

THE AGENT'S ROLE IN THE CAUSATION OF ACTION: IS MICHAEL SMITH'S CAUSAL THEORY OF ACTION IN TROUBLE?¹*

*Lucas Mateus Dalsotto***

<https://orcid.org/0000-0002-8697-6105>

lmdalsotto@hotmail.com

ABSTRACT *The goal of this paper is to find out if Michael Smith's version of the causal theory of action is able to solve David Velleman's agency par excellence challenge. Smith (2012) has claimed that his theory can deal with the challenge insofar as the exercise of the capacity to be instrumentally rational plays the intermediating role which Velleman (1992a) thinks of the agent as playing in the causation of action. However, I argue Smith misunderstands the challenge at hand, thereby failing to find the agent's proper role in action explanation. Moreover, I claim Velleman's objection puts Smith's account of the causal theory in trouble by showing it cannot reconcile the causal explanation of intentional action with our ordinary conception of agency. If Smith intends to explain what a 'full-blooded' intentional action is, I then believe he needs to incorporate into his theory a more robust account of rational guidance.*

Keywords *Action explanation. Causal theory of action. Agency. Michael Smith.*

1 I am indebted to Flavio Williges (UFSM), Sergio Tenenbaum (University of Toronto) and João Carlos Brum Torres (UFRGS/UCS) for valuable remarks and comments on drafts of this paper. I am also indebted to Wilson John Pessoa Mendonça (UFRJ), Frank Thomas Sautter (UFSM) and Leonardo de Mello Ribeiro (UFMG) for composing my thesis defense committee and making research suggestions. Finally, I would like to express my gratitude to CAPES (Coordenação de Aperfeiçoamento de Nível Superior) for the financial support.

* Artigo submetido 23/03/2018 e aprovado em 15/05/2018.

** Universidade de Caxias do Sul. Caxias do Sul, RS, Brasil.

Resumo *O objetivo deste artigo é descobrir se a versão da teoria causal da ação de Michael Smith é capaz de resolver o desafio da agência por excelência de David Velleman. Smith (2012) tem sustentado que sua teoria pode lidar com o desafio na medida em que o exercício da capacidade de ser instrumentalmente racional desempenha o papel que Velleman (1992a) imagina que o agente desempenhe na causação da ação. Contudo, argumento que Smith não compreende corretamente o desafio em questão, falhando desse modo em encontrar o papel adequado do agente na explicação da ação. Além disso, defendendo que a objeção de Velleman põe a abordagem da teoria causal de Smith em apuros ao mostrar que ela não é capaz de reconciliar a explicação causal da ação intencional com nossa concepção ordinária de agência. Se Smith pretende explicar o que uma ação intencional ‘plena’ é, então acredito que ele precisa incorporar em sua teoria uma abordagem mais robusta de orientação racional.*

Palavras-chave *Explicação da ação; teoria causal da ação; agência; Michael Smith.*

Introduction

Michael Smith (2004, p. 156) has argued that the Humean story of motivation gives rise to what is widely known in action theory as the *causal theory of action* (CTA).² Generally speaking, the idea behind CTA is that actions are caused by some desire the agent has and by some belief that she needs to do such and such in order to achieve the intended desire, which jointly cause the relevant behavior in her body. The agent’s bodily movements that are caused in some other way are not considered to be actions, but are rather events that merely take place in her. However, some philosophers (see, e.g., Velleman, 1992a, pp. 461-2; Buss, 1999, pp. 399-401; Bratman, 2001, pp. 309-12) have forcefully claimed that the causal theory faces the problem of finding no place for the agent in the explanatory order of the world.³ They have held that it omits the agent’s participation from the history of *her* action.

So the question I want to answer in this paper is as follows: Is Smith’s version of the causal theory able to properly solve David Velleman’s agency

² Hereafter I will also use CTA to refer to the *causal theory of action*.

³ For other objections to CTA, see Brian O’Shaughnessy (1973, pp. 365-6), Thomas Nagel (1986, p. 110) and Jennifer Hornsby (2004, pp. 2-4).

par excellence challenge? As I aim to argue, I believe we should respond to it negatively. My hypothesis is that Smith's account of action cannot adequately reconcile the causal explanation of intentional action with our ordinary conception of agency. Smith thinks that when the agent exercises his capacity to be instrumentally rational, he is playing his intermediating role in action explanation. Nevertheless, I will try to show it is perfectly conceivable that there are cases in which the agent can fail to *fully* participate in the action and still be instrumentally rational in the causation of it.

Below, I shall begin by presenting an interpretation of the causal theory that Smith favors (section 1). Then I shall describe Velleman's objection to CTA (section 2) and sketch Smith's reply to it (section 3). After that, I shall offer an alternative reading of the agency *par excellence* challenge (section 4) and so argue that Velleman's challenge puts Smith's account of action in trouble (section 5).

1. The causal theory of action

It has been assumed in the recent philosophical tradition that we can provide an intentional explanation of action by citing the agent's reasons for action, and that such reasons must be grounded in the agent's antecedent motivations. As a consequence of both these ideas, some philosophers (see, e.g., Mele, 2003, ch. 10; Smith, 2004, pp. 156-7) have defended the thesis that all other ways of explaining actions intuitively presuppose the existence of the Humean 'belief-desire' psychology, since it appears to be in tune with our ordinary conception of how the world works more generally. According to Smith (2004, p. 157), the Humean theory of motivation, which says that moral motivation cannot arise from moral belief alone but must also depend on a preexisting desire or conative state, and "occupies a central place in the philosophy of action". And apart from that theory, there is no other possibility available to properly explain each and every human action.

In a nutshell, the Humean theory gives rise to a prevailing theory of action⁴ known in the metaethical landscape as the *causal theory of action*, which became the "mainstream" position in action theory. Having been defended most forcefully by Donald Davidson (1980a), CTA tries to account for the difference between being an agent and being a patient in terms of the causal etiology of what happens when a body moves, given that actions are events with a certain

4 It should be noted that there are many controversial issues concerning the metaphysics of action that are beyond the scope of this paper.

causal history (Smith, 2012, p. 394). The bodily movements that we make happen while performing certain actions are causally initiated and supported by antecedents of a particular type. To explain an action, it is first necessary to identify some putative action that an agent performs and then track its effects back to some movement of her body.

