

## Symposium on *How We Get Along* Responses to Critics

J. David Velleman

New York University  
jdvelleman@nyu.edu

*How We Get Along* begins with a sob: you are crying uncontrollably. But then you get a hold of yourself and settle down to have a good cry. I ask: what has changed? Whatever it is, it's what makes the difference between behavior that is out of your control and therefore not an action of yours, on the one hand, and behavior that is in your control and therefore an action, on the other.

I contend that whereas the uncontrolled crying is simply the manifestation of hurt or grief, the controlled crying manifests something further, namely, your awareness of the hurt or grief, and your resulting perception of crying as what it makes sense to do. Had you been unable to think of what you might be crying about, you would have asked yourself "Why am I crying?", and your tears would have tapered off. Because you do know what you're crying about, however, you keep crying, but now with the concurrence of that self-understanding, which transforms your crying from mere behavior into an action.

What differentiates action from other behavior, then, is that it is guided by a disposition to do what makes sense to you — a disposition, I suggest, that arises from a drive on your part toward self-understanding. The drive toward self-understanding is not the primary motive of your actions; it is rather a second-order motive with the respect to the manner in which you act on other motives. Grief is your primary motive; the drive toward self-understanding can only oppose or reinforce your grief.

When I say that the drive toward self-understanding is what differentiates action from mere behavior, I mean that it is responsible for the features that are distinctive of action: the way we know what we're doing when we act; the way our actions are up to us, in the sense that we choose between alternatives; the way we commit ourselves to future actions in advance; and the way that our actions are guided by reasons. These features were absent from your passive crying. You may not have realized at first that you were crying; it wasn't up to you whether to cry; and although something had caused you to cry, you weren't crying for that reason. Then you saw what there was to cry about and you had a cry for that reason. It was up to you whether to have a cry, and when you went ahead, you knew what you were doing.

I initially compare this case to the crying of a "method" actor, who actually feels the emotion fictionally felt by his character and then channels it into behavior that is "in character" because it makes sense as the way his character would cry. I then complicate the comparison by pointing out that you are not following a script: you are therefore more like a improvisational method actor, feeling the relevant emotions but making up your part of the whole episode, still within the constraint of acting in character, lest the improvisation fail to make sense.

Finally, I introduce one more complication. You are not acting out a fiction, working up reactions to fictional circumstances with a method actor's tricks and then acting them out in a character that's fictional. You are improvising your actual self, reacting to your actual circumstances, in part spontaneously but also guided by how it makes sense to react; and doing what actually makes sense for you to do given those reactions. Your "good cry", I claim, is an authentic improvisation, fueled by motives that are both realistic and real, but also guided by your perception of how it makes sense to feel and to manifest your feelings.

In the remainder of the book, I develop this theatrical analogy into a theory of individual and collective practical reasoning — individual in that each actor must do what makes sense for his character to do, collective in that what makes sense for each depends on what the others are doing. Like a troupe of improvisational actors, I suggest, members of a community or society must develop shared scenarios on which their joint improvisations are based, as variations on a theme. These scenarios are their shared way of life. And I suggest that their way of life will take on the shape of a morality because of the constraints imposed by exigencies of collective improvisation.

### *Response to Phil Clark*

My theory of action has a metaethical implication. Because self-understanding is a second-order aim of every action, it provides a normative standard for action as such, irrespective of any particular action's first-order aims. Self-understanding is a normative standard for action as such in the same way as truth is a normative standard for belief as such, irrespective of what any particular belief is about. If someone's movements aren't aimed at intelligibility, in addition to his first-order aims, then he isn't acting. If he is acting, then his movements are aimed at intelligibility, and so the standard of intelligibility applies. Thus, the standard applies to any and every action.

As Phil Clark says, this theory of action and its companion metaethics are meant to implement what I call the Kantian strategy. Kant tries to derive an objective practical norm from a standard that is inescapable from the practical perspective, in that anyone must be committed to it in deciding to act. I do likewise, though I call it an aim and I identify it as intelligibility rather than universalizability. These differences are minor, but Clark thinks that there is a more significant difference. For he thinks that I fail to derive an objective practical norm from my theory.

The problem, according to Clark, lies in the specification of the aim that is inescapable from the practical perspective. It is not the aim to do intelligible things; similarly, the aim of belief is not to believe true things. A rational believer doesn't go around accumulating as many true beliefs as he can. The constitutive aim of belief is to believe things only if they are true, but their being true is not sufficient; they also have to bear on a topic on which he has reason to hold an opinion. Similarly, a rational agent doesn't go around doing as many intelligible things as he can. The constitutive aim of action is to do things only if they make sense, but their making sense isn't sufficient; they also have to be things that he has some motive for doing.

