

The Agent in Action: A Critique of Velleman

Stephanie Wykstra

Action theory faces the challenge of explaining how the following physically identical actions differ. (1) An electric probe stimulates my brain, and I raise my arm. (2) I raise my arm unprobed. The gut intuition is that *I* am, in some sense, passive in the first case and active in the second. Agent causation proponents argue that standard action theory does not account for agency and seek to explain how an agent participates in action. J. David Velleman claims that certain attempts to describe agency, namely those of Harry Frankfurt and Gary Watson, do not adequately explain how the agent participates.¹ He offers his own account in which the agent is played by a functionally identical motive: the desire to act according to reasons. In this essay, I will begin by presenting the standard theory and Velleman's (Frankfurt-inspired) revision. I will then complain that his account suffers from incompleteness, in that he limits agent functioning to conscious desires without giving convincing reason to do so. There is, I will argue, good reason to hold that the unconscious figures prominently in agent-caused action. I will press his account by showing that if he does not allow for cases of action prompted by unconscious motivations, he runs the risk of over-exclusion: actions that seem agent-caused would not count as such. Though I devote the bulk of the paper to criticizing Velleman, I will conclude by sketching my own revised account of agent-caused action.

I. Standard Action Theory

The widely accepted story that philosophers of action tell is something like Donald Davidson's account in *Essays on Actions and Events* (1980).² Desires and reasons constitute motivations that – together with beliefs about which action will achieve the desired end – lead to the formation of an intention to act. The intention to act is a decision that, given normal functioning, instigates the necessary physical processes required for its performance.

Frankfurt and Velleman claim that the standard account does not include the agent, though it takes itself to do so.³ Instead of participating, the agent is the “arena” of his motivations and intentional states, which are causally linked to actions without agent involvement.⁴ Velleman stresses that his complaint is not simply that the standard account deals with states of the agent and fails to mention the agent himself; rather, it does not explain the agent's involvement in producing the action. Something essential is left out: the means by which the agent participates in action.

II. Alienated Actions and Unwitting Decisions

Velleman writes, “What makes us agents rather than mere subjects of behavior...is our perceived capacity to interpose ourselves into the course of events in such a way that the behavioral outcome is traceable directly to us” (465). His thought seems right; it is the gut intuition that there is a difference between arm-raise₁ and arm-raise₂. In the second case it seems that I have some power with respect to raising my arm. Frankfurt's drug addicts will help crystallize the intuition that two identical actions differ with respect to whether an agent participates in his behavior.⁵

First, a distinction. That an agent participates in his own behavior might seem obvious. If I perform an action, then clearly I am the participant – who else would be the participant, if not me?⁶ There is, however, a distinction to be made between an action as a bodily happening and an action for which the actor is responsible in a non-trivial way. A mere happening in the body, such as an arm raised upon electrode probing, is a case of weak agency: a person who dances about on an electrified carpet is, in a sense, acting. But what is required for properly full-blooded agent-causation must be more than mere happenings of the body or mind, agree Frankfurt and Velleman. Weak agency, mere bodily happening, is not traceable to the actor in a way that makes the actor a responsible agent.

Consider two addicts.⁷ Both have the same first-order desire for a drug, which is to say that their physiological attraction to the drug is equally intense, and both partake on occasion. There is a difference between them, however. One of them is an *unwilling* addict who hates his addiction; he struggles against his habit to no avail. That is, his first-order desires are conflicted; he both wants and does not want to take the drug. He has second-order volitions that oppose his first-order desire to use, and he says that he tries to resist, but it is as though he is a “helpless bystander to the forces that move him.”⁸ The other, a willing addict, has various reasons for wanting to do drugs; his first and second-order desires are neither internally conflicted nor opposed.

The unwilling addict, Frankfurt claims, does not do what he claims to *really* want to do. He is not a free agent, but is instead alienated from his compulsive desire and accompanying behavior. If a person is prompted to action by a desire that he disapproves of and would rather not indulge, then he is not the originator of his action in the proper

sense.⁹ An alienated actor is not an agent, Frankfurt claims, because strong agency requires “identification” with desires and actions. The willing addict is the real agent of his drug use in a way the unwilling addict is not because he has embraced the state of affairs – his motivation, intention, and action – as his own. All of this coheres with the common intuition about the difference between arm-raise₁ and arm-raise₂. When I am coerced into movement or helpless to control addictive behavior, I am not acting as an agent.

The standard account, Frankfurt and Velleman claim, does not provide a way to discriminate the unwilling and the willing addict. In both cases, the agent has certain desires that, in light of beliefs about which actions will best satisfy his desires, lead to the formation of intentions and subsequent actions. The account provides no means to differentiate alienated and thereby coerced subjects from true agents.