CTA is thereby conceived as being a reductive⁵ and unified theory of action. Its advocates aim to construct a story to provide the necessary and sufficient conditions for determining which events are actions performed by agents and which events are merely happenings that take place in them. Even though different variants of the causal theory specify different types of events as the necessary antecedents, they all share a common assumption that the antecedents in question are mental events. Smith's (2004, p. 168) version of the causal theory, in particular, purports to tell us what makes something an action in terms of event-causal process. Briefly, it transpires like this.

(1) An agent-involving event is an action only in the case that it is caused by *the right* agent-involving psychological⁶ states and events in *the right way*.

Before saying something about Smith's account, let me make two clarifications about his defense of (1). First, *the right* psychological states and events mentioned above are mental attitudes that cause and rationalize the performance of the action, such as desires, beliefs, and intentions. Secondly, *the right way* of causation is a *non-deviant* causation, which amounts to cases in which one's desires and means-end beliefs must lead to action in *the right way* in order for an act to be intentional. Yet, given that one of the great challenges faced by the causal theory concerns cases of *deviant* causal chains, something more needs to be said about this second point.

The idea here is that the desires and means-end beliefs that would rationalize a certain action if they caused it in *the right way* may cause it in *other ways* as well. An action that was thought to be performed with the intention Φ may be caused differently. In that way, a deviant causal chain occurs when a desire and means-end belief pair is to cause a bodily movement, and the bodily movement is of the right sort to be rationalized by that desire and means-end belief pair, but the resulting event is incompatible with the desire and means-end belief pair to

5 The causal theory is reductive in at least two senses: first, because it explains the causation of the agents' bodily movements in terms of the occurrence of certain mental events; and second, because it understands that the agents' actions can be explained in terms of causal relations among events.

6 In what follows I will use the terms 'psychological' and 'mental' as interchangeable.

cause and rationalize the bodily movement in question (Wald and Tenenbaum, forthcoming).

To help us illustrate this issue, Davidson's (1980c, p. 79) example⁷ is of a climber who wants to relieve himself of the weight and risk of supporting another climber by means of a rope, and believes that by losing his grip on the rope he can relieve himself of the weight and risk of supporting another climber. Still, his coming to want and believe this makes him so nervous that he loosens the grip on the rope. In this scenario, it is plausible to say that the desire and means-end belief pair that would rationalize his action indeed caused the movement of his body. Nevertheless, it is also plausible to say that he never *chose* to loosen the grip on the rope, nor "did he do it intentionally." To some extent, the event of 'letting go of the rope' seems not to be an intentional action at all. In this way, the challenge that arises for causal theories is to find a way of distinguishing deviant from non-deviant causal chains so that the former cannot be classified as actions.

Once these clarifications have been made, the variant of CTA which Smith favors⁸ is based on the idea that one can determine whether or not an event is an action only by analyzing whether or not an agent's bodily movement is caused and rationalized in *the right way* by some desire he has for things to be in a certain way and by some belief that he needs to do Φ in order to achieve such a state of affairs. A desire for an end, and a belief in the action as a means explain why a given action was taken, and this interaction between the two causes an intention to take that action, which in turn causes the corresponding movements in the agent's body. Both of these states jointly cause the relevant behavior and manifest the causal powers that are partly constitutive of their being a desire and a belief. As Smith (2004, p. 157) asserts, "quite generally, if [a 'belief-desire' psychology] explains the action, then it follows that the things that explain the desire and the things that explain the belief explain the action too."

So the agent's role in the explanation of action is related to the causal roles of the agent-involving psychological states and events, and the exercise of her power to act can be reduced to the causal efficacy of such causal roles. The chain of events that constitutes a person's doing something is that his behavior is caused by a desire and means-end belief pair in the manner that is characteristic

7 According to Velleman (1992a, p. 463), another common example cited to explain deviant causal chains is of an "assassin whose decision to fire on his target so unnerves him as to make his trigger-finger twitch, causing the gun to fire."

8 Smith's (2012, pp. 399-400) version of the causal theory follows largely Carl Hempel's (1961, pp. 291-2) one.

of those attitudes. If in some sense our desires and beliefs are part of what we are, then they are central features of our psychology; hence they amount to the agent's participation in the action. As a result, it is not possible to account for why someone did what he did simply by saying that the particular action appealed to him. It will be necessary to indicate what it was about the action that was appealing as well.

2. The agency *par excellence* challenge

In spite of the fact that the causal theory is the default view in the philosophy of action and agency, it has always had its critics. Some philosophers (see, e.g., Velleman, 1992a, pp. 461-2; Buss, 1999, pp. 399-401) have held, for instance, that CTA faces what may be called *the problem of the disappearing agent*,⁹ since it falls victim to the problem of finding no place for the agent in the explanatory order of the world. They have claimed that the causal theory omits the agent's participation from the history of *his* action, and that nothing counts as an action unless the agent participates in it.¹⁰ This criticism is rooted in the idea that even if what happens in all or most of the episodes we consider to be actions can be explained in terms of bodily movements that are caused by an appropriate desire and means-end belief pair, this is not sufficient for a 'full-blooded' intentional action. With that in mind, the goal of this section is to expose the main ideas of Velleman's (1992a) version of the problem of the *disappearing agent*.¹¹

Velleman's core point is that many versions of CTA are premised in the mistaken supposition that human events can be divided without remainder into actions and mere happenings. These theories neglect the distinction between action and mere activity insofar as they hold a view according to which psychological states simply take place inside the agent, but he takes no active part in the action. Velleman (1992a, p. 462) contends that the causal theory describes an action from which "the distinctively human feature is missing, and that it therefore tells us, not what happens when someone acts, but what happens when someone acts halfheartedly, unwittingly, or in some equally defective way. What it describes is not a human action *par excellence*."