Clark's objection is that aiming to do things only if they make sense isn't inescapable from the practical perspective. An agent could have a specific aim, with respect to each action, to do *it* only if it makes sense. Clark compares these aims to the aims of finding particular things when one is looking for them. Not only would it be ridiculous to aim at finding lots of things, as it would be to aim at doing lots of intelligible things or believing lots of truths; it would even be odd to have the general aim of finding things if one is looking for them. All one needs,

when searching for something, is the aim of finding that particular thing. And that specific aim is sufficient to constitute one's activities as a search. So long as one is trying to find something in particular, one is engaged in a search, whether or not one has the general aim of finding things that one looks for.

The upshot, Clark contends, is that the aim of being intelligible in doing a particular thing should be sufficient, according to my theory, to constitute that thing as an action. Just as a belief that  $p$  requires no more than the aim of believing  $p$  only if  $p$  is true; so performing action  $a$  should require no more than the aim of doing  $a$  only if  $a$  makes sense. Hence action requires no universal aim that is present in every case; it requires only particular aims, specific to the actions undertaken by particular agents on particular occasions. All that can be derived from those aims are particular standards, different standards for different agents and different actions. Objectivity has therefore not been attained.

Clark illustrates his argument with an example. Imagining himself cooking shrimp Alfredo, he says, "[M]y aim in making shrimp Alfredo is to understand my behavior in respect of making shrimp Alfredo – my making it, or not making it, or making it this way or that, as the case may be. Or to put it another way, my aim is to exhibit shrimp Alfredo related behavior that I can understand." Here Clark assumes that the aim of understanding his shrimp-Alfredo-related behavior is not an instance of the more general aim of understanding whatever he does. He assumes, in other words, that the relevant aim is not the result of substituting "making shrimp Alfredo" for  $x$  in an aim with the content "For all actions  $x$ , to do  $x$  only if I understand  $x$ -ing", since the latter would be a general aim of the sort that he denies my theory requires.

Clark believes that his description of the case follows from my analysis of action, which he paraphrases like this:

**(Analysis of Action)** For  $A$ 's  $\phi$ -ing to be an action is for  $A$  to  $\phi$  in an attempt to behave in a way that makes sense as regards  $\phi$ -ing.

But this is not my analysis. My analysis, if transposed into Clark's language, would not include the last three words, "as regards  $\phi$ -ing". It would say merely "For  $A$ 's  $\phi$ -ing to be an action is for  $A$  to  $\phi$  in an attempt to behave in a way that makes sense." Given that  $A$  is  $\phi$ -ing, of course, his attempt to behave in a way that makes sense becomes an attempt to  $\phi$  (or to stop  $\phi$ -ing) in a way that makes sense — but only given that he is  $\phi$ -ing or has at least decided to  $\phi$ . Before he decided to  $\phi$ , intelligibility with respect to  $\phi$ -ing was not his aim. Indeed, my analysis is that in deciding to  $\phi$  rather than  $\chi$  or  $\psi$ ,  $A$  had the aim of doing whichever one made more sense, an aim that was not specific to intelligibility with respect to  $\phi$ -ing.

Why does Clark assume that my analysis posits the more specific aim? I don't think that this assumption flows from the analogy between action and belief. Although my belief that  $p$  has the aim of accepting  $p$  only if  $p$  is true, that specific aim results from substituting  $p$  for  $x$  in a general aim with the content "For all  $x$ , believe  $x$  only if  $x$  is true". What favors Clark's assumption is rather the case of searching, where the aim of finding  $c$ , given that I am searching for  $c$ , is not an instantiation of a more general aim. That is, I don't have the aim of finding whatever I search for. Actually, the case of searching differs from that of believing or acting in an additional respect. In the latter cases, my aim has conditional content, and so it can be satisfied, on the one hand, by believing what's true or doing what I understand, or on the other hand, by not believing what's false or doing what I don't understand. In the case of searching however, the condition attaches to my having the aim, not to its content. If I am merely considering whether to look for  $c$  without yet having decided or started to look for it, I

may not have any aim at all with respect to finding  $c$ , not even an aim conditional on looking for it, and so I have no aim that can be satisfied by my not looking for it. By contrast, as soon as the question arises whether to believe  $p$ , I have the aim to believe  $p$  only if  $p$  is true; and as soon as the question arises whether to do  $a$ , I have the aim to do it only if it will make sense.

