In his 1970 paper “How is Weakness of the Will Possible?” Donald Davidson introduced his solution to the philosophical problem of akrasia by framing it as a problem about the self:[2]

The image we get of incontinence from Aristotle, Aquinas, and Hare is of a battle or struggle between two contestants. Each contestant is armed with his argument or principle. One side may be labeled ‘passion’ and the other ‘reason’; they fight; one side wins . . . . There is, however, a competing image (to be found in Plato, as well as in Butler and many others). It is adumbrated perhaps by Dante (who thinks he is following Aquinas and Aristotle) when he speaks of the incontinent man as one who ‘lets desire pull reason from her throne’ (Inferno, Canto v). Here there are three actors on the stage: reason, desire, and the one who lets desire get the upper hand. The third actor is perhaps named ‘The Will’ (or ‘Conscience’). It is up to The Will to decide who wins the battle. If The Will is strong, he gives the palm to reason; if he is weak, he may allow pleasure or passion the upper hand.

This second image is, I suggest, superior to the first, absurd as we may find both. On the first story, not only can we not account for incontinence; it is not clear how we can ever blame the agent for what he does: his action merely reflects the outcome of a struggle within him. What could he do about it? And more important, the first image does not allow us to make sense of a conflict in one person’s soul, for it leaves no room for the all-important process of weighing considerations. In the second image, the agent’s representative, The Will, can judge the strength of the arguments on both sides, can execute the decision, and take the rap. The only trouble is that we seem back where we started. For how can The Will judge one course of action better and yet choose the other?

Davidson did not refer to the self as such, but when he spoke of “the agent’s representative”, he was invoking the same idea—the idea that a person’s psyche hosts a conversation among several voices, one of which has the distinction of speaking for the person himself.

Davidson’s solution to the problem of akrasia did not live up to its picturesque introduction. It did not identify “the agent’s representative” but merely divided the conative attitudes into prima facie, all-things-considered, and “all-out” judgments. The last of these judgments was cast in the role of The Will, insofar as it constituted the agent’s immediate intention to act, but Davidson never explained why this attitude should be conceived as representing the agent rather than simply bringing up the rear in a parade of attitudes passing through the agent’s mind.

~

A more successful attempt at identifying “the agent’s representative” appeared in the following year,
when Harry Frankfurt published “Freedom of the Will and the Concept of a Person”. Like Davidson, Frankfurt discussed cases of conflicting motives, but he offered an explanation of how one of those motives gains the authority to represent the agent himself. A motive gains this authority, Frankfurt explained, from the agent’s desire that it predominate, which Frankfurt called a second-order volition. Frankfurt illustrated this phenomenon with the example of a drug-addict who has a second-order desire to resist his craving for the drug:

The unwilling addict identifies himself, . . . through the formation of a second-order volition, with one rather than with the other of his conflicting first-order desires. He makes one of them more truly his own and, in so doing, he withdraws himself from the other. It is in virtue of this identification and withdrawal, accomplished through the formation of a second-order volition, that the unwilling addict may meaningfully make the analytically puzzling statements that the force moving him to take the drug is a force other than his own, and that it is not of his own free will but rather against his will that this force moves him to take it.

Frankfurt is usually interpreted as saying that human agency stems from the occurrence of higher-order volitions. What has not been widely noted, I think, is that Frankfurt traced the roots of agency further, to a particular interest shared by all human agents. In “Identification and Wholeheartedness”, he wrote:

It is a salient characteristic of human beings, one which affects our lives in deep and innumerable ways, that we care about what we are. This is closely connected both as cause and as effect to our enormous preoccupation with what other people think of us. We are ceaselessly alert to the danger that there may be discrepancies between what we wish to be (or what we wish to seem to be) and how we actually appear to others and to ourselves.

We are particularly concerned with our own motives. It matters greatly to us whether the desires by which we are moved to act as we do motivate us because we want them to be effective in moving us or whether they move us regardless of ourselves or even despite ourselves.

Here Frankfurt spoke, not of particular higher-order desires to be motivated in one way or another on a particular occasion, but of a standing desire to be motivated as we want, and to be so motivated because we want to be. This is, in fact, a third-order desire, for there to be correspondence and causal influence between our second- and first-order desires.

The importance of this general interest in our own motivational integrity was reflected in Frankfurt’s portrait of the paradigm nonagent, a figure that he dubbed the “wanton”:

[It] never occurs to him to consider whether he wants the relations among his desires to result in his having the will he has. The wanton addict may be an animal, and thus incapable of being concerned about his will. In any event he is, in respect of his wanton lack of concern, no different from an animal.

. . . . It would be misleading to say that he is neutral as to the conflict between his desires, since this would suggest that he regards them as equally acceptable. Since he has no identity apart from his first-order desires, it is true neither that he prefers one to the other nor that he prefers not to take sides.

