

## Hegel and Analytic Philosophy of Action

Christopher Yeomans  
*Purdue University*

*Abstract:* A primary fault line in the analytic philosophy of action is the debate between causal/Davidsonian and interpretivist/Anscombian theories of action. The fundamental problem of the former is producing a criterion for distinguishing intentional from non-intentional causal chains; the fundamental problem of the latter is producing an account of the relation between reasons and actions that is represented by the 'because' in the claim that the agent acted because she had the reason. It is argued that Hegel's conception of teleology can be used to develop the interpretivist position by solving both its and the causal theory's fundamental problems.

The analytic philosophy of action begins, as one might expect, with an attempt to delineate its subject matter. Treatments of the field often start either with Wittgenstein's formulation—"what is left over if I subtract the fact that my arm goes up from the fact that I raised my arm?"<sup>1</sup>—or with a series of examples designed to motivate the difference between action and mere behavior—e.g. my raising my arm to catch the attention of my dinner date as opposed to raising it through a muscle spasm or it being forced up accidentally by a clumsy waiter. There is widespread agreement within this tradition that the remainder in the first formulation or the differentiating element in the second is an intention of the agent with respect to the action.<sup>2</sup> Here, however, agreement ends, and a variety of formulations of what it is to be an intention—or what it means to act for reasons—have been proposed. The primary fault line in this debate separates those that advocate a causal relation between intention and action—a position most often traced to Donald Davidson—and those that deny causality in favor of what is often phrased as an 'interpretive' relation between the two—a position most often traced to Elizabeth Anscombe.

cur”—that is, to use agent-causation talk radically qualified by ‘as if.’”<sup>7</sup> So both the causal and the non-causal theorist must spell out in some informal detail the nature of the productivity that characterizes action.

Anscombe makes a first attempt in claiming that intention is characterized, at least in part, by being a kind of ‘knowledge without observation.’<sup>8</sup> This notion has puzzled many, and Anscombe herself has difficulty clearly articulating it. But there is a different element in Anscombe’s view that appears more promising as a response to the challenge of a positive account of teleology that pokes its head through the dense discussions of *Intention* at its joints. First, Anscombe claims that “The primitive sign of wanting is trying to get” (§36), which suggests a teleological criterion of striving or trying in terms of intentional attitudes rather than the more negative ‘knowledge without observation.’ More importantly, Anscombe has a wonderfully suggestive treatment of the relation between different intentional descriptions of action:

Are we to say that the man who (intentionally) moves his arm, operates the pump, replenishes the water supply, poisons the inhabitants, is performing four actions? Or only one? The answer that we imagined to the question ‘Why?’ brings it out that the four descriptions form a series, A-B-C-D, in which each description is introduced as dependent on the previous one, though independent of the following one. . . . [I]f we say there are four actions, we shall find that the only action that B consists in here is A; and so on. Only, more circumstances are required for A to be B than for A just to be A. And far more circumstances for A to be D, than for A to be B. . . . So there is one action with four descriptions, each dependent on wider circumstances, and each related to the next as description of means to end. . . . When terms are related in this fashion, they constitute a series of means, the last term of which is, just by being given as the last, so far treated as end.<sup>9</sup>

When I explain why the agent acted, my explanation therefore tracks a series of nested goals, each of which brings with it more context through the way in which it is connected with the previous goal. So explanation by D (‘He moved his arm in order to poison the inhabitants’—or ‘He moved his arm because he wanted to poison the inhabitants’) is the most complete and satisfying of the series in two ways: first, it is the final goal of the series, and so throws the most light on the internal structure of the series by identifying the process in which all four are connected parts; and second, in so doing it implicitly brings more information about the outside world into our understanding of the action. Now, Anscombe does not fill out this schema in *Intention*, and I will not attempt here to read the rest of that work in teleological terms. One of the most fascinating things about this formulation is the way in which

teleological explanations are seen to illuminate both internal, intentional structure and external contextualization in the same move. As a model of action explanation, this points explicitly to a notion of agency as *embedded self-directedness*.<sup>10</sup>

This point in Anscombe—which she leaves undeveloped—is a perfect place to take up with Hegel. Hegel clearly understands agency in teleological terms,<sup>11</sup> but to understand this element of his practical philosophy we need to look at his discussions of mechanism and teleology in the *Logic*. If a teleological account can be given of this production via animation by reasons, then it should respond to the challenge to provide a positive account of the relation underlying the ‘because’ in reasons explanation. Furthermore, if the account could give a teleological specification of the kinds of processes in virtue of which the goals or intentional descriptions of an action are related as means to end, it appears that it could also tell us something about distinguishing the right kinds of causal chains in the context of intentional descriptions. My primary interpretive thesis is that we should see Hegel as providing resources for such an account, and therefore as having much to say in the context of contemporary analytic philosophy of action.

We can best connect Hegel’s thoughts on mechanisms and teleology to this tradition through another thought of Anscombe’s. In discussing the relevance of causal factors in our understanding of human relationships, Anscombe notes that,

These causalities are mostly to be understood derivatively. The derivation is from the understanding of action as intentional, calculated, voluntary, impulsive, involuntary, reluctant, concessive, passionate etc. The first thing we know, upon the whole is what proceedings are parleys, agreements, quarrels, struggles, embassies, wars, pressures, pursuits of given ends, routines, institutional practices of all sorts. That is to say: in our descriptions of their histories, we apply such conceptions of what people are engaged in. In the context of such application, then, the causalities to which we ascribe such events can, so to speak, get a foothold. Given the idea of an engagement to marry, say, you can look for its causal antecedents.<sup>12</sup>

The multiplicity of description here is both a source of the power of the point being made and a reason to doubt whether we are increasing our understanding if we accept it. As Davidson points out, the very idea of interpreting an action as intentional through placing it in a pattern does not, in itself, advance our understanding of the nature of action unless we know something about that pattern and what differentiates it from others (after all, causation is a pattern as well).<sup>13</sup> But Hegel makes a similar priority claim for teleology

opposed to causal mechanism, in that he argues that causal mechanisms can only be individuated by the purposes that explain their functioning,<sup>14</sup> and the specificity of Hegel's conception holds out the promise of refining Anscombe's insight into a positive account of the productivity of agency.