9 This nomenclature is drawn from Alfred Mele's (2003, p. 215) work, where he also suggests that this problem can be called the problem of *shrinking agent*.

10 For more about that, see John Bishop (1989, pp. 41-3).

11 Smith's objection is also known as the agency *par excellence* challenge.

Here we have the puzzle. On the one hand, most causal theories consider that the events recounted in the story of action add up to the agent's role in the action, as components add up to a composite. On the other, Velleman believes that several roles that are actually played by the agent himself in a 'full-blooded' intentional action are played by psychological states whose participation is not equivalent to the agent himself. Strictly speaking, what he has in mind is that things in the world affect one's intention by influencing him to form it, but they affect his intention because they first affect him as an agent. One's intention to perform a 'full-blooded' intentional action is constituted by *oneself* as an agent, so that it does not move one by *itself*. Velleman (1992a, p. 462) then concludes that the agent plays at least two roles in acting: "[s]he forms an intention under the influence of reasons for acting, and [s]he produces behavior pursuant to that intention¹²."

This seems to put some versions of the causal theory in trouble, for we can think of several cases where an agent's desire and means-end belief pair causes a corresponding intention without his *full* participation in the action. If one's role in the action amounts to the description of one's psychological states in action explanation, then there are many occasions where the causal theory "omits the agent without lapsing into" causal deviance. To show this, Velleman (1992a, pp. 464-5) offers us the following example, which I quote it at length.

Suppose that [Mary] has a long-anticipated meeting with an old friend for the purpose of resolving some minor difference; but that as [they] talk, his offhand comments provoke [her] to raise [her] voice in progressively sharper replies, until [they] part in anger. Later reflection leads [her] to realize that accumulated grievances had crystallized in [her] mind, during the weeks before [their] meeting, into a resolution to sever [their] friendship over the matter at hand, and that this resolution is what gave the hurtful edge to [her] remarks. In short, [she] may conclude that desires of [her] caused a decision, which in turn caused the corresponding behavior; and [she] may acknowledge that these mental states were thereby exerting their normal motivational force, unabatted by any strange perturbation or compulsion. But [does she] necessarily think that [she] made the decision or that [she] executed it? Surely, [she] can believe that the decision, though genuinely motivated by [her] desires, was thereby induced in [her] but not formed by [her]; and [she] can believe that it was genuinely executed in [her] behavior but executed, again, without [her] help.

12 An important point to be made here is that in saying this Velleman (1992a, p. 462) does not deny the idea that "the agent's performance of those two roles consists in the occurrence of psychological states and events within him". Velleman himself subscribes his theory to some version of CTA.

To say that Mary was not involved in the constitution of her intention implies accepting that all these processes were to some extent abnormal and that, consequently, this is a deviant case. Nevertheless, that is not Velleman's (1992a, p. 465) claim. In contrast, the claim is that these processes were not abnormal in relation to the causal operation of the motives and intention involved. When desires and means-end beliefs are thought of as directly causing an intention and such intention is thought of as directly causing certain bodily movements, the desires and means-end beliefs are exercising their causal powers in *the right way*. Similarly, when Mary's desires and means-end beliefs are thought of as directly engendering her intention to sever the friendship and such intention is thought of as directly triggering her offensive tone, Mary's desires and means-end beliefs are exercising the same causal powers they exercise in ordinary cases. Thus, even if Mary's behavior is directly governed by her decision and her decision is directly motivated by her desires and means-end beliefs, this does not eliminate the idea that as her "words became more shrill, it was [her] resentment speaking, not [she]."

At least as far as I can see, the point to which Velleman's example seems to draw attention is that even among cases in which the agent's desire and means-end belief pair causes the action in *the right way*, we can still fail to have something as a 'full-blooded' intentional action. We can consciously dissociate ourselves from an intention that is operating upon us, given that being aware of it does not mean we are *causing* a certain movement of our body. The obstacle inherent to causal theories stems from the fact that all events and states of affairs are caused and explained by other events and states. Such events and states of affairs within the agent compose the causal interactions that constitute her being influenced to act. Yet the mental occurrences they describe in the agent are no more than occurrences that happen to or come over her. If one wants to know the time and believes that looking at one's watch will result in knowing the time, then this explains a glance at one's watch (Davidson, 1980b, p. 34). In this view, there is no room for the agent to make a contribution to the resulting behavior.

Hence, the agent's contribution to an action cannot just amount to the fact that his desires and means-end beliefs are exercising their ordinary causal powers.¹³ An appropriate explanation of intentional action requires more than this to account for the agent's participation in the production of action. There should be some event or state of affairs that owes its occurrence to an agent and whose resulting behavior we can trace directly to her. The problem with

13 Christine Korsgaard (2008, pp. 63-4) seems to make a similar point in her book "Self-constitution: agency, identity, and integrity".

CTA, Velleman (1992a, p. 463) says, is that it does not grasp the fact that the agent's role in the action "is to intervene between reasons and intention, and between intention and bodily movements, in each case guided by the one to produce the other. And intervening between these items is not something that the items themselves can do."

However, this is not to say that an explanatory model of action should not allude to states and events occurring in the agent's mind, nor that the causal theory should be denied as a suitable approach to explaining the action. Velleman (1992a, p. 468) himself acknowledges that "any explanation of human action will speak in terms of some such occurrences, because occurrences are the basic elements of explanation in general." Instead, his point is that the causal theory needs to reconcile the causal explanation of intentional action with our ordinary conception of agency, (i.e., the idea that we are capable of exercising control over our behavior and therefore actively participate in the causation of action). That reconciliation should spell out how the causal role assigned commonsensically to the agent reduces to or supervenes on the causal chains among events and states of affairs.