. . . . His lack of concern is not due to his inability to find a convincing basis for preference. It is due either to his lack of the capacity for reflection or to his mindless indifference to the enterprise of evaluating his own desires and motives.

The characteristic feature of the wanton is not that he merely lacks higher-order desires as to how he is motivated. Merely to lack such desires might just amount to indifference, which is not sufficient for wantonness. What characterizes the wanton is “his mindless indifference to the enterprise of evaluating his own desires and motives”, which is the lack of a third-order desire to engage in forming
second-order desires.[7]

Frankfurt envisioned this “enterprise” as requiring some motive of its own: [8]

Now what leads people to form desires of higher orders is similar to what leads them to go over their arithmetic. Someone checks his calculations because he thinks he may have done them wrong . . . . Similarly, a person may be led to reflect on his own desires either because they conflict with each other or because a more general lack of confidence moves him to consider whether to be satisfied with his motives as they are.

The wanton lacks this motive for reflective evaluation, and so he is like someone who loses interest in checking his arithmetic: [9]

One way in which a sequence of calculations might end is that the person conducting it simply quits, negligently permitting the result of his last calculation to serve as his answer. Perhaps he just loses interest in the problem, or perhaps he is diverted from further inquiry by some compelling distraction. In cases like these, his behavior resembles that of a wanton. . . .

What’s analogous to wantonness is not the lack of an opinion as to whether the last calculation is right or wrong but rather a lack of interest in the enterprise of forming such an opinion. The wanton thus lacks a third-order motive for forming second-order evaluations of his first-order attitudes.[10]

Frankfurt thus posited a single motive whose operation lies behind all human agency. Our agency arises from our concern over “whether the desires by which we are moved to act as we do motivate us because we want them to be effective in moving us or whether they move us regardless of ourselves and even despite ourselves.”[11]

But how can “the desires by which we are moved to act as we do motivate us because we want them to be effective in moving us”? How, that is, can our wanting a desire to be effective in moving us make any difference as to whether it actually is effective? And even if a first-order desire can be made effective by our wanting it to be so, how do we thereby become identified with it, so that it becomes more truly our own, attaining the status of what Davidson called “the agent’s representative”?

I am going to offer answers to both questions. Although I will trace these answers to various passages in Frankfurt’s papers, I cannot claim that they are his. They certainly do not belong to the standard interpretation of Frankfurt’s analysis of agency. The most I can claim is that they belong to an interpretation under which Frankfurt’s analysis is correct—correct as the analysis of something, that is, though not necessarily of agency.

When Frankfurt describes desires as moving us “because we want them to be effective in moving us”, he certainly seems to have a causal process in mind. Speaking of the so-called willing addict, who wants to act on his addiction, Frankfurt writes: [12]

I am inclined to understand his situation as involving the overdetermination of his first-order desire to take the drug. This desire is his effective desire because he is physiologically addicted. But it is his effective desire also because he wants it to be . . . . [1] It is therefore not only because of his addiction that his desire for the drug is effective . . . .

“Overdetermination” sounds like a causal process. Frankfurt appears to be saying that the addict’s first-order desire to take the drug is made effective by two independently sufficient causes: it is effective both “because he is psychologically addicted” and “because he wants it to be.” The latter of these causes is the addict’s second-order volition to be actuated by his first-order desire to take the drug. The question is how this cause operates: how can a second-order desire cause a first-order desire to become motivationally effective?[13] A related question is why Frankfurt describes this process as the overdetermination of the addict’s desire rather than the overdetermination of his behavior.
I think that an answer to the first question lies in the following passage introducing Frankfurt’s notion of a second-order volition:

There are . . . two kinds of situation in which it may be true that A wants to want to X. In the first place, it might be true of A that he wants to have a desire to X despite the fact that he has a univocal desire, altogether free of conflict and ambivalence, to refrain from X-ing. Someone might want to have a certain desire, in other words, but univocally want that desire to be unsatisfied. . . .

Someone who wants only in this truncated way to want to X stands at the margin of preciosity, and the fact that he wants to want to X is not pertinent to the identification of his will. There is, however, a second kind of situation that may be described by “A wants to want to X”; and when the statement is used to describe a situation of this second kind, then it does pertain to what A wants his will to be. . . . It is not merely that he wants the desire to X to be among the desires by which, to one degree or another, he is moved or inclined to act. He wants this desire to be effective—that is, to provide the motive in what he actually does. Now when the statement that A wants to want to X is used in this way, it does entail that A already has a desire to X. It could not be true both that A wants the desire to X to move him into action and that he does not want to X. It is only if he does want to X that he can coherently want the desire to X not merely to be one of his desires but, more decisively, to be his will.