Given the constraints of space, I will attempt neither to reconstruct nor to defend Hegel's argument for this priority.<sup>15</sup> Instead, I want to focus on the relation between ends and mechanical causes that develops through Hegel's argument, and in particular Hegel's claim that causal mechanical interactions serve as the means or techniques by which ends produce and thus manifest themselves.<sup>16</sup>

With respect to this analytic background, one of the most interesting things about Hegel's discussion is the way in which he interprets the causal theorist's appeal to law as a way of individuating mechanisms.<sup>17</sup> Here, Hegel is specifically thinking in terms of a kind of mechanical system that would have a law that regulated its interactions with other systems—and regulated the interactions of its parts with each another—via a certain state of equilibrium that the system tended to maintain.<sup>18</sup> The systems Hegel considers are therefore homeostatic, and such systems have been thought by analytic philosophers to present a promising model for the reduction of teleological relations to the causal feedback mechanisms that are thought to underlie them. In fact, Davidson explicitly associates the causal power of free action with just such types of systems.<sup>19</sup>

But on Hegel's view, because the law is no longer merely an arrangement of elements but rather specifies an ideal state towards which the system tends, it is not necessarily actualized within the system individuated by that law, nor is the achievement of the state necessary to individuate the system.<sup>20</sup> Therefore, within the explanatory scheme of mechanism, which, in Hegel's conception of it, demands existent objects as *explananda*, the law cannot be a proper explanation of the genesis or behavior of mechanistic systems, and therefore cannot be used to individuate such systems.<sup>21</sup> By contrast, goal-directed behavior is trying to achieve the goal, and this process is regulated by the goal (behaviors are selected which tend to achieve the goal, and are modified in such a way that they tend to achieve the goal even in the face of obstacles or changing circumstances). So the end governs the process of its own individual realization in a way that mechanistic principles do not, because ends themselves change through that influence in a way that laws do not, they are actualized in a more complete way: laws are instantiated but

ends are realized. The reciprocal interaction of end and behavior in trying makes the end an *existent* and *immanent* organizing principle in a way that laws are not, on Hegel's view.

In the contemporary literature, Larry Wright has argued for the irreducibility of teleology in similar terms:

Appropriate-but-unsuccessful behavior may well be the most central kind of teleological behavior, both conceptually and identificatorily; for it is the behavior of trying. And not only is trying one of the most emphatically teleological concepts—but trying-behavior constitutes the majority of that systematically complex behavior we are most reliable in identifying as teleological. The clearest cases of hunting, fleeing, and building consist largely of attempts—success is quite usually elusive. . . . What makes us say a predator is stalking—rather than writhing or undergoing spasms—is the systematic organization of the movements about the goal-object, or about the obvious clues to the goal-object, or about something that might be mistaken for a clue. It is this systematicity that makes the *direction* of the behavior so obvious. And the particular systematicity that gives direction to a bit of behavior is that which obtains when that behavior arises because it tends to produce a certain result. . . . [T]he behavior was plastic and persistent with respect to that result.<sup>22</sup>

Wright argues that the fact that such systematic plasticity and persistence entail the possibility of a wide range of causal processes that would count as such trying means that it is hopeless to attempt to reduce the teleological relation to an underlying efficient causal one.<sup>23</sup> Goal-directed behavior is identified instead by the pattern of the activity's responsiveness to the goal and the resources for its achievement in the environment. The *trying* is the objective phenomenon, that which is clear to all who have eyes to see.

Hegel's version of this claim is phrased in terms of a conception of self-determination as "elasticity" as opposed to the passivity involved in causal interaction, a fascinating conception whose surface I can only scratch here.<sup>24</sup> In a passage that is immensely suggestive of the import of this discussion for the question of agency, Hegel writes of mechanism as a technique for goal-directed self-determination:

[For mechanism and chemism], end is the self-determiner [*Selbstbestimmende*] that brings back into the unity of the concept the external determinedness by which it is conditioned. From this can be seen the nature of the subordination of the two previous forms of the objective process. . . . Thus mechanical or chemical technique, through its character of being externally determined, offers itself spontaneously to the end relation.<sup>25</sup>

This idea receives its most detailed formulation within Hegel's treatment of teleology through the claim that the means itself is the most significant and

complete realization of the end. In developing this thought, Hegel contrasts the way in which the interpenetration of ends and objectivity in the means contrasts with the achievement of finite, fleeting ends such as the satisfaction of desires:

Further, since the end is finite it has a finite content; accordingly it is not absolute, nor simply something that in its own nature is *rational*. But the *means* is the external middle term of the syllogism which is the realization of the end; in the means, therefore, the rationality in it manifests itself as such by maintaining itself in this external other, and precisely through this externality. To this extent the *means* is *superior* to the *finite* ends of *external* purposiveness: the *plough* is more honourable than are immediately the enjoyments procured by it and which are ends. The *tool* lasts, while the immediate enjoyments pass away and are forgotten. In his tools man possesses power over external nature, even though in respect of his ends he is, on the contrary, subject to it.<sup>26</sup>

Hegel's idea here is that in using a means for an end—especially repeatedly—we change the means in such a way that it comes to represent more completely the particular end for which it was the means. In fact, Hegel thinks that through successful use of a means we can in fact make an initially external purpose into an internal purpose. When a baseball player breaks in a glove, for example, he makes it into an object that has his own playing style and form and shape as its immanent principle. After a season, the glove comes to present his fielding in a way that perhaps no particular play could. This is true even though the end itself is necessarily subject to external influences, for example, in the way that the desire to be a good fielder might be inculcated by a parent or community. Our use of the techniques of the means is a way in which we assert ourselves in the face of that influence, a way in which we manage that influence. The *means* is therefore Hegel's way of talking about acting as objective and substantial.