Standard action theorists might respond that even if this charge does hold water, instances of compulsive or psychotic behavior are not normal. Drug addict cases do not succeed in showing the problem, because they are examples of abnormal behavior, which the standard account rules out in limiting itself to *normal* behavioral functioning.

Velleman, sensitive to this possible reply, presents a variation – the unwitting decision-maker – that he claims manifests normal functioning and succeeds in making the same argument. V-man, incensed over some trivial skirmish, ends a friendship. Reflecting upon the matter afterwards, he realizes that he had unconsciously decided some time before to end it upon a suitable pretext; his friend had become insufferable and it was too tedious to maintain a relation with him. Velleman concludes that V-man was not really the agent of his anger, because he made an unwitting decision that he did not

identify with in the sense required for true agency. Like the unwilling addict, he is alienated from his desire, and is not an agent with respect to it.

III. Velleman's Account

The unwilling addict and unwitting V-man are both cases that the standard account would allow as agent-caused, but the intuition is that they are not agent-caused. Therefore, they show that the standard account leaves out a crucial element that would allow for discrimination between cases of strong and weak agency. Frankfurt argues that agency requires identification with certain desires such that it precludes the agent's alienation from the desires. Say the commitment consisted in the agreement of second-order desires with first-order desires. But then, the agent *could* be alienated, in that he could disapprove and resist on the third-order level, resulting in the same problem of alienation. So, Frankfurt concludes, wholehearted identification must be such that "when a person identifies himself decisively with one of his first-order desires, this commitment 'resounds' throughout the potentially endless array of high orders."¹⁰

Velleman suspects that Frankfurt is question-begging; decisive commitment sounds too much like "agency" put into other words, which would not help explain where and how the agent participates any more than the standard account did. Furthermore, Velleman does not want to accept a Chisholm-like non-reductive conception of agency, which posits agency as a primitive and denies that it can be reduced to a certain state or motive.¹¹ He suggests that instead of searching for a way to show how the agent identifies with his desires, agency must be defined in terms of a state or motive that *plays the role* of the agent.

The ideal candidate for agent role-player, Velleman concludes after a brief voyage into what he calls commonsense psychology, is the desire to act according to reasons. Wanting to act intelligibly, he claims, is the agent's *modus operandi*; guided by his primary motivating desire, to act in accordance with reasons, an agent provides causal "oomph" to the motivations it considers the strongest. Its participation determines which motivations will influence actions. Thus, the agent-desire adjudicates which motivations, and thus which actions, will win. It is a functional state from which the agent cannot alienate himself; as long as he remains an agent, a person will want to act intelligibly. If he wants to act intelligibly, Velleman continues, he will, because his actions correspond to his identified reasons for acting.

IV. The Desire for Intelligibility

Velleman's account rests on his intuition that rational agents have the desire to do what is intelligible to them in the sense that they could explain it. Furthermore, he states that in giving the account, he is trying to articulate the way most people conceive of agent-caused action, not the way it *really* works. He is open to the possibility that the conception is the same as the reality, but does not insist that that is the case. Thus, when he gives his intuition that desire for intelligible action is the primary motivation of behavior, it is important that his intuition cohere with "what we ordinarily mean when we call something an action."¹²

Is there good reason to believe that humans want to be able to give an account of their behavior, and that their desire for intelligibility is strong enough to adjudicate amongst all other motivations? Or rather, is there good reason to think that this is how

we normally conceive of agency? *Prima facie*, it seems not. What about the kind of person that would describe himself as spontaneous, spur-of-the-moment, by-the-seat-of-his-pants, thrill-seeking, rash, and so forth? He revs up the bike and hits the highway for a joy ride, shucking off the burden of responsibilities despite the probable consequences. “Why did you do that?” I ask him. “I don’t know, just because,” he responds. “I wanted to be spontaneous. I don’t care if it’s irrational.” Is the desire for spontaneity tantamount to wanting to act without being able to give reasons for action? If so, then quite a few motorcyclist types would be barred from entering Velleman’s agent camp.

Velleman might respond that the desire to act spontaneously is not counter to the desire to act with reason. Instead, desire for spontaneity is just another of the myriad motivations that the primary desire moderates. When motorman hits the highway, he is throwing his weight behind one of his reasons to act, namely, his desire to hit the road.