The aim posited by Clark wouldn't be sufficient to constitute behavior as action, because it would not, for example, play any role in a choice among alternative actions, one of the roles that a theory such as mine requires it to play. Nor would it be explicable as a reflexive application of our general aim to understand the world around us. That's why my theory posits a more general aim that is indeed common to all actions.

### *Response to Tamar Schapiro*

Tamar Schapiro takes on the book's organizing metaphor, which is a special case of improvisational acting, in which the agent improvises the role of being himself — an improvised self-enactment. In order to be authentic, the self-enacting agent must portray who he really is; but he can do so by really being who he portrays. As Schapiro sums it up:

To improvise your own character well, you have to actually inhabit your character as far as possible, being yourself even as you create yourself. You have to actually have the thoughts, feelings, and motives that you enact, and your way of enacting them has to actually reflect or express the person you are making yourself into.

"[T]he idea," in other words, "is that we have to bring ourselves into accord with our enacted conceptions of ourselves, even as we bring our enacted conceptions of ourselves into accord with ourselves."

This process of reciprocal self-conception and self-enactment results from our occupying two roles simultaneously, the role of improvisational actor and the role of audience to the improvisation. The audience wants to understand the actor's performance, and the actor suits his performance to what the audience will understand, because he is, after all, identical with the audience who wants to understand it. If a self-enactor is to avoid inauthenticity, however, he has to *be* the character, not just act like him, and he has to understand who he *is*, not just who he's acting like. He can do this by enacting traits and attitudes that he has anyway; but he can also do it by acquiring traits and attitude, for example, through the well-known attribution mechanism, by which his inchoate reactions will crystalize into particular attitudes as he manifests them in the corresponding actions.

Schapiro worries that my theatrical metaphor mis-describes the phenomena of ordinary agency. Her worry is that our primary interest as the audience to a performance is not merely to understand it; our primary interests are aesthetic.

I agree. My analogy between agency and theatrical acting is not meant to be perfect. The actor and audience of a theatrical performance have many interests not shared by an agent with respect to his own behavior.

Even so, I think that Schapiro underestimates the theatrical audience's interest in understanding. She says that the audience is primarily interested in escaping into the world of the story through the suspension of disbelief. But whether they can suspend disbelief depends on whether the characters are believable, which depends on whether the actors' behavior is understandable in light of their characters' traits, attitudes, and circumstances. In a scripted drama, of course, the believability of the action has already been secured by the author: the

actors are responsible only for their gestures, postures, facial expressions, tone of voice, tempo, and so on. But, again, these features of their performance make their characters believable by making sense in light of the attributes and actions that the author has assigned them. An actor's vocal, facial, and bodily expressions have to be intelligible as manifestations of the motives and emotions behind his character's actions. That's why an actor often asks the director "What's my motivation?"

In any case, the analog of human action in my argument is not scripted drama but improvisational theater, where the actors are responsible for the action itself, not just the manner in which they enact it. They're responsible for what they say, not just how they say it. And in this case, intelligibility is more salient as an interest for both actor and audience, because the risk of incoherence is greater. As in the case of scripted drama, intelligibility subserves other interests that aren't often operative outside the theater — humor, for example — but it remains a pre-condition for satisfying those interests.

Schapiro interprets me as drawing an overly close analogy between human agents and an audience to their actions, so she raises the objection that agents have no interest in achieving the suspension of disbelief with respect to themselves. I not only concede this point; I positively assert it, since authenticity in action requires forming beliefs that will be true of oneself — "in character" with respect to one's actual character, in a broad sense that includes one's occurrent attitudes as well as enduring traits. Suspending disbelief would open the door to deception — in this case, self-deception — which is hardly conducive to autonomous agency. As Schapiro herself puts it "[T]he idea is that we have to bring ourselves into accord with our enacted conceptions of ourselves, even as we bring our enacted conceptions of ourselves into accord with ourselves."

That people are strongly motivated to bring themselves into accord with their enacted conceptions of themselves has been demonstrated time and again by social psychologists over the course of more than fifty years. It is perhaps the most robust finding of social psychology, since it is evidenced by such phenomena as the well-known cognitive-dissonance effect. That effect is not, as some people think, a response to discordant beliefs; it is a response to discord between one's attitudes and one's behavior, as manifested in the way people come to believe things that they have been induced to assert, or to want things that they have been induced to choose. Psychologists have also demonstrated that people are strongly motivated to bring their behavior into accord with their self-ascribed characters — for example, by manifesting emotions that they have been induced to attribute to themselves. Of course, the researchers who demonstrate these processes do so by manipulating subjects into believing things of themselves that aren't true — at least, not yet — and acting out attitudes that they don't yet have. But the processes don't come to light when they are operating normally. Normally, they lead people to bring their characters, self-conceptions, and behavior into genuine accord, thereby closing any gaps through which the processes themselves might show. Experimental manipulations are therefore necessary to uncover them.