3. Smith's reply to the challenge

In his article "Four objections to the standard story of action (and four replies)", Smith (2012, p. 388) purports to have answers to some of the main objections lodged against the causal theory. The core point of his response is that philosophers of action have good reasons not to put CTA away, for its alleged problems either reside in a "misunderstanding of what the standard story is supposed to be a story about, or they rest on a misunderstanding of how the standard story is properly to be told, or they presuppose views about the nature of action that are themselves implausible." Because of this, the aim of this section is to lay out Smith's reply to Velleman's (1992a, p. 461) objection¹⁴ that some variants of the causal theory "fail to cast the agent in his proper role" in action explanation.

As viewed earlier, one of the great challenges that the proponents of CTA need to deal with concerns the so-called problem of deviant causation. To solve it, it is important to determine which conditions are necessary to be met for agency, since such philosophers disagree among themselves about what is necessary and sufficient for it. In Smith's case, the fact that a bodily movement

14 For another reply to Velleman's agency *par excellence* challenge, see especially Mele (2003, ch. 10).

is caused by a suitable desire and means-end belief pair is necessary for agency but not sufficient. And this is so because in cases of deviant causal chains, the match between what one does and the content of one's desires and means-end beliefs is random. The condition of non-randomness is clearly violated in cases of deviant causal chains.

Smith (2012, p. 398) then follows Christopher Peacocke's (1979) account and suggests that "something an agent does is an action only if what he does is *differentially sensitive* to the contents of his desires and beliefs." In a nutshell, the idea of the differential sensitivity requirement is that the agent's desire and means-end belief pair must not simply cause the action, but also differentially explain it. A given mental state will differentially explain the behavioral outcome if there is a *function*¹⁵ relating the mental state to the behavior. In addition, the function relating the mental state to the behavior is 'one-to-one'; meaning that each mental state results in a different behavior. Peacocke (1979, p. 67) thus affirms that if the condition that differentially accounts for a certain event has been one of a set of conditions such that, "given the same principle of explanation employed in [that] particular case in question, then if similar conditions had been met in other [events] at other times, a corresponding *explanandum* sentence would be true."

Smith (2012, p. 398) tries to show how the requirement is supposed to work by offering an example. Imagine a situation in which a piano player wants to appear extremely nervous to the audience when he plays the piano and that, to do so, he hits a G# when he should hit a G at a given moment in the performance. In such a case, his actually hitting a G# at a given moment in the performance will be an intentional action only if, in a contrafactual scenario, he would have hit a G at that moment in the performance in the case that he did not want to appear nervous to the audience. Or suppose that, to appear nervous to the audience, he believes he could do so by hitting F at a given moment in the performance. In these circumstances, his actually hitting G# at a given moment in the performance will be an intentional action only if, in a contrafactual scenario, he would equally have hit a F at that time in the performance. In terms of a theory of action, Smith (2012, p. 398) says, the clause behind such requirement is that an agent acts intentionally only in the case that he would have realized his actual desires and means-end beliefs "in a

15 To explain the idea of *function*, Sehon (2016, p. 104) gives us the following example. Imagine a mathematical function, says " $y=x^2+1$ ", the y value is a function of the x value, in that for each inserted x -value a different y -value results."

range of nearby possible worlds in which he had a set of desires and means-end beliefs with ever-so-slightly different contents.”

The next step, then, is to hold that the differential sensitivity requirement is in line with the consideration that agents have the capacity to be instrumentally rational, where this means to have the capacity to realize their desires, given their means-end beliefs. According to Smith (2012, pp. 398-9), the requirement does not merely claim that for an agent to act intentionally she must possess the capacity to be instrumentally rational, but that she must also exercise that capacity in a certain way. After all, it is the agent's possession and exercise of the capacity to be instrumentally rational that guarantees the truth of the piano player's contrafactual, for example.

From this we can infer that what Smith (2009, p. 59) is suggesting is that, instead of two psychological states that play a causal role whenever an agent performs an action, there are, in fact, three.¹⁶ Beyond the psychological states of having a desire for a given state of affairs in the world, and a means-end belief that it is necessary to do such and such to make the world that way, we have the capacity to be instrumentally rational as well. We ordinarily attribute certain desires and means-end beliefs to the agents, and we also claim they are rational and hence will respond in a characteristic way to those desires and means-end beliefs. If an agent did not do what he did by virtue of his possession and exercise of the capacity to be instrumentally rational, then the event in question cannot be differentially explained in Peacocke's terms.

Methodologically, we can divide Smith's (2009, p. 61) theoretical account of action in two main parts. First, agents have certain desires they seek to realize in their lives and certain beliefs they take as means to get these desires. Second, agents are rational, and so they have the capacity to act in accordance with their means-end beliefs, given their desires. Unless these two parts of his theoretical account are met, there is no action at all.

For the sake of Smith's argument, let us imagine a situation in which an agent, say Charles, has a desire to lose weight by virtue of his heart disease and a belief that something he can just do, namely stop eating processed food and start doing physical exercise, would make him lose weight. Let us imagine further that, as a result, Charles stops eating processed food and starts doing physical exercise. On this view, the conditions under which Charles' bodily movements are an action should match at least the following steps:

16 According to Smith (2009, p. 62), “there is an extra psychological element in a Humean constitutive explanation of an action [...]. Every constitutive explanation of an action [...] comprises three basic psychological elements: a desire, a means-end belief, and the agent's exercise of her capacity to be instrumentally rational.”

- (2) Charles desires losing weight;
- (3) Charles believes he can lose weight by ceasing to eat processed food and by starting to do physical exercise;
- (4) Charles is instrumentally rational;
- ∴ (5) Charles stops eating processed food and starts doing physical exercise.

In that scenario, what makes Charles stop eating processed food and start doing physical exercise is the fact that he is instrumentally rational. The role of (4) in the reasoning is to guarantee the consistency of his mental states and the corresponding movements of his body. Without it, Charles could fail to perform (5). He could have the desire to lose weight and the belief he could lose weight by ceasing to eat processed food and by starting to do physical exercises, and even so not form the instrumental desire to stop eating processed food and start doing physical exercise. Absent Charles' putting (2) and (3) together, Smith (2009, p. 62) says, he will not be “instrumentally rational and so we won't be able to explain his doing anything in the way characteristic of action because he won't act.”