Two aspects of the case do not matter to Velleman. First, it does not matter that the man’s behavior is seemingly irrational. If the man sacrifices his job and home for a joyride, he might be called and call himself irrational. But Velleman does not define agency as the desire to act in accordance with the objectively best reasons (if there are such phenomena). He only insists that an agent act in accordance with reasons. Naturally, he thinks that whichever reasons the agent acts upon are, to his own mind, the strongest ones. Plato argues no one knowingly and freely chooses a seemingly lesser good; Velleman would agree that no *agent* does so.¹³ Second, it does not matter that motorman will not admit that he wants to act in accordance with reasons; whether or not he knows it, his agency entails that he *does* want to act intelligibly.

Grant Velleman his intuited human desire for intelligibility, and even allow that he might be right in citing that desire as the primary motivation that plays the role of the agent and adjudicates among all other motives. There is still a problem. Consider the following propositions, all of which Velleman arguably holds true.

- (1) The desire to act according to reasons is a conscious desire.
- (2) The reasons are reasons identified by the conscious mind.
- (3) Only where (1) and (2) combine to produce action is the action agent-caused.
- (4) Agent-caused action is brought about entirely by the conscious mind.

Take (1). In light of the suggestion that motorman does not admit that he has a desire to act in accordance with reasons, it might seem that even if he does in fact have such a desire, it is not accessible to him. His agent-desire is then unconscious. But conceivably, *every* agent could, like motorman, deny that he has a desire to act intelligibly. Then how would Velleman argue that such a desire, the agent role-player that no one admits having, drives the normal conception of agency? A conceptual framework ostensibly uses conceptual tools accessible (and admissible) to the conscious mind.

I think that this problem may be circumvented by claiming that it is not necessary that an agent experience a desire for intelligibility in those exact terms; if, say, the agent knows himself to want to act freely, that might be enough. Commonsensically, it seems that someone who wants to act freely wants to have some power with respect to options. Having power with respect to options entails siding with one course rather than another, and it is very hard to understand how someone would side with one or the other course, if not upon reason. Even flipping a coin – heads, left; tails, right – is giving the self a

reason to go one way rather than the other. Thus, acting freely entails acting with reason, and the desire to act freely entails the desire to act with reason. (1), then, seems tenable, couched in some form.

Now for (2). Velleman writes:

...reasons for a particular action are considerations by which the action could be explained and in light of which it would therefore make sense...when a desire appears to provide the strongest reason for acting, then the desire to act in accordance with reasons becomes a motive to act on that desire, and the desire's motivational influence is consequently reinforced....(479)

From this passage, it seems clear that Velleman takes the reasons picked out by (1) to be conscious, identifiable reasons. When the agent (or rather, that which plays the role of the agent) picks out certain reasons, those reasons become the reasons for the subsequent action. Thus, the identified reasons, mediated by the agent, direct action. If the agent had picked out other reasons, then those reasons would be the powerful ones, and the action might well be different. Thus, Velleman holds (3) and (4), which just follow from (1) and (2): all of the important functioning that leads up to an action takes place in the conscious mind, which is to say that the functioning is cognitively accessible to the agent. Now I will begin to call Velleman's demand for agential cognitive access into question.

V. The Unconscious Mind

According to Freud, the conscious self is only the tip of the iceberg when it comes to what goes on in the human mind, and in particular, to what motivates human behavior. Without getting into too much detail, consider his theories about the subconscious, which posit that much of human behavior is explicable in terms of unconscious motivations. I disinter Freud, and could also point to Jung, Husserl and others, not as their disciple but only to direct our attention to the huge impact their ideas have had. Velleman wants to

deal with “our” conception of agency; it seems, then, that we are obliged to tip “our” hat to the popular psychology of the past century and a half.¹⁴ Fortunately, however, talk of the unconscious does not require explicit recourse to Freudian theory. I will argue that there are good reasons to believe that many action-producing desires and motivations are unconscious.

Before the argument, two clarifications.¹⁵ First, I am going to assume that there is always a “main motivator” that prompts action, which may either be conscious or unconscious. Note that henceforth, when I refer to a main motivator as conscious or unconscious, I am not referring to the agent’s being aware (conscious) or unaware (unconscious) of the motive, but instead to the main motivator’s status. If a motivator is conscious, then the agent has the capacity to know of it; if it is unconscious, the agent is not able to access it.¹⁶ There is a fact of the matter about 1) what the main motivator is and 2) which realm it belongs to.¹⁷

Second, I will assume the following about Velleman’s stipulation that the agent is played by the “desire to act in accordance with reasons...to do what’s intelligible to them.”¹⁸ Acting intelligibly means having self-understanding, where self-understanding requires a correct identification of one’s main motivator. Therefore, to spell out what Velleman merely implies: self-understanding is only possible where the main motivator is conscious. Now I turn to the argument: there is, *contra* Velleman, reason to believe that agential main motivators are, at least some of the time, unconscious.