The end of this passage is subject to at least two readings. The second way of wanting to want to X is a desire not just for some desire to X but for an effective desire to X, which is what Frankfurt would call a second-order volition. On the one hand, the passage could mean that this second-order volition must be a desire for the effectiveness of a particular first-order desire that one already has and, happily, wants to have. The second-order desire must refer to an already existing desire and must wish, of that desire, that it be effective. On the other hand, the reason why wanting an effective desire to X entails having a desire to X may be, not that it requires an already existing desire to X that can serve as its object, but rather that it just is a desire to X. Wanting an effective desire to X is, by definition, wanting that one actually do X as the result of a desire to do so. It therefore involves wanting oneself to X, which is a reflexive way of wanting to X.

I think this second reading is preferable, because it doesn’t require that one’s second-order desire be a desire for a first-order desire that one already has. According to this reading, both ways of wanting to want to X are desires for a first-order desire that one does not already have. What Frankfurt is saying, on this reading, is that in wanting an effective first-order desire to X, which one does not already have, one already does have a desire to X, after all — namely, this very second-order desire.

This reading helps to explain the following passage from a different paper: “Suppose a man wants to be motivated in what he does by the desire to concentrate on his work. It is necessarily true, if this supposition is correct, that he already wants to concentrate on his work. This desire is now among his desires.” Why “now”? The answer, according to this interpretation, is that the desire to concentrate on his work just consists in his second-order desire for such a desire to motivate what he does.

This interpretation also enables us to explain how a second-order volition can cause a previously existing first-order desire to become motivationally effective though it was not effective before. A second-order volition has a first-order motivational force matching that of the first-order desire whose effectiveness it endorses. A pre-existing desire to X can thus be reinforced by a second-order desire that it be effective, because the latter desire is also a desire (for oneself) to X. Hence, a second-order volition can cause a first-order desire to become motivationally effective by reinforcing it.

Thus, when Frankfurt says that the case of the willing addict involves “overdetermination of his first-order desire to take the drug”, he means that the addict has two desires to take the drug: his addictive urge to take the drug and a desire that he take the drug because of that urge, which is a desire for himself to take the drug and hence a reflexive desire to take it. Because the addict has two desires...
to take the drug, the fact that he wants to take it is overdetermined—which, I think, is what Frankfurt means by the "overdetermination of his first-order desire to take the drug."

This explanation necessitates a subtle clarification about the content of second-order volitions. A second-order volition that one be effectively moved by a first-order desire cannot have the content that one be effectively moved by the first-order desire alone. The content of a second-order volition must be that one be effectively moved by the first-order desire as reinforced by this very volition. Otherwise, the volition would tend to be self-frustrating.

For suppose that the willing addict formed a second-order volition to be effectively moved by his addiction alone. If he ended up taking the drug because his addictive urge was reinforced by his second-order volition, then he would not have taken it because of the addictive urge alone, and so he would not have satisfied the volition, after all. His volition would be necessarily self-frustrating in the sense that its motivational effectiveness would tend to cause it to go unfulfilled: in order to be fulfilled, it would have to be motivationally inert.

Surely, a volition of this form is inappropriate for creatures like us, to whom "[i]t matters greatly . . . whether the desires by which we are moved to act as we do motivate us because we want them to be effective in moving us or whether they move us regardless of ourselves and even despite ourselves."[19] If we want it to be the case that our desires are effective because we want them to be, then we had better not want them to be effective all on their own; our desire for them to be effective had better leave open the possibility of their being effective because of our hereby wanting them to be so.

Indeed, given our general desire that our motives for action be effective because we want them to be, our second-order volitions may not only leave open that possibility but positively favor it. Our desire that some first-order motive be effective may be, more specifically, the desire that it be effective partly because of being reinforced by this very desire. Second-order volitions would then incorporate the concern that is fundamental to agency, in Frankfurt's view. Wanting our first-order motive to be effective when reinforced by this very volition would be an instance of our concern that motives actuate us because we want them to, rather than regardless of or despite ourselves.

The resulting model of agency can be summarized as follows. Our concern that motives generally actuate us because we want them to leads us to reflect on our desires and to form a second-order volition that one of them actuate us because of our hereby wanting it to. This volition is a desire for us to do that which the first-order desire is a motive for doing, and so it reinforces that desire, which consequently actuates us because we want it to.

How can this process constitute our identifying with the first-order desire? How does it authorize the desire to represent us in the contest among our motives? I do not believe that an answer to these questions can be extracted from Frankfurt's own words. Frankfurt's conception of identification remains opaque throughout his writings on agency. I am going to propose an answer, partly on the basis of hints in the texts but largely on the strength of its own plausibility. I think of it as the answer that might have been in Frankfurt's mind as he wrote or may be in ours as we read, insofar as we find him persuasive.