In light of all this, Smith (2012, p. 399) then maintains that the agency *par excellence* challenge rests on a misunderstanding of how the causal story is properly to be told, since it claims something that many defenders of CTA “had already seen and built into their theories.” He thinks that the agent's exercise of his capacity to put his desire and means-end belief pair together is fully consistent with the intermediating role that Velleman imagines the agent to cast in the causation of action. The distinction between a deviant and non-deviant cause is that a non-deviant cause not only initiates the action, but also continues to control and maintain that action throughout its performance. In this way, something is intentionally done by the agent only in the case that he is instrumentally rational. Smith thus holds that there is no reason to put the causal theory away, and hence philosophers of action should stick with it.

4. Outline of an alternative reading

In spite of the numerous controversial issues concerning what constitutes human agency, philosophers of action have widely held that it involves the initiation of action by the agent herself.¹⁷ But what exactly does it mean? In Smith's view, that initiation can be explained in terms of causation by the agent's psychological states and events, and his capacity to put his desire and

17 For an appropriate overview about it, see the entry on *agency* in Stanford Encyclopedia of Philosophy, which is available at: <https://plato.stanford.edu/entries/agency/>.

means-end belief pair together in order to make his body move in a certain way. A person causes his own action through the proper exercising of his capacity to be instrumentally rational, which allows the actions to take place.

Smith defends the idea that agency itself can be reduced to the causal efficacy of the agent's mental states and events. He is quite sure that this kind of explanation succeeds in solving Velleman's objection in the sense that it finds a place for the agent in the explanatory order of the world. Yet I have serious doubts as to whether Smith's theory succeeds in doing such a thing. It strikes me that his understanding of the agency *par excellence* challenge does not really capture what Velleman is objecting to. The goal of this section, then, is to broadly offer an alternative reading of the challenge.

As already viewed, Smith (2012, p. 398) presupposes that the heart of Velleman's criticism is that, in order to participate in the causation of action, the agent needs to intervene between her reasons and intention, and between her intention and bodily movements. On the one hand, Smith seems to be right in thinking that way. Indeed, Velleman (1992a, p. 475) maintains that the problem with the majority of causal theories is that they omit the agent of the story of action by not mentioning "anything that plays his intermediating role." But on the other hand, the issue brought up by the agency *par excellence* challenge seems to be something deeper than Smith thinks it to be. At least as I see it, the point is that the agent intervenes throughout this process only in the case that he reflects on the reasons "vying to govern his behavior", thereby taking sides with some of them rather than others (Velleman, 1992a, p. 477).

In a rough way, we live, so to speak, under the pressure of a large variety of laws, duties, obligations, social expectations, and personal commitments, all telling us what to do. Some of these demands lead us to think that the spectrum of possibilities for choosing what to do is considerably low and restricted. Nevertheless, even if the world establishes many of the external and internal conditions of our choices, we are still capable of changing our minds whenever we want. We believe we have the capacity for active self-determination of our choices and decisions. And this is so, Velleman (1992a, p. 466) says, because we understand that what "makes us agents rather than mere subjects of behavior – in our conception of ourselves, at least, if not in reality – is our perceived capacity to interpose ourselves into the course of events in such a way that the behavior outcome is traceable directly to us."

Still according to Velleman (1992a, p. 471), these considerations are strongly supported by the ordinary thought that when reflecting upon the reasons that compete for governance of her behavior, the agent occupies a *critical detachment* standpoint from those reasons. In that position, the agent understands that his

acts are not determined simply by the strength of the desire to which he is antecedently subject, but that it is under his control to decide what he has most reason to do. It is up to the agent to either reinforce or refuse the influence of his reasons “by throwing her weight behind some of them rather than others¹⁸.” And he can only do this because he is capable of grasping his reasons and deciding which ones he wants to *guide* his actions. It turns out that people can make good sense of themselves as agents only in the case that they believe they are able to choose what they have most reason to do. Yet it must be clear that it is wrong to think of people as going through such a process of reasoning each time they act. Moreover, they can also fail to grasp their reasons and then participate in the production of action incompletely. To some extent, this seems to be the case of Mary’s intention to sever the friendship.

Velleman’s (1992a, p. 478) idea here is that the agent is playing his intermediating role in the causation of action in the sense of being able to set the inner conditions of his choices, and to view his desires and intentions as providing suggestions he may take or leave. Once reasons have varying strengths, the agent has the power to decide which of them he wants to form his intention for acting. The agent’s involvement in his behavior amounts to his capacity to adjudicate conflicts among the reasons he has, and to determine which of them provides him with the strongest reason for acting. In a way, this is the full and active participation that people commonly attribute to the agent in action explanation. The agent’s practices of deliberation and rationalizing are designed for the action as he conceives it to be, and any account of reality different from such “conception will not help us understand the logic of these practices” (Velleman, 1992a, p. 465). Unless causal theories take this thought seriously, they cannot explain intentional actions at all.

A suitable conception of agency thus needs to explain how we are *guided* by our reasons, rather than just how we are *moved* by them.¹⁹ In order to provide a complete explanation of intentional action, an account of *rational guidance*²⁰

18 Korsgaard (2008, p. 69) seems to share a consonant view in stating that “[s]uppose that I decide to get some work done on my book today. At this moment, now, I decide, I *will*, to work; at the next moment, at any moment (importantly, maybe even at this moment), I will certainly *want* to stop. If I am to work I must *will* it – and that means I must determine myself to stay on its track. Timidity, idleness, and depression will exert their claims in turn, will attempt to control or overrule my will, to divert me from the work. Am I to let these forces determine my movements? At each moment I must say to them: ‘I am not you; my will is this work’.”