Take the following example. Late one night I sit in my room weighing the pros and cons of getting a cup of coffee. On the one hand, I have a paper to write and caffeine would make it easier to stay awake. On the other hand, too much coffee is bad for my

health and I would feel guilty, get dehydrated, and so on. In the end, I decide against coffee and opt for water and an invigorating stroll instead, which happens to take me past the coffee shop.¹⁹ Smelling the coffee and witnessing the energized swillers, I pause on the sidewalk contemplating how important it is that I write a good essay tonight, and go in to buy some coffee.

What happened? My desire for coffee is subject to the primacy of the immediate. The desire for coffee increases in the presence of the object, and so does the corresponding reason to get it. Say that I am not really aware of the smell and sight of the coffee shop – that is to say, I am not deliberately concentrating on the sensory input. Nevertheless, my desire responds to the proximity by increasing in intensity. The phenomenon is probable because it happens all the time; advertising depends on subliminal messages to increase appetite without arousing conscious awareness and resistance to the craving.

When I buy the coffee, I identify the essay's importance (and the coffee as aid to that effect) as my reason for acting as I do, despite resolving not to buy coffee minutes before. I am misidentifying the real motive; my real reason for acting is unidentified, while the reason that I cite is a false front. I contend that at least on some occasions, attempts at introspection do not yield the truthful reasons for acting, but only red herrings. This may be, as in my case, due to self-delusion.

Consider V-man again. An unconscious desire leads to an unwitting decision, which prompts action. V-man cites the wrong reasons, the false fronts: he has been wronged, the friend has misbehaved, and so forth. Velleman claims that V-man is not the full-blooded agent of his behavior because his actions are not his own in the right sense.

I may conclude that desires of mine caused a decision, which in turn caused the corresponding behavior; and I may acknowledge that these mental states were thereby exerting their normal motivational force, unabridged by any strange perturbation or compulsion. But do I necessarily think that I made the decision or that I executed it? Surely, I can believe that the decision, though genuinely motivated by my desires, was thereby induced in me but not formed by me; and I can believe that it was genuinely executed in my behavior but executed, again, without my help (464).

Grant Velleman, if he likes, the right to exclude behavior brought about by unconscious main motivators from his inventory of agent-caused action. The next section will show that this exclusion of the unconscious from his agent realm might lead him into difficulties.

VI. Velleman's Problem of Over-Exclusion: Agent Unaware

Of the agent's actual progress to action, Velleman is vague, and it requires some guesswork to flesh out his position. At his most informative, he writes, "when a desire *appears* to provide the strongest reason for acting, then the desire, and the desire's motivational influence is consequently reinforced."²⁰ He also describes the process the following way: "for potential determinants of behavior to be critically reviewed, to be embraced or rejected, and to be consequently reinforced or suppressed [by the motive that plays the agent]."

Consider the possible meanings of a desire *appearing* to provide the strongest reason. What sort of attention must the agent lend to the process in order for the desire to *appear* to him one way or another? There are two main alternatives. (1) He is fully aware of his desires, surveying them and inscribing them in a mental pro and con list. That is, they almost literally *appear* to him. (2) He is only aware of his desires in a

fleeting, under-the-surface way. Given enough time he could label them, but in throwing his weight behind one of them, he is not necessarily aware in a fully blown sense.

The first reading coheres with Velleman's description of the process as a "critical review," which sounds formal and deliberate, but I do not think that Velleman could plausibly defend it. The second option, it seems, better describes the majority of actions that Velleman would, or should, call agent-caused. Most actions require quick decisions. I am not referring to reflex-type actions like veering off the road to avoid an oncoming vehicle, which would rely upon conditioned impulses, but rather of real decisions that are made *almost* without thinking. To take stock of myriad facts and corresponding desires and motivations, a person does not – in deliberation and full awareness – make a pro and con list, but instead throws his weight behind a motivation quickly. Given time, he would be able to spell out his motivations, but does not do so at the moment of action. Think of a clocked chess player who, under time pressure, must move quickly: he relies upon intuitions that are well-packed with reasons, but he is not completely aware of those reasons as he makes the move. If he tries to act in the way (1) describes, he will always lose on time.

Given these considerations, it seems fair to read Velleman's use of "*appearance*" as cohering with some version of (2). The agent need not be fully aware in acting.²¹ He would, given enough time and desire, be able to identify his competing desires and account for why he threw his weight behind a certain motivation, but he is not required to do so in the moment of decision. I think that Velleman would agree; in any case, he would have to give a reason for disagreeing with what seems to be the more plausible reading of his description.