In Hegelian terms, the means is a way in which the reflection-into-self of goal-directed self-determining is elasticity, and so it is not surprising that Hegel holds of the means that in it, the end's "reflection-into-self is itself also self-external and a reflection outwards."<sup>27</sup> That is, the means is a particular way in which the goal-directed system opens itself up to external influence by selecting and directing that influence, a way in which it is active towards the force of that influence rather than merely passive. In a remarkable passage that illustrates this function of the means, Hegel writes,

that the end posits itself in a *mediate* relation with the object and *interposes* another object *between* itself and it, may be regarded as the *cunning* of reason. . . . [The end] puts forward an object as means, allows it to wear itself out in

its stead, exposes it to attrition and shields itself behind it from mechanical violence.<sup>28</sup>

It is through the means, therefore, that self-determination is protected from the alienation of the mechanical relationship of cause and effect. But more to the current point, the end represents the constancy of systematic orientation that is maintained behind but also through the manifold and changing causal interactions that constitute the goal-directed behavior in the means.

On the one hand, this point seems to suggest that some causal theorists may be correct in thinking that they need not provide a specific criterion for distinguishing intentional from deviant causal chains, since the intrinsic plasticity with respect to the processes employed in goal-directed self-determination entails a diversity of such processes. But on the other hand, what lets the causal theorist off the hook is, on Hegel's view, a fundamentally distinct and irreducible conception of productivity. To see this, consider another version of the example of the waiter who wants to startle his employer. One can imagine a situation in which the waiter feels his nervousness coming on, and recognizes it as a technique that might be exploited to do what he might otherwise not have the nerve to do. He refrains from controlling his nervousness and instead lets it play out next to the stack of glasses. The glasses are spilled and the employer is startled. In that case, I think we would say that the waiter intentionally startled his employer; even though the causal chain is non-standard, it is not deviant in the relevant sense. What accounts for the difference between this case and Wilson's case, in which the nervousness is a deviant chain? Locating the difference in an *additional* causal chain of a certain sort is unlikely to help (e.g. control from some higher-order mental state), because that chain as well can either be employed or not by the goal-directed activity in question. But this is the key to the answer: the waiter in this second example acts intentionally because the control of the intention is represented by the plasticity with which the realization of that intention is pursued, and the way in which new means are recognized when they present themselves.<sup>29</sup> That is, the plasticity is a way of talking about the way in which the purpose illuminates the context and presents processes as means; this is the cunning of reason. And notice that it does so by introducing the interaction of goals and context in the way that Anscombe's account of the relation between intentional descriptions suggested. This suggests that Anscombe's myriad patterns for the relation of intentions to actions might be understood as variations on the basic form of the teleological relation.

A complaint about a contemporary teleological account can help to further bring out the interest of Hegel's account here. In commenting on George Wilson's teleological account of agency, Abraham Roth complains that teleological considerations explain only general facts, not specific events. But the action explanations that we make appear to be explanations of specific actions, and these actions track an objective relation between reasons and actions that constitute those actions as intentional. Causation is a singular relation, and so the relation between reasons and actions must be causal.<sup>30</sup> Now, Hegel does not share these intuitions about either teleology or causation. In fact, Hegel holds that teleology promises a better articulation of the singular individuality of its *relata* than mechanism does, and it is interesting to see why.<sup>31</sup>

Hegel holds that one distinctive feature of teleology is the fact that the reciprocal interaction between end and means results in the modification of the end in the process of its realization. Hegel thinks that the necessity of the end's own transformation follows from its role as the immanent organizing principle of the means:

Now insofar as end is this total reflection of objectivity into itself and is so immediately, in the first place, the self-determination or particularity as simple reflection into self is distinct from the concrete form, and is a determinate content. From this side end is finite, although in respect of its form it is infinite subjectivity. Secondly, because its determinateness has the form of objective indifference, it has the shape of a presupposition, and from this side its finitude consists of being confronted by an objective, mechanical and chemical world to which its activity relates itself as to something already there; its self-determining activity is thus, in its identity, immediately external to itself and as much reflection outwards as reflection into self. . . . Accordingly, the movement of end can now be expressed as having for its aim to supersede its presupposition, that is the immediacy of the object, and to posit the object as determined by the concept. . . . Positively, [this process] is the realization of the end.<sup>32</sup>