19 According to Velleman (1992b, pp. 4-5), if philosophers intend to make sense of what ‘acting’ consists in, then it is important to be clear that “one and the same action could be due to a confluence of [the story] of motivation and [the story] of rational guidance.”

20 From Jay Wallace’s (1999, p. 238) outlook, rational guidance can be widely understood as “the fact that our behavior is controlled by our deliberative understanding in cases in which we succeed in complying with our judgments about what there is reason to do.”

must tell us how we act for a reason, given that we want to be able to make sense of our actions. It seems reasonable to say that a central part of what an explanatory model of action should do is to show how the causal connection between the agent's attitudes and his bodily movements ensures that *he* guides and secures the resulting action. In acting for a reason, the agent acts not just because his attitude towards the reason is more like belief than disbelief, but also because the proposition in question counts in favor of his action rather than against it. The agent's attitudes, Velleman (1992b, p. 7) points out, are thought of as "having propositional objects that intrinsically favor a particular action, and that their favoring the action is conceived as crucial to their behavior influence."

From this follows that having an appropriate account of rational guidance is an essential condition for solving the problem of deviant causation. Without such account, even among cases in which one's desire and means-end belief are exercising their causal powers in *the right way*, "we can still have cases in which we have something that it is not a 'full-blooded' intentional action" (Wald and Tenenbaum, forthcoming). Moreover, as much as I can see, this stems from the fact that the agent is not in control of how his desires take place in him.

In light of this, I then think that the claim behind the agency *par excellence* is that reconciling the causal explanation of intentional action with our ordinary conception of agency will be possible only if causal theories can explain how agents are *guided* by their reasons as well. It is hard to understand how merely *responding* to reasons might count as a proper exercise of our capacity of active self-determination. Smith seems to see agency as an essentially explanatory phenomenon and, in my view, this threatens the success of his account.

Only from an explanatory perspective does agency not contribute to accounting for the choices and decisions that people make when acting. And as I tried to show just now, the explanation of the intentional action is also an explanation of how the action appeared reasonable to the agent from his point of view (Tenenbaum, 2007, p. 35). That is, how it was the result of the exercise of his deliberative capacity of deciding on which course of action he had most reason to take in the circumstances he found himself. The things we cite in explaining a certain action must be such as to make it possible to understand the normative factors that the agent took into account in deciding what to do.

Before we move on, I think that one thing needs to be made clear. In arguing that the agent plays his intermediating role in the causation of action by reflecting on the reasons he has, Velleman is not begging the question of agent-causation.²¹ Remember, he claims that an adequate causal explanation of

21 For present purposes, I will not address the positive part of Velleman's theory here since this would require an exclusive work.

intentional action should account for how the role assigned commonsensically to the agent supervenes on the causal network among events and states of affairs. In his own theory, Velleman (1992a, p. 474) tries to reduce the agent's participation in the action to mental phenomena – such as desires, values, or decisions – “whose existence we can assume without presupposing that agent-causation occurs.” For him, finding a place for the agent in the explanatory order of the world does not imply denying event-causation.

5. An unresolved issue: the agent's role in action explanation

On the assumption that my reading of Velleman's objection does in fact stand, it is therefore reasonable to conclude that Smith at best misunderstands the objection, for he fails to grasp in depth what is behind it. But I believe we can go further and say that the agency *par excellence* challenge puts Smith's version of CTA in trouble by showing that his theory, at least as it is currently built, is not capable of reconciling the causal explanation of intentional action with our ordinary conception of agency.

Smith (2012, p. 400) affirms that those who are interested in discussing the plausibility of the causal theory should turn their attention to questions like: “Can we convincingly spell out the differential sensitivity solution to the problem of [deviant] causal chains? Can we really turn that account into a plausible and satisfying account of what it is for an agent to possess and exercise the capacity to be instrumentally rational?” Following his suggestion, I guess that these two questions should be answered in a negative way regarding Smith's approach, since answering them otherwise would require an account of rational guidance that his theory does not seem to have at this moment. In this section, I aim to hold that Smith's version of the causal theory falls victim to Velleman's agency *par excellence* challenge by not finding the agent's proper role in action explanation.

As already seen, Smith believes that Peacocke's differential sensitivity requirement is capable of ruling out cases of deviant causal chains. The core idea of the requirement is that if a given desire differentially explains why one does something, then it causes one's action in a *non*-deviant way. Otherwise, it does not. In the case of Davidson's climber, for instance, we have a deviant causal chain because his desire to relieve himself of the weight and risk of supporting another climber does not account for any feature of his loosening the grip on the rope. Once he *does* such a thing through nervousness, it is plausible to say that the speed with which he opens his hand or the way in which he loosens the grip on the rope are not sensitive to any feature of the climber's desire.

Everything that happened would have been similar even if the climber's desires and means-end beliefs had ever-so-slightly different contents.

Yet it is not clear how the differential sensitivity requirement would be able to rule out deviant causal chains at all. Scott Sehon (1997, pp. 208-9) has forcefully argued that the requirement is too strong, for it is plenty acceptable to state that we could have an intentional action even if desires and means-end beliefs with even-so-slightly different contents produce the same movement in the agent's body. For him, the differential sensitivity requirement claims something that does not fit properly with our ordinary experience. In many situations, the requirement is counterintuitive because it makes us classify as deviant some cases that are to us truly intentional actions. And these cases deny Peacocke's (1979, p. 79) idea that the function relating the mental state to the behavior is "one-to-one", meaning that each mental state results in a different behavior.

To clarify his point, Sehon (1997, p. 209) gives us the following example. Let us imagine a pitcher who has an excellent and accurate control over the speed of his pitches and whose maximum pitching speed is 75 mph. In this way, the pitcher's intention for how fast to pitch the ball will differentially explain the ball's speed. If he aims to pitch the ball at 55 mph, the ball will go 55 mph, and if he aims to pitch the ball at 73 mph, the ball will go 73 mph. Let us imagine further that he aims to make a new move and then pitch the ball at 78 mph. In such a case, the ball would have gone 75 mph, because this is his maximum pitching speed, at least at this time. Intuitively, this last attempt to pitch the ball seems to be clearly an intentional action to us, though it contradicts Peacocke's idea that the function relating cause to effect is "one-to-one". I thus think it makes sense to state that, as Smith endorses the differential sensitivity requirement, to some extent Sehon's argument puts Smith's version of the causal theory in trouble as well. Moreover, this seems to strongly suggest a negative answer to Smith's first question posed above.

