the agent at the time of the action. The agent's control over his action consists in just what the typical compatibilist supposes: the causal efficacy of certain of one's reasons. We secure autonomy by *weakening* this relation, supposing it to be nondeterministic: the agent might have performed a different action in identical circumstances because there was a non-zero (and perhaps pronounced) chance that other reasons-bearing states had been efficacious in producing the action which they indicated.<sup>3</sup> According to many critics, indeterminist event-causal approaches falter just

---

espoused a compatibilist view incorporating agent causation (1999, 257-277). But these are atypical. Agent causationists generally assume that an agent-causal event is essentially uncaused, or at least causally undetermined. (See, e.g., O'Connor 2000, 52-5.) More on this anon.

<sup>3</sup> The most prominent advocate of this avenue has been Robert Kane. Kane has developed his view in a number of writings culminating in *The Significance of Free Will* (1996).

here, in the fact that the autonomy of the control they posit is, as it were, secured by an *absence*, a removal of a condition (causal determination) suggested by the manifestly inadequate varieties of compatibilism. If there is no means by which I can take advantage of this slack in the flow of events, its presence can't confer a greater kind of control, one that *inter alia* grounds moral responsibility for the action and its consequences.

## II Clarke's 'Integrated' Agent-Causal Account

Such is the basic dialectic that leaves many incompatibilists seeking a third way. In *Libertarian Accounts of Free Will*, Randy Clarke pursues the path of *integration*, arguing that it is possible to combine the advantages of each of our two familiar approaches into a single, more complicated picture (135, 145-6). The way to do so is to suppose that directly free actions are caused *both* by the agent (*qua* substance cause) and by the event of the agent's having reasons so to act (for simplicity, 'reason R1' corresponding to 'action A1,' and so on). R1 is nondeterministic: it has an objective causal tendency of probability measure between 0 and 1. R2 also had a nonzero probability to cause A2, and perhaps there were other possibilities as well.

If we said nothing more, then it would merely be a happy coincidence that R1 and the agent both cause the same outcome, when they do. Clearly, for this picture to be plausible it is necessary that the activity of the agent is somehow bound up with the activity of his corresponding reasons. Clarke proposes that the tie is achieved via a

---

Since we are not giving anything like a full hearing to this approach to freedom, we will not discuss the ways in which Kane augments the bare bones indicated in the text.

fundamental law of nature. In the presence of a ‘live’ agent causal capacity, it is a law of nature that:

- i) whatever action is performed will be caused by the agent,
- ii) a reason R1 will cause an action A1 only if the agent causes it, and
- iii) the agent will cause an action only if a corresponding reason causes it.

We do not collapse clauses (ii) and (iii) into a biconditional since it is not clear that Clarke either wants or should say that in every case there is some reason such that *it* will cause action A1 if the agent does. Suppose a case where two reasons R1-A and R1-B both have a propensity to cause A1, and no other reason states of the agent have this propensity. And suppose that the agent in fact causes A1. Must it be the case that both R1-A and R1-B also caused A1, or might only one of them have done so? Clarke doesn’t say, and it’s not obvious to us that he *should* assume that both *potential* reasons for acting a certain way are actual reasons if either is.

As we have seen, Clarke intends for his integrated agent-causal account to serve as the best of both worlds: an approach to libertarian agency that offers what many take to be the most straightforward account of reason-explanation conjoined with robust control over the outcome of indeterministic processes. He does so by combining a causal model of reason-explanation, in which reasons are the causes of decisions, with agent causation. But there are several significant difficulties with Clarke’s account, difficulties that would likewise plague any approach that weds agent causation with certain features of event-causal libertarianism in the way Clarke has proposed. We enumerate these difficulties in turn below.

First: suppose I proposed to you that some entity,  $\alpha$ , causes events of type  $E1$ . Then suppose I explained further that  $\alpha$  causes these events only in conjunction with other causal events; that is,  $\alpha$  causes  $E1$ -events only when other causal events cause  $E1$ -events as well. So  $\alpha$  is never the only cause of  $E1$ -events. Further, the kind of  $E1$ -event caused in cases of this sort of  $\alpha$ -causation seems to vary primarily with variations in these other causal events, not with variations in  $\alpha$ . Given such an explanation, you may begin to find the idea of  $\alpha$ -causation suspect, at least with respect to these  $E1$ -events. Why, you may wonder, even posit  $\alpha$  as a cause at all in these scenarios? It is, after all, the other causal events that are doing all the explanatory work in the causal story I have given you.