Frankfurt states "The Problem of Action" as follows: [20]

[An] explication of the nature of action must deal with two distinct problems. One is to explain the notion of guided behavior. The other is to specify when the guidance of behavior is attributable to an agent and not simply, as when a person's pupils dilate because the light fades, to some local process going on within the agent's body. The first problem concerns the
conditions under which behavior is purposive, while the second concerns the conditions under which purposive behavior is intentional.

Frankfurt points out that the contrast between “instances in which purposive behavior is attributable to a creature as agent and instances in which this is not the case” can be applied to the behavior of lower animals, such as spiders.[21] The appearance of the contrast in the behavior of spiders constrains possible solutions to the problem of agency: “The conditions for attributing the guidance of bodily movement to a whole creature, rather than only to some local mechanism within a creature, evidently obtain outside of human life. Hence they cannot be satisfactorily understood by relying upon concepts which are inapplicable to spiders and their ilk.”[22]

An analysis of action in terms of reflective attitudes, which are presumably unavailable to spiders, would appear to violate this constraint. Frankfurt therefore suggests that there must be general conditions of agency whose satisfaction in humans requires these sophisticated attitudes but whose satisfaction in spiders does not: “While the general conditions of agency are unclear, it may well be that the satisfaction of these conditions by human beings depends upon the occurrence of events or states which do not occur in the histories of other creatures.”[23]

To my knowledge, Frankfurt never attempts to clarify what he here calls “the general conditions of agency”, which pick out the behaviors that are attributable to spiders rather than to sub-arachnoid mechanisms, just as they pick out the behaviors that are attributable to people rather than sub-personal mechanisms. But here is a passage that I find suggestive:[24]

Behavior is purposive when its course is subject to adjustments which compensate for the effects of forces which would otherwise interfere with the course of the behavior . . . . The behavior is in that case under the guidance of an independent causal mechanism, whose readiness to bring about compensatory adjustments tends to ensure that the behavior is accomplished. The activity of such a mechanism is normally not, of course, guided by us. Rather it is, when we are performing an action, our guidance of our behavior. Our sense of our own agency when we act is nothing more than the way it feels when we are somehow in touch with the operation of mechanisms of this kind, by which our movements are guided and their course guaranteed.

Here Frankfurt makes explicit what is elsewhere only implicit—and yet, I think, pervasively implicit—in his writings on agency: his approach to the topic is phenomenological, informed primarily by observations about “our sense of our own agency”, of what it is like to be an agent. And what it is like to be an agent, Frankfurt says, is to be somehow “in touch” with the mechanisms guiding our behavior.[25]

Suppose that the phenomenological notion of being “in touch” with what guides one’s behavior expresses “the general condition of agency”, as Frankfurt conceives it. If this general condition of agency requires for its satisfaction in humans the occurrence of events or states that do not occur in other creatures, as Frankfurt suggests, the reason should be that those events or states are necessary in order for humans, but not for other creatures, to be in touch with the mechanisms guiding their behavior. Maybe spiders achieve agency with less sophisticated mental equipment, in other words, because they need less equipment to be in touch with their own mechanisms of behavioral guidance. So, at least, we might hypothesize on the basis of these few passages.

What would account for such a difference between humans and spiders? Two of the papers under consideration—“Freedom of the Will” and “Identification and Wholeheartedness”—begin with the distinction between human persons and lesser creatures, including the lower animals. In the first paper, the difference is attributed to the presence in persons of hierarchically ordered attitudes. But in the second, it is attributed to the presence of reflective consciousness, and Frankfurt remarks in a footnote, “The notion of reflexivity seems to me much more fundamental and indispensable, in dealing with the phenomena at hand, than that of a hierarchy.”[26] Perhaps, then, reflective consciousness is
what makes additional events and states necessary to put persons in touch with the mechanisms guiding their behavior, thus enabling them to satisfy the general condition of agency.

I have now inched my way along a branch of Frankfurtian texts that can no doubt be lopped off behind me at any one of several junctures. Rather than trust my weight to the texts any longer, I am going to leap off into thin air. What follows cannot be supported by textual evidence at all; it is a speculative reconstruction.

According to the standard interpretation of Frankfurt, the difference between the motives with which one is “identified” and the motives to which one stands as a “helpless bystander” is that one wants the former to be effective in determining one’s behavior but wants the latter not to be effective.[27] What produces the opposite of identification with one’s motives (a condition for which Frankfurt has no negative term correlative to “identification”) is the presence of a negative higher-order attitude.

I want to suggest an alternative to this interpretation.[28] In the phenomenological sense of “identification” that I attribute to Frankfurt, the opposite of identification—call it dis-identification for now—is produced by reflective consciousness itself.