Now consider a case of action with unconscious main motivator X. The action takes place without the actor enumerating reasons for and against acting. He is unaware of all of his competing desires, some conscious and some unconscious, but nonetheless he throws his weight behind unconscious X, much like V-man does.²² Velleman does not allow such an action to count as agent-caused, because “the decision, though genuinely motivated by my desires, was thereby induced in me but not formed by me.”²³ Consider V-man’s evil twin VE-man. VE-man consciously decides to break with his friend and makes the same pretence of irritation in order to do so; Velleman calls him an agent. What distinguishes V-man from VE-man? Simply this: VE-man is aware of his decision, and V-man is not.

If *awareness* distinguishes agents from mere actors, however, Velleman seems to have a problem. We agreed that (1) is a practically insensible demand: decisions, in the real world, are made a) quickly and b) are based on intuitions, not on pro/con lists. It seems, then, that if Velleman is demanding that all agential actions be decisions made in full *awareness*, he would be excluding a certain range of action from the agent-caused category. I would go so far as to claim that he would be excluding *most* actions: all but the getting-engaged or buying-a-house variety.

In order to avoid such a drastic exclusion, Velleman needs to provide a way to distinguish actions prompted by motivations of which the agent is unaware, but which are (or would be upon reflection) available to the conscious mind, and actions prompted by unconscious motivations. Recall the coffee case. My reason for buying coffee is the following: I throw my weight behind my coffee craving. There are two scenarios: in the first, I am unaware of my conscious craving. In the second, I am unaware of my

unconscious craving.²⁴ How could Velleman allow the first scenario to be agent-caused, and not the second?

His best possible reply is to add a further stipulation to his account of what full-blooded agency requires. Call it the Accessibility Clause, or AC.

(AC): The agent, played by the desire to act in accordance with reasons, must have cognitive access to the reasons it acts upon.

Cognitive access, I take it, is a success term indicating that the agent would, if he tried, be able to correctly match his desire with the action it prompts.²⁵ Because the unconscious is not available to conscious scrutiny, the agent cannot access unconscious desires. With this requirement, it seems, Velleman effectively bars unconscious reasons from serving as agential main motivators. If an agent throws its weight behind an unconscious reason, then he would not have the *potential* to identify that reason as his main motivator. He would not have the cognitive access (AC) requires, and would therefore fail to be an agent. The clause would allow cases of action prompted by unaware but conscious coffee craving to count as agent-caused, because the coffee drinker would, if allowed to settle into introspection, recognize that the craving was the true reason for action. The unconscious coffee craver, on the other hand, might be able to guess that his craving caused his action, but his success would not reflect cognitive access to the desire.

(AC), I contend, is an *ad hoc* tautology. In order to allow actions prompted by (conscious) motivators of which the agent is *possibly* unaware, and bar (unconscious) motivators of which the agent is *necessarily* unaware, Velleman needs to do more than simply claim he doesn't allow the latter. I see little convincing reason to believe that

(AC) is a justifiable stipulation. However, in the interest of fairness to hypothetical Velleman, I will suggest a possible argument for (AC).

When an agent is unaware of a conscious motivator, Velleman might argue, the motivator is lurking just below the surface of deliberation. Although the decision is made quickly, or intuitively, without a full-scale critical review, the agent is still capable of identifying the motivation, such that she is able to throw her weight behind it without “it” being a great mystery. One might compare such conscious unawareness to seeing an object in one’s peripheral vision as opposed to directly. The viewer has a certain awareness of it, such that if he turns to the side, he might easily concentrate on it. There is, Velleman might continue, no reason to think that an agent is even *able* to throw her weight behind an *unconscious* motivator in this kind of way. Such a motivator is not just lurking below the surface of awareness, still readily accessible, but is instead buried somewhere at the time of the decision. It is hard to see how an agent could *mediate* between such desires and subsequent actions, at all. It seems, rather, that unconscious desires necessarily boil into actions without the agent’s endorsing them in the least, in the manner of V-man’s eruption. Thus, (AC) prevents V-man from counting at an agent: because his decision to rupture is unconscious, he has no potential to cognitively access the motivator at the time of his action, and thus no power to endorse it. If it had been conscious but below awareness, then even though he had not consciously done a full-out critical review of the situation, he would have at least been throwing his weight behind a motivation appearing *somewhere* in “his mind’s eye.”