So, astute thinker that you are, you decide to ask me for other examples in which  $\alpha$  exercises causal power. For you reason that if there are certain similarities between aspects of  $\alpha$ -causation of  $E1$ -events and other instances where  $\alpha$  exercises its causal power (say, to cause events of type  $E2$ , or  $E3$ , etc.), then you would have good reason to accept that  $\alpha$  was doing some causal work in causing the  $E1$ -events as well. This would give you better grounds for accepting the plausibility of  $\alpha$ -causation in the case of  $E1$ -events. But my response to this query, rather than helping my case, just further cements your doubt: I answer that  $\alpha$ -causation occurs only in the kinds of cases I have already described. In other words, the only effects  $\alpha$  ever causes are events of type  $E1$ , and  $\alpha$  causes these kinds of events only in conjunction with the other causal factors in the way I have described above.

At this point, it appears that you would have good grounds for rejecting the very idea of  $\alpha$ -causation. For it is much simpler to posit only the other causal factors as

causes, and it is prudent to avoid positing an instance of causation that does little to no explanatory work in the causal account.

But the description of  $\alpha$ -causation parallels Clarke's own theory regarding agent causation. In his account, the agent is never a lone cause of the effects she produces; changes in the effects caused (theoretically) directly by the agent are tracked only by changes in the *other* causal factors; and there are no other instances of agent causation apart from these cases of conjunctive causation. So if  $\alpha$ -causation is dubious, then so is Clarke's particular species of agent causation.

But suppose we waive this concern and grant *ex hypothesi* that free decisions are agent-caused in a way that corresponds to Clarke's integrated account. Clarke's libertarian agent-causal approach has some *prima facie* appeal because one may assume that in each case of agent causation, the agent's causing of a decision is prior, in some non-temporal but relevant sense, to the event causation of that same decision. Clarke denies (144) that the agent causation is *causally* prior to the event causation, and with good reason, for this would run counter to his robust causal account of reason-explanation. If there is any relevant priority, then, it would have to be *explanatory priority*. That is, the agent causation of the decision would have to somehow explain the event causation of that same decision in a satisfactory way. If the agent's causal exertion were explanatorily prior to the event causation, then we would have some reason to conclude that the agent is the ultimate *explanans* for the decision and a relevant *explanans* for the event causation of that decision.

But even a cursory investigation into Clarke’s position reveals that there can be no explanatory priority in this respect.<sup>4</sup> For explanatory priority—however it is ultimately articulated—seems to require asymmetry, and there is no asymmetry of this kind in favor of the agent within Clarke’s account. That is, we cannot say that an instance of event causation occurred *because* an instance of agent causation occurred in any stronger sense than we can say that an instance of agent causation occurred because an instance of event causation occurred. On Clarke’s account, the explanation goes both ways, with neither officially “trumping” the other. As such, we cannot hold agent causation of a decision to be explanatorily prior to the corresponding event causation of that same decision.

This lack of explanatory priority in the direction of agent cause to event cause seems to render Clarke’s response to a specific challenge far weaker than could be given by a more traditional agent-causal account. In one place, Clarke discusses an imagined “deterministic integrated [agent-causal] account,” wherein persons are posited as the agent causes of decisions that are also caused by deterministic event causes within a fully deterministic universe (163-4). Here is his criticism of this **imagined** position:

---

<sup>4</sup> See chapters 8-10 in Clarke, especially 170-176. See also the criticism below regarding a suggested implicit explanatory asymmetry in favor of event-causation within Clarke’s account.

Such an account allows that each instance of agent causation—if not itself causally determined by events that occurred long before the agent ever existed (and hence over which the agent has never had any control)—is made inevitable by these events as they causally determine the action that is agent-caused. In contrast, the libertarian integrated account, as has been said, provides in a straightforward sense for agents' originating or initiating their free actions. Because there is this absence [of any determining causes of the actions that are said to be free], on the libertarian version alone, we get origination, in a straightforward sense, of actions by their agents. (163)

What we find especially conspicuous here is the fact that on *neither* the deterministic integrated account nor Clarke's nondeterministic integrated account does the agent, in virtue of agent causation, ultimately and asymmetrically determine which action will be performed, for this kind of final determination would involve granting some *priority* to the agent as cause. Furthermore, desire-belief states are just as causally efficacious over the agent's action, and just as causally independent of the agent's causal activity, on one account as they are on the other. (What is more, in the deterministic case, it appears that the laws of nature will necessitate *independently of the law of agent causation* that the reasons will cause the action.) Since we find certain objections to any deterministic account to be cogent, we also find the similarity between these two accounts to be good evidence against Clarke's approach, his response notwithstanding.