Being the subject of a desire usually entails being the subject of various thoughts symptomatic or expressive of the desire. Being thirsty, for example, entails thinking thirsty thoughts: looking around for quantities of liquid; wondering if they are potable; considering ways of reaching them, avidly imagining their taste; and so on. All of these thoughts are framed from the perspective of a potential drinker, but none explicitly represents the occupant of that perspective. They are framed from the point of view of a potential drinker who remains out of the picture, at the unrepresented origin of that point of view. Of course, the thoughts symptomatic of thirst may include the first-personal thought “I’m thirsty,” but that thought is in the first instance an atomic expression of thirst, like smacking one’s lips or crying “Water!” rather than a compositionally analyzable attribution of thirst to oneself.

The difference between that expression of thirst and the attribution of thirst to oneself defines a continuum of possible thoughts that include awareness of one’s thirst in various degrees of explicitness. Sometimes one looks for a drink without yet knowing that one is thirsty; sometimes one looks for a drink while knowing about one’s thirst but not attending to it at all; sometimes one attends equally to the possible drink and the dryness of one’s throat or the urgency of one’s craving; sometimes one focuses on the thirst to the exclusion of the prospects for slaking it.

Across this continuum, one becomes progressively less engrossed in the activities motivated by thirst. At the former end are the cases in which one “loses oneself” in gazing at the cool drink being served at the next table, or in peeling an orange, or in assaulting the shell of a coconut. In the middle of the continuum are the cases in which one undertakes such activities with cool self-possession. At the latter end are cases in which one is distracted by one’s thirst from the very activities that it would motivate. Cases of the first kind can end with the thought “Oh, I must be thirsty”: noticing that one’s attention has become engrossed in the pitcher of water carried by a waiter, one belatedly becomes aware of one’s thirst. Cases of the last kind can end with the thought “Stop thinking about how thirsty you are and get a drink!”

This last thought is naturally couched in the reflexive second person, because it occurs when one has put a distance between oneself and one’s thirst—that is, between one’s reflecting self and one’s thirsting self. Attentively reflecting on one’s thirst entails standing back from it, for several reasons. First, the content of one’s reflective thoughts is not especially expressive of the motive on which one is reflecting: “I am thirsty” is not an especially thirsty thought, not necessarily the thought of someone thinking thirstily. Second, attentive reflection is itself an activity—a mental activity—and as such it requires a motive, which of course is not thirst. Reflecting on one’s thirst is therefore a distraction from acting on one’s thirst, and in that respect it is even a distraction from being thirsty. Most importantly,
though, consciousness just seems to open a gulf between subject and object, even when its object is the subject himself. Consciousness seems to have the structure of vision, requiring its object to stand across from the viewer—to occupy the position of Gegenstand.

The distancing effect of reflective consciousness can be represented in terms of perspectives or points of view. The perspective one occupies as the subject of thirst is not the same as the perspective one occupies as the subject of reflective thinking about being thirsty. In the former perspective, one’s thirst is out of view, at the unrepresented point of origin, from which it issues in thirsty thoughts and actions directed at possible sources of drink, which dominate the field of view. In the latter perspective, one’s thirst is placed within the field of view, whose unrepresented origin may now contain reflective curiosity or concern but is unlikely to contain thirst, since thirst is not a motive for reflecting. The reason why becoming reflectively aware of one’s thirst tends to make one less engrossed in thirsty activities is that such awareness draws one away from the perspective in which thirst occupies the governing point of origin rather than the passive field of view.

These remarks return me to my current suggestion, that we are distanced—or, as I put it, dis-identified—from our motives by reflective consciousness itself. The more conscious we become of a motive, the more it becomes the object of our thought; and the more it becomes the object of our thought, the less we think from the perspective of its subject; and the less we think from the perspective of the motive’s subject, the less engrossed we are in the activities that it motivates.

This suggestion bears on what I have interpreted Frankfurt as believing about the general condition of agency. My interpretation is that what makes for agency in general, according to Frankfurt, is the phenomenological condition of being “in touch” with the mechanisms guiding one’s behavior. And the feature that distinguishes a person from a lower animal now turns out to take the person out of touch with those mechanisms, since what distinguishes him as a person is reflective consciousness, which opens a gulf between him as the subject of reflection and his motives as the object. A lower animal is already in touch with its motives in a way that a person is not, because the animal lacks the capacity for stepping back to a reflective distance. A person is continually falling out of touch with his motives, by becoming reflectively conscious of them.

That is why a person’s ability to satisfy the general condition of agency “depends upon the occurrence of events or states which do not occur in the histories of other creatures.” In order to satisfy the condition of agency, a person must overcome a disadvantage that other creatures do not face—namely, the distancing effect of reflective consciousness. It is in order to bridge this reflective gap that human agency requires events and states not present in other creatures. Their function is to return the person and his motives to a self-rapport of a kind that lesser creatures are not capable of losing. How can second-order volitions perform this function? My earlier account of their motivational efficacy explains how.