Anyway, perhaps to avoid this kind of criticism presented by Sehon, Smith (2012, p. 399) qualifies his approach by maintaining that the differential sensitivity requirement indicates that an agent can act intentionally only in the case that he exercises properly his capacity to be instrumentally rational, where this means to achieve his desires, given his means-end beliefs. If a person fails to do what she does because of her capacity to put her desire and means-end belief pair together in order to make her body move in a certain way, "then her desires and beliefs with even-so-slightly different contents will not stand in the modally rich pattern of connections that they must stand in if they are to satisfy the differential sensitivity requirement." For an event to be intentionally caused,

it is necessary that the agent is capable of maintaining the consistency between his mental states and the corresponding movements of his body.

Nevertheless, Smith's idea of the agent's capacity to be instrumentally rational seems to fail to find the agent's proper role in action explanation by classifying some bodily movements as intentional actions without his full participation in the behavioral outcome. And that flaw stems from the fact that the agent's exercise of his capacity to put his desire and means-end belief pair together is just a psychological state of causal efficacy. It is a psychological state that does not help the agent reflect on the strengths of his reasons and take sides with some of them rather than others. Further, it does not contribute to explaining the choices and decisions that people make when acting, nor does it show how the action appeared reasonable to the agent from his point of view. As a result, I believe that if from one perspective Smith's theory clarifies how the agent's reasons *move* him into action, from another it cannot suitably explain how he is *guided* by them. Thus, in order to explain what a 'full-blooded' intentional action is, a more robust account of rational guidance is needed.²²

My argument that the exercise of the capacity to be instrumentally rational cannot play the proper role in the action that Velleman imagines the agent to play is as follows. The capacity to be instrumentally rational, as Smith presents it, is only the idea that the agent plays his intermediating role in the exercise of action insofar as he takes the means to achieve the intended goal. This amounts to stating that if I desire to drink some coffee and I believe I can get it by going to the nearest coffee shop, then what accounts for my intermediating role in the action is my going to the nearest coffee shop to buy the coffee. Or, if I desire to gift my wife with a jewel and I believe I can buy it by going to the jewelry store *Y*, then what accounts for my intermediating role in the action is my going to the jewelry store *Y* to buy the jewel that I want to. In both of these situations, we clearly have cases of intentional actions in which the agent actively participates in the resulting behavior.

Nonetheless, in order for the agent to play his intermediating role in the action, the capacity to be instrumentally rational does not require him be consistent with his desires and means-end beliefs over time. That is, he does not need to be consistent with those values and goals he imagines as constituting him as a unified agent over time. Instead, what the capacity to be instrumentally rational requires is just that every time an agent has a new desire, he takes the means to achieve the new intended desire. But if whatever an agent instrumentally

²² My point here is just that, in considering Velleman's argument, it makes sense to say that Smith's theory lacks a more qualified account of rational guidance.

does is the means to the end he is going to pursue, then how can he be *guided* by his reasons when anything he instrumentally does counts as following them? Or how can he understand those normative factors that led him to choose those reasons and not others? From my outlook, these questions raise the problem that Velleman is calling attention to, for it is certainly possible that there are cases in which the agent can fail to fully participate in the action and still be instrumentally rational in the causation of it. That is, the agent's desire and means-end belief pair can be exercising their ordinary causal powers in order to get the intended end without *any* contribution from him.

To view these cases it is not necessary to imagine cases of "psychological compulsion, physical addiction, and emotional disturbance" (Smith, 1992, pp. 325-6). It is possible to have simpler counterexamples in which the agent fails to actively participate in the production of action without lapsing into deviant causation. As I understand the issue, this is the reason why Velleman presents Mary's case. By the time that Mary was arguing with her old friend, to some extent she wanted to sever their relationship, and she believed she could do this by alienating her friend, raising her voice, replying sharply to his offhand comments, and so on. In that case, Mary's desires and beliefs were exercising their ordinary causal powers in *the right way*. And in making her body move in a certain way, Mary was exercising her capacity to be instrumentally rational in order to achieve her desire, given her means-end beliefs. Of course, these causal relationships between Mary's reasons and intention, and between her intention and bodily movements, may have been mediated by a number of subconscious intentions. However, this should not be a problem, since most of our actions are mediated by a number of subconscious intentions, many of which we will not be able to access even through later reflection.

Regarding Smith's version of the causal theory, we can thereby conclude that Mary's example is a case of intentional action. Yet this seems to be inconsistent with our ordinary conception of agency. As Velleman sought to show, in Mary's case we have a situation where the agent feels as though she has not participated in the resulting action. Mary was not able to see her behavioral outcome as having been done by *her*²³ because she was not able to intervene among her desires, intention, and bodily movements. If Mary had been able to resist her anger she would have rejected it as a reason for acting, and then she would

23 To say this does not mean that Mary is not responsible for what she did. As Velleman (1992a, p. 466) says, her "responsibility for the action in question arises from [her] having failed to prevent or control it rather than from [her] having truly initiated it. And [she is] responsible for having failed to prevent or control the action because it would have yielding to various measures of self-scrutiny and self-restraint that [she] could have initiated."

have seen her action as a production of *her*, where this means that she would have made sense of her action from her point of view. It follows that, in order to actively participate in the action, the agent cannot just be determined to act by the strength of the desire to which she is antecedently subject. More than that, she needs to reflect on the reasons she has and then decide which ones she wants to *guide* her actions. So I believe that Smith fails to find the agent's proper role in action explanation, and this suggests a negative answer to his second question posed at the beginning of this section as well.