Compare these results with the kind of response a more traditional agent-causal theorist would give if his account were placed alongside such a deterministic integrated

account in a similar challenge. This theorist claims that in the indeterministic scenario he envisions, the agent is the sole and ultimate producer of his or her actions, in stark contrast to the deterministic proposal. This contrast serves to differentiate the two kinds of accounts to a much greater degree than does the contrast between Clarke's approach and the deterministic account, rendering the traditional agent causationist's response much stronger than Clarke's "more straightforward origination" response to the same challenge.

As we have seen, the lack of any causal or explanatory asymmetry in favor of agent causation within Clarke's account removes much of the initial plausibility of his approach and forces him to give a comparatively weak response to the objection that his account does not sufficiently distinguish itself from a certain brand of determinism. We shall now argue that certain features of Clarke's account suggest a relevant kind of asymmetry in the *opposite* direction, in favor of the event-causal side of his theory. We have already noted his conclusion that actions must be event caused if they are to have adequate reason-explanations. In another place, Clarke appeals to event causes as the entities responsible for an agent's action or decision occurring *when* it does, an appeal that again prioritizes event causation when there is explanatory work to be done (197-9). More evidence of an implicit asymmetry is his assertion that an event may be an action in virtue of being caused solely by certain kinds of events, while *no* event is an action simply because it is agent-caused; rather, *agent-caused events are actions solely because they are even-caused by the appropriate kinds of events.*<sup>5</sup> Since, according to Clarke,

---

<sup>5</sup> See, for example, 136, 145, and 215.

actions are actions in virtue of being event caused by the relevant types of events, it follows that in a world where agents possessed an agent-causal power that did not work in tandem with event causation, the effects of agent causation would not be actions.<sup>6</sup> This is not a point to be taken lightly: Clarke's official position is that event causation is necessary, and in certain circumstances is sufficient, for action, while agent causation absent event causation is neither necessary nor sufficient for action. Thus, as in the case of both reason-explanation and the explanation of the timing of actions, Clarke's position as to what constitutes an action involves a prioritization of event causation over agent causation. Finally, consider again a point noted in our exposition of Clarke's account: there may be more than one reason favoring a course of action, each having some probability of causing it, while only some such reasons actually do contribute causally to its occurrence. For simplicity, we consider a case involving two such reasons, R1-A and R1-B. R1-A and the agent each are causes of A1, though it might have been that R1-B and the agent were causes of A1. Since the agent's causing A1 does not suffice for it's also being true that R1-A also caused A1, it cannot be that the latter depends upon the former. However, it also seems plausible here to assert the reverse dependency: in the

---

<sup>6</sup> It is open to Clarke to postulate that such a world is not possible, since agent causation by its very nature must be reason governed. This is an attractive position for the agent causationist to take (one that was emphatically endorsed by Thomas Reid, for example), but it doesn't seem to mesh well with Clarke's particular account of the way reasons explain actions, on which reasons and agents are independent, co-occurring factors. This position is better suited to the alternative agent-causal account considered below.

actual course of events, it was *RI-A* that led the agent to cause A1, though R1-B might have done so as well.

Clarke's overall approach to libertarian agency thus places such a fundamental emphasis on event causation by reasons that one may justifiably claim that there is an explanatory asymmetry in favor of that event-causal activity over the agent's causal contribution. But if the event-causal activity is prior in this sense, then Clarke's account does not succeed in the task for which it was created: such priority would render it ill-equipped to ground sufficient agent control and hence unable to overcome his objection to event-causal libertarian accounts.

The final difficulty we wish to articulate centers on the very foundation of Clarke's integrated account. On Clarke's model, the causal co-determination of a free decision by both the relevant agent and a proper event cause is no sporadic or contingent matter. Rather, as articulated at the beginning of this section, the two are linked together nomologically: there is a natural law stipulating that every time an agent directly causes a free decision, an event causes that very same decision.