This and other self-consistency phenomena do not require self-absorption, as Schapiro supposes. On the contrary, they are the result of largely unconscious processes. If they weren't, we would already have known about them without the help of social psychologists. My view is that practical reasoning is the conscious tip of this unconscious iceberg. In order to illustrate how it can work without self-directed attention, I quoted an extended passage in which a philosopher expressed thoughts that were ordered not logically but *psychologically*, so as to

express an intelligible process of thought [pp. 20-21]. As the author wrote this passage, he managed to make psychological sense to himself without attending to his psychology.

Schapiro concludes her commentary by examining a sentence about *akrasia* in one of my footnotes. There I wrote that if an agent mis-characterizes his own motives, “he may try to enact a disposition he doesn’t have — in which case he may fail to carry it off, thereby suffering *akrasia*, or weakness of will” [p. 15, fn. 7]. Schapiro says, “This makes it sound like the failure involved in *akrasia* is simply a failure to attribute to yourself the motives you actually have.” But, she complains, I don’t treat this mis-attribution as a failure, because I say that we should engage in wishful thinking, for example, by believing we are brave in order to steel our nerves

What I wrote, however, is that *akrasia* occurs when we extend the habit of effective wishful thinking to cases in which the wish won’t come true:

Without intentionally trying to exploit the power of wishful thinking, the agent can fail to distinguish between the cases in which his thoughts will be self-fulfilling and the cases in which they won’t. Accustomed to thinking ahead of the facts about himself and relying on them to catch up, he can fail to notice when they are no longer following” [p. 92].

In other words, the mis-attribution is indeed a failure if the agent fails to act accordingly and, by that means, to bring his character into accord with it.

Schapiro says, “It does seem natural to think that my response [to my own *akrasia*] might be to say, ‘I don’t know what I’ve been doing!’” Schapiro doesn’t recognize this view of the case as mine. “I do think,” she says, “that clear-eyedness in this sense might well be at stake in every action, but I do not see evidence that this is what Velleman has in mind.” It is precisely what I have in mind, and I am glad to find that Schapiro agrees.

#### *Response to Justin D’Arms*

Justin D’Arms correctly explains how my theory of self-improvisation extends beyond actions to reactions, and through them to values. What’s valuable, I say, is what it makes sense to value — to desire, in the case of the desirable; to admire, in the case of the admirable; to detest, in the case of the detestable; and so forth.

What it makes sense to value is prior in the order of determination to what is valuable. For example, something funny makes sense to laugh at, not because it’s funny, but because it’s like other things that make one laugh and unlike things that make one wince or gag; because it’s shocking to one’s expectations, amenable to one’s prejudices, titillating but not gross; and so on. I then distinguish between merely laughing at something and finding it funny. Finding something funny is being sensitive, in laughing at it, to whether doing so makes sense. It is, in other words, laughter guided by the thing’s being funny, though it doesn’t rise to the level of a judgment to that effect. D’Arms explains this view admirably and invites me to join him in considering some problem cases.

In one sort of case, you laugh at a lame joke because you are drunk, while judging that the joke isn’t funny. By my definition, finding the joke funny would require laughing at it partly because of the perception that laughing at it made sense. Since you see that this joke isn’t funny, your laughter must not amount to finding it funny, so defined. D’Arms points out that your laughter and your judgment are “rationally at odds”, and he argues that they wouldn’t be at

odds unless your laughter was to some extent evaluative, which would make it a “finding” of humor in the joke. So my definition of “finding funny” seems too restrictive.

In my view, however, the laughter and the judgment are at odds, not because the laughter is evaluative, but because, in light of your judgment, the laughter doesn’t make sense. That’s all it takes for them to be at odds. No “finding” needed.

In the next case, you laugh at a joke told at a party by a woman to whom you are attracted, but when you retell the joke the next day, you find that it isn’t funny. Did you find it funny the night before, in my sense of “find funny”? It all depends. You may have found it funny by mistake, because you misperceived it as having qualities that made it intelligible to laugh at. Perhaps you perceived it as clever, novel, erudite when in fact it was hackneyed and low-minded. You mistook pretty for witty. That can happen (not to me, dear). In the light of day, you see your mistake, and so you rightly find the joke unfunny. You find it to have different qualities on different occasions because one “finding” was based on erroneous perceptions.