As I explained earlier, a second-order volition to be actuated by a desire to drink is a desire for oneself to act on that lower-order desire, possibly as reinforced by this very higher one. It is therefore a desire for oneself to drink, which is a reflexive way of wanting to drink. That is why the volition can reinforce the lower-order desire to drink, thereby enabling it to take effect in one’s behavior.

Because the second-order volition is a reflexive desire to drink—that is, a desire for oneself to drink—it enables one to combine perspectives that were previously disjoint. In wanting oneself to drink, one occupies a perspective in which one is simultaneously the subject of a desire to drink and the subject of reflective consciousness of oneself as wanting to drink. Reflection on one’s first-order desire to drink may have drawn one out of one’s perspective as subject of that desire, but reflexively wanting oneself to drink introduces a desire to drink into one’s new perspective as subject of reflection. The higher-order desire enables one’s reflecting self and one’s thirsting self to share a perspective, because one now thirsts reflectively. Therein lies the identification, which closes the reflective gap, putting one in touch with the mechanisms guiding one’s behavior and hence fulfilling the general condition of agency.
As I have said, I do not attribute this conception of identification to Frankfurt; I attribute it primarily to myself as a reader of Frankfurt: it is what I have in mind when I find myself persuaded by Frankfurt's talk of identification. Maybe it is what other readers have in mind, too. For those whose reading of Frankfurt is informed by some such conception of identification, I want to indicate three significant consequences.

First, this conception of identification insulates Frankfurt's view from a well-known objection that was raised by Gary Watson and addressed by Frankfurt on several occasions. The objection is that second-order volitions cannot succeed in identifying the agent with a first-order desire unless the agent is already identified with the second-order volition. If he can make the "analytically puzzling statement" that a first-order desire is "a force other than his own", then he should be able to make a similar statement about a second-order volition. And if the second-order volition is a force other than his own, then its endorsement of a first-order desire can hardly make the latter "truly his own". How, then, does the agent become identified with his second-order volition? If the answer is that the second-order volition becomes his own by being endorsed at a higher order, a vicious regress of volitions will ensue.

This regress disappears under the conception of identification that I have introduced. For one thing, the agent is identified with his desires by default, until he steps back to reflect on them. And for each reflective step back, the agent can bridge the resulting distance by forming a reflective desire, with which he will be identified by default. No third-order desire is required unless and until the agent takes a second step back, to reflect on his second-order volition. When the agent stops stepping back, the regress stops with him.

Secondly, the agent's second-order volition can endorse itself, at least implicitly, since it may be a desire for his first-order motive to be efficacious, not by itself, but with reinforcement from this very desire. Reflective desires represent a general solution to the problem of reflective detachment—a way of being reflectively conscious of a desire without stepping back from it. They provide this solution because they provide a perspective in which the agent is simultaneously the subject of the desire and the subject of reflection upon it. Of course, an agent can step back from a reflective desire, to a perspective of higher-order reflection. But then he can close the resulting gap, with another reflective desire, for which there is no problem of reflective detachment, unless he chooses to create one by reflecting further. So my conception of identification has some good news for readers of Frankfurt: the vicious regress is gone. Now for the bad news.

The bad news is that Frankfurt's theory, as I have now interpreted it, does not solve the problem stated by Davidson—the problem of picking out "the agent's representative" among the competing motivational forces within him. The agent's representative is supposed to speak with the authority of the agent, articulating what will count as his best judgment and executing what will count as his decision. Representing the agent is thus a normative matter of being authorized, as his proxy, to make commitments for which he will be responsible.

But I have now interpreted higher-order volitions as identifying the agent with his motives, not in the normative sense of authorizing them to act as his proxies, but in the phenomenological sense of putting him "in touch" with them, by bridging the reflective gap. Under my interpretation, Frankfurt's theory becomes a theory of how to stay engaged or even engrossed in one's activities, despite the distancing effects of reflective consciousness. As such, it may no longer pick out a proper part of the psyche that (in Davidson's words) "can execute the decision and take the rap." Indeed, Frankfurt's theory may no longer be a theory of the self, under my interpretation. Any motive, even the most fleeting whim, can be the object of reflective identification; and what an agent identifies with at one moment, he can step back from at the next. Surely, being the object of the agent's identification is too ephemeral a status to carry the title of self.
Of course, Frankfurt himself has moved beyond the theory of the self expounded in the papers that I am interpreting here. His new theory, while still invoking the notion of identification, focuses mainly on the fixity of the self, its role as a stubborn bulwark against ambivalence and vacillation. Perhaps, then, Frankfurt himself believes that what his earlier papers offered was not, after all, a theory of the self.