What we find distressing here is that Clarke's conjectured law identifies a link between agent causation and event causation, but it gives not even a hint of an explanation for this link. Whence comes this lawful regularity in the simultaneous occurrence of an event cause with every exercise of agent-causal power? That every decision that is agent-caused must simultaneously be event-caused, with absolutely *no*

*causal relation* between the agent-causal activity and the event-causal activity, is *prima facie* mysterious. Moreover, he *cannot* stipulate a causal link between these two causes, for the resulting picture would undermine either his account of reason-explanation or his account of agent control, depending on which of the two causes was also stipulated to be a cause of the other. Considering the strong (nomological) tie that is supposed to hold between the two types of causation, this inability to offer a causal explanation of the regularity in question affords him little chance of devising any adequate explanation at all.

Let us sum up the difficulties enumerated above. First, given that (i) the agent is never a lone cause of his decisions, (ii) changes in the effects supposedly caused by the agent are tracked only by changes in event-causes, and (iii) there are no additional instances of agent causation apart from the theorized conjunctive instances, the very postulation of agent causation (like the postulation of our imagined  $\alpha$ -causation) seems suspect. Second, the fact that there is no causal or explanatory asymmetry in favor of the agent removes much of the *prima facie* plausibility from Clarke's account, since the agent is not deemed the sole ultimate *explanans* for his decisions in a way that best answers to Clarke's own objection to event-causal libertarianism. In addition, this lack of asymmetry affords him a rather weak response to the objection that his account is not sufficiently different from a certain species of determinism. Third, certain features of his account suggest a strong explanatory asymmetry in favor of the event-causal side of Clarke's model. But such prioritization of event causes over the agent would prevent his approach from achieving its goal of grounding sufficient agent control. And finally, the

nomologically necessary link between agent causation and event causation that is crucial to Clarke's account is not only unexplained, but—given his inability to appeal to a causal relation between the two relata—seemingly inexplicable. We find these reasons persuasive enough to warrant the rejection of Clarke's integrated agent-causal libertarianism.

### III Revisiting Traditional Agent-Causal Accounts

One might be tempted to conclude that any agent-causal libertarian account is doomed to failure. After all, if we must accept an event-causal approach to reason-explanation, and if the difficulties with combining such an approach with agent causation are as severe as has been argued above, then the prospects for agent-causal libertarianism do not look good.

Fortunately for friends of agent-causal approaches, however, it is not clear that we must accept an event-causal account of reason-explanation. Clarke best supports his claim that traditional agent-causal accounts do not secure an adequate *non*-causal account of reason-explanation by the following type of argument:<sup>7</sup> on more traditional agent-causal approaches, reasons are not the causes of free decisions; but if some desire-belief state *DB* were a cause of a free decision, then even if the non-causal conditions for reason-explanation were *not* met by *DB* in that scenario, *DB* would still qualify as a reason for that free decision; moreover, were *DB* a cause of a free decision and some

---

<sup>7</sup> Here we are presenting a generic and loose reconstruction of a line of argument found at Clarke (2003), 140-1.

independent reason *DB*’ met the non-causal conditions for reason-explanation, *DB* would and *DB*’ would not qualify as a reason; therefore, the account of reason-explanation involved in the more traditional agent-causal approaches is not adequate—it is neither necessary nor sufficient for reason-explanation—even in cases where there is no cause of the decision apart from the agent. This result leads Clarke to propose his integrated view of agent causation, in hopes of providing an adequate account of reason-explanation by appealing to reasons as event causes of the decisions they explain.

To assess this sort of argument, we need to consider a specific version of the ‘traditional’ agent-causal view. Let us focus on the following account.<sup>8</sup> An agent’s freely deciding to do action A consists in his causing the coming to be of an executive intention whose content has the form, *A-for-the-sake-of-G*, where G is the goal of some reason the agent had consciously in view at the time of the action. (Thus, we espouse a component, rather than a product, view of actions, including the basic action of forming a decision: the causing of an event is partly constitutive of the action.)<sup>9</sup> By “executive intention,” we mean intentions to immediately undertake the action, which intention normally will directly result in the initiation of the intended action. Our sort of agent causalist claims that such a decision and the corresponding action are explained by the agent’s having had the reason with content G, even though the decision and action were not causally

---

<sup>8</sup> O’Connor (2000), Ch.5, defends a similar view. It is modified in a way reflected here in O’Connor (2005).