D’Arms favors a version of the story in which you realized at the time that you were laughing at the joke partly because you were attracted to the teller. Laughing would then have made sense to you in light of circumstances, and its making sense would have guided you to have a good laugh, just as the intelligibility of crying guided you to have a good cry. Do I want to say, then, that you found the joke funny?

First consider a slightly different case, to which D’Arms also alludes. Sometimes the humor lies not just in the joke but also in the delivery. When the joke falls flat the next day, you say, “Well, it was funny when *she* told it.” In that case, what was funny the night before was not the joke itself but the whole performance, joke plus delivery. If you were guided by a perception of the performance as making sense to laugh at, then you indeed found it funny — the whole performance, not just the joke. What you subsequently find unfunny is a different performance.

Now return to the case in which you laugh in part because of the joke-teller’s attractions, which also help to make it seem intelligible to laugh at. Should I treat this case like the last, by saying that you found something funny, namely, the joke plus the teller’s attractions? Surely, her looks were not funny-making.

Faced with the last two cases, I might respond that the delivery of utterly unfunny content can also make you laugh — Mel Brooks reading the phonebook, for example — whereas good-looking women elicit a smile but never so much as a chuckle. But that distinction seems irrelevant to the question what you found funny in this particular case, where a woman’s looks were partly responsible not only for making you laugh but also for guiding your laughter. A better response is that you laughed *at* the delivery as well as the joke in the first case, whereas you laughed at the joke alone in the second, the teller’s looks being a circumstance rather than a target of your laughter. The distinction between the circumstances of a response and its target is psychological: there is a psychological fact as to what you are laughing at as opposed to other things that are causally implicated in your laughter. The same distinction has to be drawn in the case of many responses: your indigestion is partly responsible for your anger, but you’re angry *at* the other driver; you’re afraid *of* the dog, but also partly because of the darkness.

More pressing than the question whether you found the joke funny is whether it *was* funny by your lights, that is, deserving of laughter from someone with your sense of humor. As D’Arms points out, this case shows that what makes sense for you to laugh at need not be what is funny for you, or by your lights, since you ultimately concluded that the joke wasn’t funny but that, under the circumstances, laughing at it made sense. The case therefore appears to refute my

view that what's funny is what it makes sense to laugh at, or more generally, what's valuable is what it makes sense to value.

D'Arms offers me a solution to this problem. The solution is that we can find regularities in how our responses are regularly affected across the board by such things as priming, emotional contagion, heightened arousal, intoxication, and the like. We learn that having laughed at the last joke disposes us to laugh at the next, that having been angry at one meeting disposes us to be angry at the next, and in general that responses tend to carry over from one target to another. We learn that drink makes us quicker to laugh, but that also it makes us quicker to cry or to lose our tempers. We learn to recognize the damping effect of depression, the amping effect of anxiety, effects that apply to many responses alike. All of these psychological mechanisms, once recognized, can contribute to the intelligibility of a particular response on a particular occasion, but they can also be factored out to yield a judgment of what response makes sense other things being equal — “equal” meaning apart from effects not peculiar to that response. Whether a particular telling of a joke is funny is determined by all of the circumstances. But whether the joke itself is funny depends on whether it is in itself such as to make laughter intelligible, where “in itself” means apart from regular sources of interference. The joke is funny only if laughing at it makes sense other things being equal. In sum, the joke in itself wasn't funny because laughing at it made no sense apart from circumstances that alter many responses in predictable ways.

Having offered me this solution to the current problem, D'Arms worries that it undermines my remarks about the interdependence between different responses with respect to their intelligibility. Laughter is dampened by disgust, I say, and so laughing at a joke makes less sense if it is disgusting. Shouldn't I now say that the disgust can be factored out, so that the joke turns out to be funny other things being equal?

No, I shouldn't. If an anodyne joke were told over a disgusting meal, the dampening effect on laughter might be factored out as distortion, and the joke itself could still be such as made sense to laugh at, hence funny. But in this case the joke itself is disgusting, and so it is not funny in itself. We can say that it would be funny if it weren't so disgusting, but we cannot say that it is funny though not when told in those circumstances.

No author could ask for more attentive or more constructive commentators than Phil Clark, Tamar Schapiro, and Justin D'Arms. Their commentaries have given me a welcome opportunity to rethink, revise, and clarify the views presented in *How We Get Along*. I am grateful to them and to the editors of *Abstracta* for a thoroughly enjoyable and enlightening philosophical exchange.