I will first reply to Velleman by further arguing for the prevalence of unaware motivations in cases of agency, such that the agent *must* be unaware of the motive in

order to do the action at all. If she turns her head, so to speak, to focus on the peripheral motivator, she will not pursue the same course of action. Such unawareness, I will argue, is no different from the necessary unawareness of unconscious motivations. In the next section, I will build my own (tentative) account of how an action prompted by an unconscious main motivator can be a case of full-blooded agency, responding to Velleman's doubt that an agent can in fact throw her weight behind an unconscious motivator.

Consider the following example. A man sits at a bar. Call him Barman. He is alone, and wants to change seats in order to sit by an attractive woman at the other end of the bar, so he does. When asked why he moved, he gives a guilty smile and shrugs; it turns out that Barman is happily married. Let's suppose that if he had conducted a full-scale review of his motives, he would have realized that his purpose was contrary to the way that he prefers to live; he does not really want to pursue other women. In making his move, however, he was successfully averting his attention from his main motivator. Clearly, Barman is an agent with respect to his shifting chairs. Just as clearly, his main motivator is to pursue the woman (that might not be clear, but let's suppose that it is). He is *necessarily* unaware of his main motivator; if he were aware of it, he would not have shifted chairs.

Is Barman's main motivator conscious or unconscious? That is, does he have cognitive access to it or doesn't he? I do not think that it matters. He throws his weight behind the main motivator without being aware what his main motivator is. He would, as I did in the coffee case, identify false front reasons if asked to explain his action. Or if he is particularly savvy and honest about his own predilections, he might give a guilty smile,

realizing or guessing at his real motivation, and then move back to his chair (especially if it's his wife who shows up to pose the question). In either case, he is endorsing his main motivator in the sense required for agency without its being consciously identified. There are many examples of such behavior, and it would seem arbitrary and improbable to rule them out as cases of full-blooded agency.²⁶

Recall (AC). Given that such examples show that the agent will not act if she, in full awareness, identifies the main motivator, it seems that Velleman must further explain *why* cognitive access to motivators is crucial to agency. Why is such access important, if it is never (in such examples) taken advantage of, if taking advantage of it would *thwart* the action? I again contend that (AC) is simply a convenient *ad hoc* clause aimed at ruling out unconscious agential motivators without good reason to do so.

VI. The Unconscious in Action

Ascribing unconscious action-prompting desires to full-blooded agents is, I gather, not the norm for philosophers of action. This may be, in part, because Freud is widely discredited these days, and he is the main point man for the topic. More plausibly, the unconscious is overlooked because there are myriad problems in constructing an account for its functioning. Because I have, in this essay, tried to force Velleman into admitting that unconscious motivations can prompt agential action, it seems that I must sketch out a plausible account of how such prompting would take place. The main problem – suggested in Velleman's hypothetical reply – is the following. The pre-analytic intuition about agency is, as Velleman puts it, "What makes us agents rather than mere subjects of behavior...is our perceived capacity to interpose

ourselves into the course of events in such a way that the behavioral outcome is traceable directly to us.”²⁷ If, however, the main motivator is unconscious (or, conscious but unaware, since I wish to equate the two), how does the agent mediate? It seems that without deliberation and awareness, the agent is simply the vessel of desires without actually endorsing them in a traceable way. If he does not recognize the unconscious desires, then how can he throw his weight behind him in the way that Velleman requires?

My rough sketch of how the unconscious acts takes its cues from an article by Ruth Weintraub.²⁸ She argues that a conscious action has only conscious beliefs and desires as motivations; unconscious beliefs and desires *do* have causal efficacy in agent full-blooded decision-making, but only by triggering conscious “counterparts,” under the guise of which the agent acts. We couldn’t say, however, that counterparts are the real motivations, since in the absence of the unconscious prompters they would not be activated. This account harkens back to my coffee craving case: my craving corresponds to the conscious reason for acting that I identify, namely, that I write a better essay. If I did not have the craving, I would not drink the coffee. If I did not have an essay to write, I would probably drink coffee rationalized under some other guise. My conscious reason does not truly account for the action, but serves merely as a false front.

Weintraub points out that the unconsciously suspicious husband who calls home is not directly motivated by his suspicions, but instead by a conscious rationalization (to ask if he forgot his portfolio). The conscious reason to call is the result of the unconscious reason’s influence. In throwing his weight behind the counterpart reason, I argue, the husband is also endorsing the unconscious reason for acting. To deny that this is the case would be to make the mistake of dividing the agent into two selves, the

conscious and unconscious. As Weintraub writes, “The mistake is removed by viewing the conscious and the unconscious as belonging to one agent, and this allows us to say that the agent is mistaken about his unconscious beliefs...we utilize the interplay [between the conscious and unconscious] to uphold the unity of the agent.”²⁹

Unconscious motivators belong to the agent, in some cases, just as much as conscious motivators do. The fact that they work via conscious counterparts, false fronts, does not matter to the agent-causal nature of the action. In the next section, I will show that certain unconscious motivators do not count as agential action prompters, and explain why.