<sup>9</sup> This contrasts with the product view reflected in our earlier exposition of Clarke’s agent-causal theory. For brief discussion, see Clarke 2003, 25 and 138-9.)

produced by it. The explanation here is grounded, not in an external causal relation, but in an internal, contentful relation between the reason and the intention the agent produces.

Now let us return to the scenario that Clarke believes to bring out the insufficiency of this non-causal account of reason-explanation of free decisions. We imagine there is a *DB* that causes a decision whose content matches not *DB* but *DB'*, an entirely different reason that did not causally affect the decision. Here, Clarke says, the reason that explains the decision is *DB* and *not DB'*. Clarke's judgment is plausible because we tend to think that causal explanations constitute the strongest explanations, and it is reasonable to count the *explanans* in the strongest plausible reason-explanation of some decision as the, or at least the dominant, reason for it. Therefore, given an example where there is a desire-belief causal explanation of some decision, we are inclined to endorse that causal explanation as providing the true reason for the decision. And since the decision is stipulated to be free, Clarke concludes that in the case just described the cause would count as a reason for a *free* decision.

But are we to interpret the example such that *DB* is a cause of the agent's *causing* (the coming to be of) the intention, or of the intention itself (in addition to and independently of the agent)? Assume first that *DB*'s effect is only the intention itself (and Clarke seems willing [141] to settle for this horn). In this case, it seems that what we have is a scenario in which there are two wholly distinct, unlinked, nondeterministic causes of a single event, the coming to be of an executive intention with a particular content. The scenario is bizarre, involving a wild coincidence of causes, but it does not

seem threatening to the traditional agent-causalists' view. It seems that what we should conclude is that the posited situation supports two wholly distinct explanations for that common effect, and the larger action which it results in. As *DB* has no causal relevance to the agent's causing the intention's coming to be, it cannot be a rival explanation to the explanation forged by the noncausal, internal link between *DB'* and the intention.

Let us then suppose that *DB* in Clarke's scenario is a cause of the complex event which we identified as the free decision: the agent's causing the coming to be of the intention.<sup>10</sup> An insufficient response would simply note that, in this case, what is caused is not a *free* decision, since the ultimacy of the agent's control is compromised, a product (albeit an indeterministic one) of other factors whose efficacy he does not control. This response is insufficient because Clarke can concede the claim while denying its relevance to his central contention. If we allow that in Clarke's scenario, *DB* is explanatorily relevant to the decision but *DB'* is not, then it will be implausible to say of an otherwise similar scenario minus the activity of *DB*, that *DB'* is now a perfectly good reason-explanation of the action as envisioned.

We believe there is a better reply to Clarke's scenario, one which denies that *DB* explains the (unfree) decision to the *exclusion of DB'*. Clarke's scenario is an unusual and complicated one, and this is reflected in its proper explanation, which, we suggest, is two-tiered. To fend off a charge of special pleading, we develop the point through a

---

<sup>10</sup> For our part, we find it hard to credit that it is *possible* for events to produce causal relations between other events. But we set aside this doubt here.

consideration of an analogous case involving the wholly event causal production of a decision. Suppose reason R2 causes R1's causing decision D1. Which reason explains the occurrence of D1, R1 or R2? Were this scenario possible (which it would be if Clarke's were), we'd say: both, in different ways. R1 is the immediate cause of D1 and thereby is explanatorily relevant to D1. But R2 provides a 'deeper' explanation, explaining why in this case R1 is causing D1 in the process of explaining the occurrence of D1.

We think the agent causalist can plausibly give a broadly similar reply to Clarke's case, at least if one is not antecedently disposed to reject the possibility of noncausal varieties of reason-explanation. In that case, the agent is acting for the sake of goal G, since he caused the executive intention with that very content in full awareness of the reason it matches. But in this case, there is also a deeper and different sort of reason-explanation of his doing so, namely the causal activity of the reason R2, a reason that, unlike R1, has no internal connection to G. Here, the agent is not *freely* deciding to A *solely* for the sake of G, as he has been caused to do so by one of his own reasons. Nonetheless, he *is* deciding to A *for the sake of G*.

We conclude that Clarke has not demonstrated the failure of the non-causal form of explanation on which certain traditional agent-causal theories rest. Yet we agree with Clarke that there *is* need of greater integration between the elements of agent control and reason-explanation within such theories, as we shall now explain.