Here, the end confronts conditions that it presupposes in the form of different mechanical systems lying about in the context of expression. But it constitutes that context by forming and sorting those mechanisms according to its own ends; the context is constituted by the point (*Sache*) for which it is the context.<sup>33</sup> But precisely because this forming and sorting takes place as a presupposing, it takes the form of the initial establishment of context that is, to an important extent, invisible still from the perspective of the end itself. In human agency, we do not normally recognize the way in which our goals determine our context, since our attention is more naturally directed

to changing that context to achieve our goals, but often this feature is very visible to others. For example, when we say of someone that he has a 'chip on his shoulder,' part of what we mean is that he is looking to constitute external events as insults to which he can respond, and even though he does not realize this fact, it is clear to others who then try to avoid saying or doing anything that could be construed as a slight to him. This is a classic case in which the non-intentional background of intentional behavior is 'behind the back,' and is perhaps revealed to the agent herself only by breakdowns such as alienating contexts (e.g. oppressive factory work) where the dominance of one agent's goals in constituting the context results in a failure of agency due to other agents' inability to form their context in this way. In these breakdowns, our subconscious strategies for making the world into the context of our actions are revealed precisely through their unexpected ineffectiveness. Hegel, as one might expect, holds that the true teleological relation must make this forming and sorting explicit. That is, it must show that forming and sorting to be a positing—an active contribution of the agent—rather than merely a presupposing—a passive receptivity to a given fact. Thus a theoretical problem for the notion of expression is made into a practical problem for individual expressers (whether agents or single-celled organisms). Or, to put it in terms of Hegelian phenomenology, the implicit contribution of the subject's self-understanding is thematized and thus made into part of the problem that the action is attempting to solve. In so doing, the problem becomes not a form of skepticism to be refuted but a potentially alienating element that is to be managed.

On Hegel's view, if the end is to be truly immanent, it must take on its full character in the processes that constitute the means as an object; these are its activity, the process of its objective realization. Even though the end is to be conceived as directing the process of engagement with its context, the process nonetheless exposes the end to external conditions. These conditions create an intelligible distinction between an internal, basic character of the end and the more particular form of its external realization; they are the way in which the teleological idea differentiates itself from its own expression. This happens only in the engagement of the means by the end, but may easily generate the illusion that there must be some prior end acting mechanically on the means. It is precisely to combat this illusion that Hegel differentiates teleology from force and cause:

The end [*Zweck*] is therefore the subjective concept as essential striving and drive to posit itself externally. . . . End may indeed also be defined as force and

cause, but these expressions fulfill only an incomplete side of its significance; if they are to be predicated of it as it truly is, they can be predicated only in a way that supersedes their concept: as a force that solicits itself to externalization, as a cause, which is the cause of itself or whose effect is immediately cause.<sup>34</sup>

Teleological production is therefore quite clearly intended by Hegel to be a replacement for the notion of a self-cause that has played such an important role in understandings of agency. But more to the point, the intensive interaction between goal and causal context effectively rebuts the charge that purposes can only explain general facts instead of specific events, since the explanatory purposes evolve to be quite specifically related to the action they explain during the process of its production. The problem of explaining particular actions arises only if we treat such purposes as what Hegel would term abstract universals. But on Hegel's view, this is a conception more appropriate to the covering laws of purportedly singular causal connections than to internal purposiveness.

This point is important for understanding the differences between causation and teleological productivity as a model of action. One might think that long-term goals would be best represented as abstract universals, e.g. survival or reproduction, and thus that the abstract/singular distinction was independent of the causal law/goal-directedness distinction. But the explanations of specific actions—and *a fortiori* of long-term projects—will be more satisfying to the extent that, e.g. survival in a particular way as a particular kind of person is invoked as an explanation, and this points to the fact that the realization of such general goals modifies and specifies them such that, at best, the general term is just a shorthand reference to the more specific form of the goal. But this recursive transformation of the goal in context makes the unilinear picture of causation under covering laws inapt as a general description of the kind of productivity at issue. If Hegel is right about this, then goal-directed productivity cannot be articulated as any specific and distinct kind of causal process, and on this point Hegel has common cause with Wright in rejecting the reduction of teleology to efficient causal mechanisms.<sup>35</sup> For Hegel, the important contrast is the local, minimally diachronic causal linkage as opposed to the global, more fully diachronic systematic orientation of the pattern of causal linkages. The systematic orientation of the system itself changes through its realizations of the goal (and the frustration of attempts at realization)—this is why the plasticity of the system is an articulation of this recursiveness—but this change cannot necessarily be represented in the abstract as adoption of one specific kind of causal chain as opposed to an-

other (though such a change may be involved if such a change would be an appropriate technique for the reorientation of the system).

To realize our ends, we use our bodies as means in mechanical relations to other objects. But this action on the world has a reaction on us and changes us, and in turn modifies our ends. As this recursive process continues the end and the means are further and further integrated into each other, so that what begins in infancy as a relatively external relation between the body and purposes can develop into a tighter fit. The case in which one's end is realized exactly as planned is just a limit case of this schema (both in theory and in practice). Even when the end is perfectly realized, it is still changed in the sense that its realization puts it into connection with a host of conditions and events surrounding the action. If the plan is perfect, then this just means that the agent completely put the end into relation with such conditions in the planning process, thus transforming the end internally in anticipation of the resources and contingencies that would be encountered in the external world. But such a perfectly planned and realized end is a limit case—not the paradigm. The paradigm cases are those in which we modify our ends both beforehand in anticipation of the context of externalization, and then subsequently as our knowledge or control of that context turns out to be imperfect. I plan, for example, an itinerary for a family trip taking into account the usual weather for this time of year and my children's current nap schedules, but then this must be adjusted on the fly if the weather changes drastically or my toddler sleeps poorly the night before. This reciprocal interaction between goal and context can go either way, of course. If the goal changes much more than the context, we call this resignation or inauthenticity, rather than persistence.