Finally, if Frankfurt’s theory is intended to solve the problem of reflective awareness, as I suggest, then it may turn out to be a half measure, stopping short of a complete solution. Although we can bridge the reflective gap and get “in touch” with our motives by means of higher-order volitions, we can eliminate the gap entirely by becoming so engrossed in an activity that we stop reflecting and lose ourselves.

There is at least one philosophical tradition that recommends transcending reflective awareness in this manner. It is the Daoist tradition, especially as represented in the Zhuangzi. In my interpretation of Frankfurt, his theory of agency becomes a prolegomenon to that work.

The spiritual ideal expressed in the Zhuangzi is one of effortless action, as described by the phrase wu wei. The word wei means “action”, and wu wei is its negation—literally, “nonaction”. But “nonaction” does not mean doing nothing at all; it means acting without deliberate intention or effort—spontaneous activity: “It describes a state of personal harmony in which actions flow freely and instantly from one’s spontaneous inclinations—without the need for extended deliberation or inner struggle—and yet nonetheless accord perfectly with the dictates of the situation at hand, [and] display an almost supernatural efficacy.”

A prime example of spontaneous activity is the exercise of a “knack”, which is exemplified throughout the Zhuangzi by stories such as this:

Butcher Ding was cutting up an ox for Lord Wenhui. Wherever his hand touched, wherever his shoulder leaned, wherever his foot stamped, wherever his knee pushed—with a zip! with a whoosh!—he handled his chopper with aplomb, and never skipped a beat. He moved in time to the Dance of the Mulberry Forest, and harmonized with the Head of the Line Symphony. Lord Wenhui said, “Ah, excellent, that technique can reach such heights!”

The butcher sheathed his chopper and responded, “What your servant values is the Way, which goes beyond technique. When I first began cutting up oxen, I did not see anything but oxen. Three years later, I couldn’t see the whole ox. And now, I encounter them with spirit and don’t look with my eyes. Sensible knowledge stops and spiritual desires proceed. I rely on the heavenly patterns, strike in the big gaps, am guided by the large fissures, and follow what is inherently so. I never touch a ligament or tendon, much less do any heavy wrenching!”

Although we would regard carving oxen as the purposeful application of a skill, Butcher Ding has gone “beyond technique” and his movements are due to “spiritual desires” that “proceed” as if on their own. This process requires a shedding of ordinary purposes and precepts, as another artisan explains:

Woodworker Qing carved a piece of wood and made a bell stand, and when it was finished, every one who saw it marveled, for it seemed to be the work of gods or spirits. When the marquis of Lu saw it, he asked, “What art is it you have?”

Qing replied, “I am only a craftsman—how would I have any art? There is one thing, however. When I am going to make a bell stand, I never let it wear out my energy. I always fast in order to still my mind. When I have fasted for three days, I no longer have any thought of congratulations or rewards, of titles or stipends. When I have fasted for five days, I no longer have any thought of praise or blame, of skill or clumsiness. And when I have fasted for seven days, I am so still that I forget I have four limbs and a form and body. By that time, the ruler and his court no longer exist for me. My skill is concentrated and all outside distractions fade away. After that, I go into the mountain forest and examine the heavenly nature of the trees. If I find...
one of superlative form, and I can see a bell stand there, I put my hand to the job of carving; if not, I let it go. This way I am simply matching up ‘Heaven’ with ‘Heaven.’ That’s probably the reason that people wonder if the results were not made by spirits.”

The performance of artisans like Butcher Ding and Woodworker Qing is guided by an inexpressible knack. Wheelwright Pien says: “You can’t put it into words, and yet there’s a knack to it somehow. I can’t teach it to my son, and he can’t learn it from me.”[39]

The way to exercise such a knack is not to keep one’s eye on an ultimate goal, or to follow the precepts of a method, or even to focus on one’s actions themselves. On the contrary, Woodworker Qing must forget external goals (“congratulations and rewards, titles or stipends”), forget evaluative judgment (“blame or praise . . . skill or clumsiness”), and indeed forget himself: “I forget I have four limbs and a form and body.”

Such forgetfulness is necessary because spontaneous action is inhibited by distinctions between good and bad, right and wrong—or shi and fei, often translated as “That’s it” and “That’s not”.[40]

Yi’erzi visited Xu You.

“What riches did you get from Yao?” said Xu You.

“Yao told me: ‘Be sure to devote yourself to Goodwill and Duty and say plainly ‘That’s it, that’s not.’”

“Then what do you think you’re doing here? When that Yao has already branded your hide with Goodwill and Duty, and snipped off your nose with his ‘That’s it, that’s not,’ how are you going to roam that free and easy take-any-turn-you-please path?”