#### IV An Alternative Approach to Integration

There are at least three shortcomings to an agent-causal theory that adopts the non-causal account of reason-explanation sketched above as the *whole* story on how reasons explain free decisions and actions.

- (i) A purely non-causal account of reason-explanation seems unable to illuminate apparent facts concerning the prior relative likelihood of an agent's choosing this or that.
- (ii) It is implausible that *every* free decision has the requisite content identifying the goal for which the agent is acting.
- (iii) It is plausible that reasons often influence free decisions in ways that are not consciously recognized.

We now briefly sketch a modified picture that, we believe, preserves the ultimacy of a free agent's control while better accommodating each of the above contentions.<sup>11</sup> Its central strategy is to identify a way that an agent's reasons may causally influence his free decisions without constituting, even in part, a condition that causally *produces* them. This requires our thinking of the agent-causal capacity as (like other causal capacities) something which is *structured* by external influences, with an evolving probability of being exercised in this way or that as a function of environmental factors. As we will suppose, the predominant determinants on the agent-causal capacity are the agent's own reasons (consciously grasped or not): corresponding to every possible course of action an agent might choose in a given episode there will be one or more reasons promoting the

choice of that action. These reasons differ in causal strength and their strength will possibly fluctuate in continuous fashion as an agent explicitly considers some such reasons, or deliberates in general, or simply has a relevant memory triggered by the chance observation of an environmental stimulus. All those reasons that directly contributed to the likelihood of an agent's A-ing are reasons *on* which the agent A-ed, when he does, though some could be negligible influences. Yet the agent causationist can and should insist that, though these reasons structured the agent's capacity and made him objectively likely to act as he did, in the end, nothing causally produced his decision.

We noted earlier that, where the content of an agent's decision identifies a goal for which the agent is acting, a goal that matches one of his prior reasons for so acting, there is an explanatory link between the reason and the action. If the agent freely and directly brings about an intention to A for the sake of G, then the agent is indeed motivated to attain G, and this fact is explained by the agent's having had a corresponding reason in view at the time of his decision. This purely teleological type of explanation is distinct from but consistent with the structuring-cause form of explanation advanced above. We may say that an agent acts *on* any reason which helps to structure his agent-causal capacity in a way that promotes the action chosen and that an agent acts *for* only those reasons that (in some cases) match the goal identified in the intention's content. We observe that while any reason of either type will help to explain the occurrence of a decision to A and subsequent A-ing, the latter type of reasons, when

---

<sup>11</sup> For the record, while we both find this kind of modified account to have promise, one of us is more confident in its tenability than is the other. For a fuller development of the picture sketched above, see O'Connor (2005).

available, uniquely explain the fact that the decision had the specific goal-directed content that it did.

We propose that these are some basic elements out of which a better integrationist theory of free will might be developed. We close with the following observation: how many of an agent's reasons he is consciously aware of; to what degree he is so aware; to what extent the reasons motivating his actions are ones whose presence and strength he himself has influenced through prior choices—all these seem relevant to an agent's freedom, and all come in degrees. Might not freedom of choice, then, itself be the sort of thing that admits of degrees?

*Indiana University*

**BIBLIOGRAPHY**

- Clarke, Randolph (2003). *Libertarian Accounts of Free Will*. New York: Oxford University Press.
- Ginet, Carl (1990). *On Action*. Cambridge: Cambridge University Press.
- Goetz, Stewart (1988). "A Noncausal Theory of Agency," *Philosophy and Phenomenological Research* 49, 303-16.
- Kane, Robert (1996). *The Significance of Free Will*. New York: Oxford University Press.
- Markosian, Ned (1999). "A Compatibilist Version of the Theory of Agent Causation," *Pacific Philosophical Quarterly* 80, 257-277.
- McCann, Hugh (1998). *The Works of Agency*. Ithaca: Cornell University Press.
- O'Connor, Timothy (2000). *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.
- (2005). "Freedom With a Human Face," *Midwest Studies in Philosophy* XXIX, 207-227.
- Thomas Pink (2004). *Free Will: A Very Short Introduction*. Oxford: Oxford University Press.
- Taylor, Richard (1966). *Action and Purpose*. Englewood Cliffs: Prentice Hall, 109-115.