VII. Too Much Action?

Recall Velleman’s initial reason to revise the standard account of agency: the old description of “what happens when someone acts” does not discriminate between arm-raise₁ and arm-raise₂ – the first electrically caused, and the second “chosen” – and must therefore be refined. While conceding Velleman his serviceable definition of an agent – played by the desire to act in accordance with reasons – I have pressed him to admit that he is wrong to limit agent functioning to conscious motivators. I provided a rough account of how an agent might continue to fulfill Velleman’s requirement while throwing his weight behind an unconscious motivation, under the guise of its conscious counterpart.

The crucial question is the following: does my revision of Velleman’s account also accomplish the original goal of discriminating between arm-raise₁ and arm-raise₂, or between the willing and unwilling addicts? I contend that it does.

Consider, first, the examples that I used to illustrate my points. Barman exemplifies a case of full-blooded agency in which he obfuscates his real motivation.³⁰ It is available to his conscious mind, but if he were aware of it, he would refrain from acting. He might cite a false front reason as I do in the coffee craving example, which is also a case of full-blooded agency. In citing a false front as a reason to buy coffee, I am still prompted by my agent-playing desire to act intelligibly, even if I identify the wrong reason as my main motivator. As an agent, I consist of both unconscious and conscious desires, where my unconscious desires are no less my own. Velleman's V-man, I hold, is no different from the other two cases. He has an unconscious desire to act a certain way, throws his weight behind a counterpart reason, and in doing so endorses his unconscious desire to break with his friend.

If a widened range of action is permitted as agent-caused, do the unwilling addict and the electrode arm-raiser also rank as agents? After all, they will be prompted to act by motivations of which they are unconscious. My response is that they are not cases of agency, because they do not throw their weight behind their motivational desire in the appropriate way. Consider the unwilling addict. He says, as we have seen, that his drug desire seems foreign to him. When he does drugs, he does not consciously endorse his action. The requirement for Velleman's agent, which I uphold, is that the agent be motivated by the desire to act in accordance with *some* reason. Even if that reason is a counterpart for an unconscious desire (think of the husband calling about his portfolio), it is still an endorsed reason to act. Likewise, knee-jerk reflexes, dancing on electric carpets, being prompted by wires to raise an arm, and so forth, do not count as agent-caused actions because the agent is not endorsing a main motivator, either conscious or

counterpart, that makes his actions intelligible to him. That, I think, is a legitimate way to continue to exclude intuitively non-agential actions, while including actions prompted by unconscious motivations.

VIII. Conclusion

To give a thorough account of what it means for an agent to act, confronting myriad difficulties that I have not even mentioned, would be an unpalatably momentous task. In my installment, I hope to have revealed that Velleman's account is incomplete. He needs either to expand his description of agency, allowing the unconscious to motivate agent action, or to better explain the omission. More specifically, he needs to address my main attack by thrusting a warranted wedge between actions prompted by unconscious motivators and actions prompted by conscious motivators of which the agent is unaware.

References

- Chisholm, Roderick. Person and Object: A Metaphysical Study. London: G. Allen & Unwin, 1976.
- Davidson, Donald. Essays on Actions and Events. Oxford, Eng.: Oxford UP, 1990.
- Frankfurt, Harry. The Importance of What We Care About. Cambridge, Eng.: Cambridge UP, 1988.
- Rubinstein, Benjamin B. "On the Psychoanalytic Theory of Unconscious Motivation and the Problem of its Confirmation." Noûs 14.3 (1980): 427-442.
- Velleman, J. David. "What Happens When Someone Acts?" Mind 101 (July 1992): 461-81.
- Watson, Gary. "Free Agency." Journal of Philosophy 72 (Apr 1975): 205-220. Reprinted in Free Will. Ed. Gary Watson. Oxford: Oxford University Press, 1982. 96-110.
- Weintraub, Ruth. "Unconscious Mental States." Philosophical Quarterly 37 (1987): 423-432.

¹ Velleman, J. David. "What Happens When Someone Acts?" *Mind*, July 1992.

Frankfurt, Harry. *The Importance of What We Care About*, 1988.

Watson, Gary. "Free Agency" in his *Free Will*, 1982.

² Velleman, 1.

³ Frankfurt, Ch. 6.

⁴ Velleman, 1.