Before ending this debate, it is worth drawing attention to one thing. In fairness to Smith (2012, p. 400), it is possible to say he is right in claiming that there is nothing that requires us to abandon CTA, given that Velleman himself subscribes his theory to it. Nevertheless, this does not disqualify the relevance and strength of the issues raised by the agency *par excellence* challenge. On the contrary, it makes it clear that if causal theories intend to provide an appropriate explanation of intentional action, then they need to find a way of introducing in their explanatory models the role that the ordinary parlance attributes to the agent when acting. But, as I tried to argue, Smith's approach does not cover the grounds where our conception of agency is found, even by appealing to the agent's exercise of his capacity to be instrumentally rational. And this seems to me to be a flaw of his theory. Once our actions are designed as we conceive them to be, any explanatory model of reality different from such conception will not help us have an accurate understanding of what happens when one is acting. For that reason, I defend the idea that the agency *par excellence* challenge puts Smith's account of action in trouble by showing it cannot reconcile the causal explanation of intentional action with our ordinary conception of agency.

Conclusion

In this paper, I have sought to answer the following question: Is Smith's version of the causal theory able to properly solve Velleman's agency *par excellence* challenge? On the one hand, Velleman's (1992a, p. 463) objection is that in order for the agent to play his role in the causation of action he needs to intervene between reasons and intention, and between intention and bodily movements. On the other hand, Smith (2012, p. 399) imagines that the agent plays this role by exercising his capacity to put his desire and means-end belief pair together in order to make his body move in a certain way. However, I have argued that Smith misinterprets the agency *par excellence* challenge, thereby failing to find the agent's proper role in action explanation. I have also defended that his theory cannot reconcile the causal explanation of intentional action with

our ordinary conception of agency and, therefore, we should respond to the question above in a negative way. To lay out all this, the route taken in this work was as follows. I began the discussion by providing an overview of a variant of the causal theory that Smith favors. Then I described Velleman's objection to CTA and outlined Smith's reply to it. After that, I offered an alternative reading of the agency *par excellence* challenge. Finally, I argued that Smith's version of the causal theory falls victim to Velleman's challenge because the agent can fail to fully participate in the action even if he is instrumentally rational in the production of it.

References

- BISHOP, J. "Natural agency: an essay on the causal theory of action". Cambridge: Cambridge University Press, 1989.
- BRATMAN, M. "Two problems about human agency". *Proceedings of the Aristotelian Society*, pp. 309-326, 2001.
- BUSS, S. "What practical reason must be if we act for our own reasons". *Australasian Journal of Philosophy*, Vol. 77, Nr. 4, pp. 399-421, 1999.
- DAVIDSON, D. "Actions, reasons, and causes". In: D. Davidson. *Essays on actions and events*. Oxford: Clarendon Press, 1980a. pp. 3-20.
- _____. "How is weakness of the will possible?". In: D. Davidson. *Essays on actions and events*. Oxford: Clarendon Press, 1980b. pp. 21-42.
- _____. "Freedom to act". In: D. Davidson. *Essays on actions and events*. Oxford: Clarendon Press, 1980c. pp. 63-81.
- HEMPEL, C. "Rational action". *Proceedings and Addresses of the American Philosophical Association*, Vol. 35, pp. 5-23, 1961.
- HORNSBY, J. "Agency and actions". In: H. Steward, J. Hyman (eds.). *Agency and action*. Cambridge: Cambridge University Press, 2004. pp. 1-23.
- KORSGAARD, C. "Self-constitution: agency, identity, and integrity". New York: Oxford University Press, 2008.
- MELE, A. "*Motivation and agency*". Oxford: Oxford University Press, 2003.
- NAGEL, T. "*The view from nowhere*". Oxford: Oxford University Press, 1986.
- O'SHAUGHNESSY, B. "Trying (as the mental pineal gland)". *Journal of Philosophy*, Vol. 70, pp. 365-386, 1973.
- PEACOCKE, C. "*Holistic explanation: action, space, interpretation*". Oxford: Clarendon Press, 1979.
- SCHLOSSER, M. "Agency". *The Stanford Encyclopedia of Philosophy*. 2015. Available at: <https://plato.stanford.edu/entries/agency/> (Accessed June 3rd, 2017).
- SEHON, S. "Deviant causal chains and the irreducibility of teleological explanation". *Pacific Philosophical Quarterly*, Vol. 78, pp. 195-213, 1997.

- _____. “Free will and action explanation: a non-causal, compatibilist account”. Oxford: Oxford University Press, 2016.
- SMITH, M. “Valuing: desiring or believing?”. In: D. Charles, K. Lennon (orgs.). *Reduction, explanation, and realism*. Oxford: Oxford University Press, 1992. pp. 323-360.
- _____. “The possibility of philosophy of action”. In: M. Smith. *Ethics and the a priori: selected essays on moral psychology and metaethics*. New York: Cambridge University Press, 2004. pp. 155-177.
- _____. “The explanatory role of being rational”. In: D. Sobel, S. Wall (eds.). *Reasons for action*. New York: Cambridge University Press, 2009. pp. 58-80.
- _____. “Four objections to the standard story of action (and four replies)”. *Philosophical Issues*, Vol. 22, pp. 387-401, 2012.
- TENENBAUM, S. “Appearances of the good: an essay on the nature of practical reason”. New York: Cambridge University Press, 2007.
- TENENBAUM, S.; WALD, B. “Reasons and action explanation”, In: D. Star. *The Oxford handbook of reasons and normativity*. Forthcoming.
- VELLEMAN, D. “What happens when someone acts?” *Mind*, Vol. 101/403, pp. 461-81, 1992a.
- _____. “The guise of the good”. *Nous*, Vol. 26, pp. 3-26, 1992b.
- WALLACE, J. “Three concepts of rational agency”. *Ethical Theory and Moral Practice*, Vol. 2, pp. 217-42, 1999.