In this way, Hegel's view has important advantages over Rowland Stout's contemporary teleological view of agency, which similarly holds that the deviant causal chains are deviant because they eliminate teleological processes by introducing an element that is not sensitive to the means-end relation. In Stout's theory, causal chains which can be broken down into component events are distinguished from causal processes that, though they require "a certain kind of structure of stages to occur," nonetheless, "at any one moment when a process is happening, what is happening is the whole process, not just part of it."<sup>36</sup> Stout then defines a 'mechanism' as "a persisting, self-standing, entity within the underlying conditions of a process . . . that . . . has the capacity either to undergo or to cause the characteristic stages of that process."<sup>37</sup> To deal with cases of deviant causal chains, Stout wants to "include the condi-

tions that constitute non-interference as part of the underlying nature of the process,” so that it is clear that the interference present in the deviant chain cases ends the relevant teleological process and thus erases agency from the event.<sup>38</sup> But this inclusion of a non-interference condition within the essential properties of the process is implausible both because of the necessity of interference or resistance in *some* form or other,<sup>39</sup> and because it therefore appears to rule out the need for the plasticity and persistence that is central to any teleological process. This explains why plasticity and persistence fall out of Stout’s account in favor of a less substantive *sensitivity* to the goal. The former, however, is a messier but more informative and intuitive account of teleological processes. So, for example, Stout is forced to distinguish between the process of the sun warming up the stone (if there is nothing to interfere) and the process of the sun stopping the stone from cooling down (if a cooling wind interferes with the former process) But it is deeply counterintuitive to individuate processes in this way; given the omnipresence of interfering conditions in the real world, this method of individuation would multiply basic processes beyond all reason.

The basic problems that Stout faces concern the relation between structure and its underlying conditions; in Hegelian terms, between an essence and its appearance, or between the idea and the context of its expression. On the one hand, Stout has the insight that interference can change the nature of a process. But on the other hand, his concern to establish the continuous presence of an underlying condition leads him to argue that such a change results in a completely new process.<sup>40</sup> As a result, the elasticity, in Hegel’s sense, of goal-oriented activity is effectively denied. This also comes out in the way that events happening during a process are ruled out for that reason by Stout as inputs to the process, which is another deeply counterintuitive consequence of his characterization of processes.<sup>41</sup> In addition, by Stout’s own admission, his insistence on the *self-standing* nature of the mechanism means that the capacities that it grounds are understood much more broadly than is normal (e.g. that not only the watch but also the spring within it has the capacity to tell the time). This is essentially the problem of the arbitrariness of individuation that Hegel holds follows from the false (because rigid) independence of mechanisms. If Hegel is right that the end that specifies the basic character of a teleological process and the means that represents the developed articulation of that character reciprocally interact with each other, then he has a theory that can, in principle, articulate the individuality of a

mechanism without either of the two counterintuitive results of Stout’s view (the fragility of processes in the face of interference and the over-extension of the concept of mechanism), both of which follow from the rigidity of independence on Stout’s conception.

Hegel’s avoidance of the first result is just his inclusion of plasticity and persistence in the nature of a teleological process, as described above. It is only in those cases in which the interference overwhelms the capacity of the teleological process to maintain its goal-orientation that we should think of the teleological process as ending, and thus, in the case of agency, the causal chain as being deviant rather than intentional. So, for example, in Wilson’s original case of the waiter who accidentally startles his employer because he becomes nervous at the thought of doing so, the fuller picture of the waiter’s psychology will be dominated by causal chains that tend to either avoid the result or bring it about in a different way at a different time. The causal chain that extends from the desire to startle to the startling through the nervousness is therefore deviant both in the sense that it contrasts with the preponderance of the agent’s orientation and in the sense that nervousness that is the key link in the chain is actually a byproduct of the contrary desires to avoid the startling or to execute it at another time. In the modified example, the agent recognizes this function of the nervousness and then the preponderance of his psychology would be directed at non-interference with this mechanism or indirect nurturing of it (fanning the flames of his nervousness, to mix metaphors). Of course, the possibility of self-deception muddies the waters here, but the criterion for distinguishing deviant from standard intentional chains is the systematic orientation of the agent’s intentional stance, including those causal chains that are not directly productive of the action. The same point can be made diachronically. In the first case, we can imagine the agent recognizing his nervousness and mobilizing resources to combat it (breathe deeply, think happy thoughts), where these attempts to stop the nervousness are ineffective or backfire. In the second case, that recognition is followed by the opposite response, which is further develop that nervousness.

We can see Hegel’s avoidance of the second problem (the over-extension of the concept of mechanism) by using these ideas to address the example of the watch. On Hegel’s view, we ought to individuate the watch as a whole as the mechanism to tell time rather than just the spring because the watch as a whole includes means to adjust the functioning of the watch to maintain correct time in the face of variable circumstances, whereas the spring itself

does not. So the watch as a whole is plastic and persistent with respect to the goal of telling time if the functioning of the watch modifies itself to track the goal of telling correct time, or it is individuated as the key physical artifact of the human process of telling time by virtue of the way that it affords the wearer the opportunity to make such adjustments herself. In the first case, the watch tells the time, and in the second we tell the time by means of the watch.

Hegel's claim that the end is necessarily modified in the process of its own development is the direct result of a continuous line of argument regarding reflection as expression that begins in the logic of reflection and continues through the subjective logic, which argument is oriented by the problem of arbitrariness in individuation, where individuation is argued to be parasitic on the general principles that explain the behavior of the individual. From this perspective, Hegel claims that causal laws are instantiated by particulars in virtue of their falling more or less arbitrarily under certain general kinds.<sup>42</sup> Goal-directed activity does not leave it to external reflection to decide its fate; rather, it sets about realizing its own goal. However, it does not do so just as it pleases, but under circumstances existing already, given and transmitted from its past and environment.<sup>43</sup>

If this is right, then it makes sense to formulate a Hegelian teleological account that sees agency in the plastic and persistent goal-directed reorientation of causal chains, as represented by the dynamic pattern of those chains. Though it is important to Hegel's own modal theory that possibilities be represented as actualities in the context of action, and thus that the distinctions in responses to circumstances that distinguish goal-oriented processes from causal ones, and one goal-oriented process from another, are primarily conceived globally and diachronically as the differential response to present but diverse and changing conditions, Hegel's general framework nonetheless provides a constructive way of thinking about these possibilities from the analytic perspective, which emphasizes counterfactuals as a way of bringing in the significance of alternate possibilities. This point can be developed by considering another form of Roth's complaint about the generality of teleology, which is his claim that, unlike causation, explanation by purposes cannot support singular counterfactuals:

If . . . Lisa went for a bike ride because she wanted to exercise, then (all things being equal) she would not have gone for a ride had she not wanted exercise. . . . That particular bicycle ride depended upon Lisa desiring in the way that she did at that time. If this is not to be a bare counterfactual, we need to tell some story of how the reason is actually related to the action so that the

counterfactual receives some support. To provide such a categorical basis for the counterfactual, the Causalist may simply appeal to the causal relation between attitudinal reason and action.<sup>44</sup>

But, Roth argues, the teleological relation will not support such specific counterfactuals. It strikes me that from a Hegelian perspective, there are at least two difficulties with this argument. The first is that it seems incredible that simply appealing to the idea of a causal relation as support for a counterfactual gives any insight into that counterfactual or leaves it any less "bare" than before. In fact, it seems more likely that the counterfactuals will inform our understanding of the causal relation. Second, it is well known that the causal relation cannot be understood as a necessary condition along the lines assumed by Roth, since over-determination cases are unavoidable counterexamples. Just as one bomb may have caused the collapse of a house before a second bomb had a chance to do the same, Lisa may also have wanted to impress her roommate, a desire that would have caused her to go for a bike ride in the absence of her desire to exercise, but was not causally active in the face of that desire.

By contrast, explanation by purposes does support counterfactuals, and in a much more relevant sense than even the purported one that we have just seen to fail. This is what distinguishes the normal case of the assassin killing the target through a direct shot from Davidson's example of the stampeding pigs: achieving the goal of killing the target is more certain and less exposed to contingency through the direct shot, and so it is easier to see it as an example of trying to kill the target. We therefore believe that the shooter who instead startled the pigs *would have compensated for whatever error or circumstance caused her to miss, had she been made aware of it beforehand*.<sup>45</sup> This is a counterfactual that ties the agent to her context in virtue of her competence—a way of talking about her skills as means—and thus is a counterfactual that articulates a substantial element of our understanding of agency.<sup>46</sup> It does not, of course, have the logical simplicity of Roth's suggested counterfactual, but I take it to be the Hegelian judgment that it is none the worse off for it.

These two Hegelian theses—the significance of the means and the necessary transformation of the end—articulate the shape of self-determination as elasticity. In the contemporary understanding, they are ways of understanding the plasticity and persistence with respect to the goal that defines all goal-directed behavior as the way in which the goal selects behaviors that conduce to its realization. And it is this conception of plasticity and persistence that



provides the content of the positive account of teleological production, and thus also a way of distinguishing intentional from deviant causal chains.

Finally, since plasticity and persistence are related to the way in which goals bring context into the productive relation, the teleological account does so specifically as an account of *embedded* agency. It does so because the external context that for the causal theorist primarily serves as the conditions under which the beliefs and desires that cause action are formed plays a more expanded role in Hegel's account of productivity, and a role which is more in line with interpretivist theories. The diversity and meaningfulness of natural and social context that Anscombe so powerfully makes central to the understanding and explanation of action are organized by a substantial articulation of the force of the 'because' in action explanations by reasons or intentions, and an articulation that is, broadly speaking, *productive* in the way that the causal theorist expects. In Hegel's theory, context plays the role of the condition of the development of intentions, but this development is recursive, and occurs just as much through the influence of the context on the realization of the intention as in its influence on the initial formation of the intention (if this latter process is considered to be something contrastively internal and subjective). It is a model of productivity that therefore embeds the agent in her context as the medium of the self-development of her own goals, and thus embeds the context in the agent as a system of means for self-expression. In this way, relatively external purposes that begin either as abstractly subjective or as merely given by natural or social features of the context can be internalized through the continual working of the goal on the context. It is not necessary that such internalization be successful—this teleological account makes the problem of alienation a central concern in the theory of agency, and this is particularly apparent in Hegel's discussions in the *Phenomenology*. But the account also holds out the possibility of grounding a theory of free will as the non-alienated experience of nature and community, which must surely be Hegel's conception of being at home with oneself in the other.<sup>47</sup>

#### NOTES

1. Ludwig Wittgenstein, *Philosophical Investigations* (Oxford: Blackwell, 2009), §621.
2. I mean this formulation to be neutral between the different ways that an intention might be related to an action (e.g. as future-directed goal or as the present understanding of the agent of her action). Though there is much discussion about these different forms and

whether any one is reducible to any other, this issue is tangential to the concerns discussed here. The classic presentation of the different ways in which intention can be related to action is in Anscombe's *Intention* (Ithaca, NY: Cornell University Press, 1963), §1.

3. Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), pp. 9–10.
4. *Ibid.*, p. 80.
5. Davidson credits Daniel Bennett with the example (*Essays*, p. 78).
6. For a helpful yet brief discussion of this challenge, see George Wilson, *The Intentionality of Human Action* (Stanford, CA: Stanford University Press, 1989), pp. 169–70.
7. *On Action* (Cambridge: Cambridge University Press, 1990), p. 13.
8. Anscombe, *Intention*, §8. This notion of an intrinsic mark of action is the guiding intuition behind earlier critiques of the causal view according to which the logical connection between intention and action could not be modeled by the causal relation, insofar as that relation requires discrete and logically independent events for its *relata*. Unfortunately, reliance on this notion for the characterization of agency raises the problem of the connection of non-observational first-person knowledge with observational third-person knowledge, a connection that is crucial for the conception of a public realm of meanings that must be central to any robustly interpretivist account (see *Intention*, §29). This criterion of non-observational knowledge is plausibly interpreted in terms of a strict introspectability of the relevant reasons for an action, which is then put to use in rejection of the causal view through the plausible observation that for a large number of paradigmatically intentional actions, there simply is no introspectable reason-state (a desire or belief) to serve as the cause of the action. (See Rosalind Hursthouse, "Intention," in *Logic, Cause & Action* [supplement to *Philosophy* Vol. 46 (2000)], pp. 86–88.) But the strictness of that criterion has seemed overly high a bar to many. Davidson, for his part, points out that, "Those who have argued that there are no mental events to qualify as causes of actions have often missed the obvious because they have insisted that a mental event be observed or noticed (rather than an observing or a noticing)" (Davidson, *Essays*, p. 12). This is connected with the fact that Davidson denies that the belief or desire is, strictly speaking, an event at all or need be in order to play its role in the event-causal story. In any event, Anscombe's clearest presentation of how an intention is non-observational knowledge is in terms of a kind of formal priority of maker's knowledge over its performance, a formulation grounded in Aquinas's characterization of practical knowledge as "the cause of what it understands," which brings us around full circle to the question of the productivity of action, even though this cause is not to be understood as "a mere *extra* feature of events whose description would otherwise be the same" (Anscombe, *Intention*, §48).

9. Anscombe, *Intention*, §26. For other discussions of the teleological aspect of intentionality, see also §44 (no practical reasoning without a prior end in view) and §47 (intentionality as inherently tied to teleology in the form of "further" goals).

10. Anscombe herself rejects what she considers to be the restriction of her view to a teleological account, in claiming that Aristotle's understanding exceeds our own in allowing at least for four kinds of causes—and yet even four is not enough ("The Causation of Action," in *Knowledge and Mind*, ed. Carl Ginet and Sydney Shoemaker [Oxford University Press, 1983], p. 175).

11. This is clearest in the Introduction to the *Philosophy of Right*, e.g. PR, §28. See also §122 of that work.

12. Anscombe, "The Causation of Action," p. 190.

13. Davidson, *Essays*, p. 10.
14. WL 12, pp. 154–55/SL pp. 734–35. The ‘WL’ reference is to Hegel’s *Wissenschaft der Logik*, vol. 12 in *Gesammelte Werke*, ed. F. Hogemann and W. Jaeschke (Hamburg: Meiner Verlag 1978–). The ‘SL’ reference is to *Hegel’s Science of Logic*, trans. A. V. Miller. (Atlantic Highlands, NJ: Humanities Press, 1969). Of course, mechanisms are more complicated entities than forces (e.g. the force of gravity), and the overall argument of the objective logic has been that forces must be understood in terms of mechanisms, i.e. as ways in which mechanisms are related to one another. But Hegel argues that forces of repulsion and attraction are insufficient to individuate the mechanisms that are supposed to be the objects of those forces, and thus that some other principle is required for individuation before explanation by forces can get a grip on the interactions of entities.
15. For a good reconstruction of this argument, see James Kreines, “Hegel’s Critique of Pure Mechanism,” in *European Journal of Philosophy* Vol. 12, No. 1 (2004), pp. 38–74; and section 4.2 of my own *Hegel on Freedom and Reflection* (unpublished manuscript).
16. In the following, I help myself to an association between causation and mechanism in Hegel (in particular, to an identification of mechanical interaction with causal processes). An extended argument for this position can be found in the interpretation of both notions in chapter 10 of my *Freedom and Reflection: Hegel and the logic of Agency* (Oxford: Oxford University Press, 2011).
17. To be clear, Hegel’s concern in the *Science of Logic* is quite a bit broader than the concern about proper as opposed to deviant causal chains as a matter of agency. In his discussion of causation, mechanism and teleology Hegel is concerned with the more fundamental question of how we individuate or explain anything at all. But if Hegel is right that *in general* mechanical systems of causal interaction are individuated only by reference to the ends that they manifest, then there is every reason to approach agency from this perspective. Again, I will not attempt here to judge the truth of this antecedent claim, but rather to explore the consequent project against the background of this analytic dispute about agency.
18. WL 12, pp. 139–42/SL, pp. 718–21.
19. Davidson, *Essays*, p. 65.
20. WL 12, p. 146/SL, p. 25.
21. Here I agree with Kreines’s reconstruction in “Hegel’s Critique.”
22. Larry Wright, *Teleological Explanations* (Berkeley: University of California Press, 1976), pp. 48–49.
23. *Ibid.*, pp. 29–31.
24. WL 12, p. 139/SL, p. 17.
25. WL 12, pp. 159–60/SL, p. 40.
26. WL 12, p. 166/SL, p. 47.
27. WL 12, pp. 162–63/SL, p. 43.
28. WL 12, p. 166/SL, pp. 46–47.
29. It is tempting to think that this plasticity must then invoke a counterfactual element, in the sense that the determination of plasticity rests on claims about what would have happened in other situations, where as in the actual situation only a single causal chain occurred. But I think that the metaphor of the causal “chain” is misleading here, since the shape and direction of influences on a given action is almost never unilinear in the way that the metaphor suggests. Instead, the agent is almost always managing and identifying heterogeneous influences coming