⁵ Frankfurt, ch. 1.

⁶ Frankfurt, 60. Idea from Terence Penelhum, "The Importance of Self-Identity," *Journal of Philosophy* LXVIII (1971).

⁷ Frankfurt, 17.

⁸ Frankfurt, 21.

⁹ An explicit discussion of the kinds of "wanting" is beyond the scope of this paper, but I am assuming that there are at least two kinds of desire. The first is isolated from other considerations, for instance, my desire to eat the entire cake. The second is an all-things-considered desire: given that I will ruin the birthday party, feel extremely sick, and tarnish my reputation for moderation, I do not want to eat the entire cake. Thus, the unwilling drug addict might want drugs in the first sense but not the second.

¹⁰ Frankfurt, 21.

¹¹ Chisholm, Roderick. *Person and Object: A Metaphysical Study*, 1976.

¹² Velleman, 466.

¹³ I do not wish to come down on either side of the complex debate on whether there is such a thing as *akrasia*, nor do I wish to imply that Velleman takes a position. I am merely pointing to the distinction between objective (universally recognized) and subjective (first-person only) reasons for acting; Velleman does not require that intelligible reasons for acting be universally accessed or universalizable for the action to be one of full-blooded agency.

¹⁴ "My aim is to explain...action as we conceive it to be..." (Velleman, 466).

¹⁵ In presenting my case, I will advance a few discussion-facilitating assumptions that Velleman does not make himself. Though he would not, I think, have reason to dispute my groundlaying, it would be of slight import if he did: the assumptions are not critical to my argument.

¹⁶ I am not going to take a strong position on whether hypnosis provides a means of accessing unconscious desires. I suspect that it does, but articles such as Rubinstein's shed doubt on the accuracy of such "discoveries." I will assume that if a main motivator is unconscious, then at least at the time of the action, the agent cannot know of it. Whether she is able to "discover" it later, in psychotherapy or trance-like introspection, is a rather irrelevant issue that I won't discuss.

^{16a} Rubinstein, Benjamin B., "On the Psychoanalytic Theory of Unconscious Motivation and the Problem of its Confirmation," *Noûs*, Vol 14, Issue 3, (1980), 427-442.

¹⁷ Why believe there is a "main motivator" corresponding to every action? One might think that myriad desires, together with many pertinent facts, combine to motivate action. I agree, but it seems plausible that one of those desires looms larger than the others as a prompting force. Because the assumption facilitates my discussion without being crucial in its own right, I will move on.

¹⁸ Velleman, 478.

¹⁹ It's at least possible that I unconsciously engineered the route to take me past the coffee shop, but I won't insist.

²⁰ Velleman, 479, my emphasis.

²¹ I need to spell my intuition out further. Why wouldn't it be possible for an agent to be aware of the main motivation, just one fact and not time-consuming in the manner of a pro and con list? I would claim that in order to be fully aware of a main motivation, an agent would have to also be aware of all, or at least many, of the competing desires/motivations that proved weaker than the chosen motivation. It would seem paltry awareness indeed to be able to specify *what* the main motivation is, but have no immediate defense of *why* it is as such. In order to know *why* the main motivation is what it is, there must be deliberate scrutiny of the other options; according to my argument, in most instances there is not time to conduct such a deliberate (and leisurely) perusal.

²² In section VII, I will provide a possible account of how this happens.

²³ Velleman, 464.

²⁴ It might seem a bit far-fetched to suppose that I could be unaware of a craving, either conscious or unconscious, especially when I am in the process of acting upon it as my main motivator. Isn't that like being unaware of pain, even as one screams? However, I would guess that for someone, highly distracted by other goings-on, it might be possible. In any case, I could pick a motivation other than craving, if I wanted to.

²⁵ Cognitive access, like epistemic access, does not require infallible functioning: if it is at least very probable that the agent would be able, upon reflection, to correctly identify his conscious reasons for acting as he does, then he has cognitive access to those reasons.

²⁶ Procrastination is another kind of example; the main motivator is recognized, but the reasons NOT to act are obscured. Say I have a big paper due, but I decide to go to a concert instead. I focus on the pleasure of music, not on the pain of last-minute writing.

²⁷ Velleman, 465.

²⁸ Weintraub, Ruth. "Unconscious Mental States," *The Philosophical Quarterly* Vol. 37 (1987), p. 423-432. Though Weintraub concludes that the unconscious is of little importance to decision theory, she makes some helpful assertions along the way.

²⁹ Weintraub, 429.

³⁰ To describe how this obfuscation, also called repression, takes place would require more knowledge and time than I currently have. It would, however, be an interesting topic to return to.