- from different directions, and it is the systematic pattern of orientation of these influences—the way that they all “point to” the goal—that reveals their plasticity. I think that this is in keeping with Hegel’s own modal theory, in which the possibilities that most philosophers represent counterfactually are instead understood to be actual features of the situation.
30. Roth then proposes a “situational environmentalist” account in which a singular causal relation is posited between the action and a mental state in which contextual features are represented. Thus the mental state forges an explanatory tie between the action and its circumstances, and thus provides a way of accounting for complex action over time and in context. Though very suggestive, Roth’s positive account is not well developed in this article. Indeed, it is not clear that it extends at all beyond Davidson’s original account, in which beliefs about the environment and desires for the realization of certain goals are conditions that combine with events in the environment to explain action. In fact, Roth identifies as paradigmatic precisely those homeostatic systems that Davidson also understands as models of the causal power of free will (“Reasons Explanations of Actions: Causal, Singular, and Situational,” *Philosophy and Phenomenological Research* Vol. 59, No. 4 [December 1999], pp. 839–74, 852).
31. Due to considerations of space, I am forced to ignore here the interesting but complex relation between teleological etiologies of kind and of individual virtues, though this relation is relevant to the point here and to Roth’s objection to Wilson. For insightful discussions of this point, see Larry Wright, “Revisiting Teleological Explanations: Reflections Three Decades On,” in *Functions: Selections and Mechanism* (Dordrecht: Synthese Library Publications, forthcoming); and Rowland Stout, *Things that Happen because They Should* (Oxford: Oxford University Press, 1996), pp. 99–112.
32. WL 12, p. 161/SL, pp. 41–42.
33. Cf. Wright, “Revisiting Teleological Explanations,” pp. 10–11: “What stands out in the flux of data is the systematic variation of a pattern’s detail with the conditions affecting the significance of that detail. We thus distinguish its locus, the agent, from the neighboring detritus as a *source* of activity, that is, as a creature *doing* something in the robust sense of that term.”
34. WL 12, p. 160/SL, pp. 40–41.
35. Wright, *Teleological Explanations*, pp. 29–31.
36. Stout, *Things that Happen because They Should*, pp. 50, 53.
37. *Ibid.*, p. 61.
38. *Ibid.*, p. 58.
39. I do not have the space to argue for this axiom of goal-directedness here, but it is central to Hegel’s own conception of teleology (since teleology is a pattern of mechanical interaction, and such interactions are defined in part by the resistance of the parts of the mechanism to external influence (e.g. WL 12, p. 139/SL, p. 17) and his rejection of standard forms of the substance-accident model as a way of making sense of productivity (WL 11, pp. 394–96/SL, p. 557).
40. Stout, *Things that Happen because They Should*, pp. 59–60.
41. *Ibid.*, p. 87.
42. Interestingly, this is precisely the characterization Roth gives of the attempt to explain particular instances of behavior under general teleological relations. Roth finds it inadequate

because the fact that the particular instantiates the general merely has to be assumed ("Reasons Explanations of Actions," p. 850n23).

43. With apologies to Marx, of course (*The Eighteenth Brumaire*).

44. Roth, "Reasons Explanation of Actions," p. 851.

45. It is tempting to think that the employment has to be understood as a particular element within the picture, or a particular link between elements. How else could it be explanatory? This is the mechanistic picture of explanation, and it points to the importance of Hegel's critique of the same for his broader project. The employment is rather to be understood as the dynamic systematic orientation—the plasticity and persistence—as it changes over time in response to environmental factors. In this respects, Roth's "environmental situationalism" is on the right track. But as a result there is no need to reify this reciprocal interaction between goal and context into a specific mental state, as Roth does. Doing so only exposes his view to the attacks of the non-causalists about the instrospectability and individuation of such states, without materially advancing our understanding of the relation between goal and action.

46. The language here was originally suggested to me by Larry Wright. See also "Revisiting Teleological Explanations," pp. 13-14, on the relevance of competence to virtue-etologies for actions. This is similar to Rowland Stout's analysis (*Things that Happen because They Should*, pp. 91-92), and is naturally connected with John Martin Fischer and Mark Ravizza's notion of guidance control (*Responsibility and Control: A Theory of Moral Responsibility* [Cambridge: Cambridge University Press, 1998]).

47. The author would like to thank Matt Talbert, two anonymous reviewers, and the participants in the 2009 Eastern APA Hegel Society of America session on Hegel and Analytic Philosophy for helpful feedback.

## Rethinking Autonomy in Hegel's Earliest Writings

Dean Moyar  
Johns Hopkins University

*Abstract:* This essay investigates the themes of autonomy and conscience in Hegel's earliest writings. Though these themes play a large role in Hegel's mature philosophy, they are largely absent from the writings in his Frankfurt and Jena periods before the publication of the *Phenomenology of Spirit*. The essay argues that essential elements of the mature position on autonomy and conscience can already be found in the treatments in the early writings of moral motivation, moral conflict, formal freedom, and intersubjectivity.

When one thinks of Fichte's response to reading Kant, or of Schelling's response to reading Fichte, Hegel's development to full philosophical originality seems torturously slow. Of the dozen years (1793-1805) that constitute that journey, the two best documented phases are the years in Frankfurt under the influence of Hölderlin and the early years in Jena under the influence of Schelling. Much less clear is the path that led Hegel from his Schelling-inspired position of 1801-1803 to the dramatically different view expressed in the *Phenomenology of Spirit* and essentially carried into his mature philosophy. This transformation, from the Schellingean language of intuition and *Indifferenz* to his own conception of self-consciousness as negativity, is the indirect concern of this paper. My main purpose here is to go further back, before the Frankfurt period, to uncover a dimension of Hegel's early thinking that played a major role in that final phase of his development almost a decade later. A central, and still rather puzzling dimension of the final turn is a thorough reintegration of the standpoint of the autonomous individual into the social conception of Spirit.<sup>1</sup> The autonomous agent of conscience appears in a very prominent place in the *Phenomenology*, namely at the end of the "Spirit" chapter, where this conceptual figure is the historically latest and systematically most advanced shape of Spirit. The discussion of conscience is