To “roam that free and easy take-any-turn-you-please path” is not a matter of taking a random walk through life; it is rather a way of attaining effective activity (“nonaction”). It can even be a solution to the problem of “how to rule the world”: “Let your mind wander in simplicity, blend your spirit with the vastness, follow along with things the way they are, and make no room for personal views—then the world will be governed.”[41]

A “wandering mind” is conducive to effective activity because of its responsiveness to the objective demands of one’s materials and circumstances—its capacity to “follow what is inherently so.” Thus, Butcher Ding “relies on the Heavenly patterns” and Woodworker Qing “examines the heavenly nature of the trees.” This responsiveness to “things the way they are” is what requires the mind to be emptied of external goals and evaluative judgments.

Ultimately, emptying the mind of goals and judgments leads to forgetfulness of the self. Here is another description of the same phenomenon: [42]

“I make progress”, said Yan Hui.

“Where?” said Confucius.

“I have forgotten about rites and music.”

“Satisfactory. But you still have far to go.”

Another day he saw Confucius again.

“I make progress.”

“Where?”

“I have forgotten about Goodwill and Duty.”
“Satisfactory. But you still have far to go.”

Another day he saw Confucius again.

“I make progress.”

“Where?”

“I just sit and forget.”

Confucius was taken aback.

“What do you mean, just sit and forget?”

“I let organs and members drop away, dismiss eyesight and hearing, part from the body and expel knowledge, and go along with the universal thoroughfare. This is what I mean by 'just sit and forget.'”

“If you go along with it, you have no preferences; if you let yourself transform, you have no norms. Has it really turned out that you are the better of us? Oblige me by accepting me as your disciple.”

The Zhuangzi’s conception of spontaneous activity has been compared to the “flow” experience described by the psychologist Mihaly Csikszentmihalyi.[43] Csikszentmihalyi conducted research in which subjects were prompted to record their activities, and their feelings about them, at regular intervals during the day. He then identified a category of “optimal experiences” that occur in the course of highly challenging activities in which the subject exercises appropriate skills.

According to Csikszentmihalyi, “flow” begins as follows: [44]

When all of a person’s relevant skills are needed to cope with the challenges of a situation, that person’s attention is completely absorbed by the activity. There is no excess psychic energy left over to process any information but what the activity offers. All the attention is concentrated on the relevant stimuli.

As in the “knack” stories of the Zhuangzi, evaluative judgment is suspended: [45]

In normal life, we keep interrupting what we do with doubts and questions. “Why am I doing this? Should I perhaps be doing something else?” Repeatedly we question the necessity of our actions, and evaluate critically the reasons for carrying them out. But in flow there is no need to reflect, because the action carries us forward as if by magic.

Also as in the “knack” stories, awareness of the self disappears: “[O]ne of the most universal and distinctive features of optimal experience [is that] people become so involved in what they are doing that the activity becomes spontaneous, almost automatic; they stop being aware of themselves as separate from the actions they are performing.”[46] Csikszentmihalyi goes on to explain that this loss of self-consciousness “does not involve a loss of self, and certainly not a loss of consciousness, but rather, only a loss of consciousness of the self.”[47]

As we have seen, Frankfurt regards reflective awareness as the distinctive characteristic of humanity. A spiritual ideal of transcending reflective awareness would thus be, in Frankfurt’s terms, an ideal of transcending what makes us human. But transcending what makes us human is just what the Zhuangzi and Csikszentmihalyi recommend.

According to the Zhuangzi, the Daoist sage “has the physical form of a human being but lacks the human essence.”[48] He explains: [49]

Judging “That’s it, that’s not” is what I mean by “the essentials of man.” What I mean by being
without the essentials is that the man does not inwardly wound his person by likes and dislikes, that he constantly goes by the spontaneous and does not add anything to the process of life.

The “likes” and “dislikes” mentioned here appear to be based on judgments of right and wrong, or “That’s it” and “That’s not”, and so they appear to be reflective attitudes. They no doubt include motives of the sort that result from thinking about “Goodwill and Duty”, or “skill or clumsiness”—self-critical instances of “That’s it” and “That’s not”. Hence, lacking a human essence, in the Zhuangzi’s sense, must entail lacking that “concern with our own motives” that makes us “care about what we are.”[50] It therefore entails lacking what Frankfurt identifies as the source of human agency, according to my interpretation.

Creatures who lack human reflectiveness are at an advantage in attaining effortless action: [51]

The kui [a mythical one-legged beast] said to the millipede, “I go hippety-flopping on one foot, and there’s nothing like it! How do you manage those ten thousand feet of yours?”

The millipede said, “It’s not like that . . . . I just put my heavenly mechanism into motion. I don’t know how it works!”

The millipede said to the snake, “I use this mob of legs to walk but still don’t match up to you with none at all. How do you do it?”

The snake said, “The heavenly mechanism does it. What could be easier? What use would I have for legs?”

The millipede and snake are like Frankfurt’s spider, who attains agency without all of the complications that are necessary for human beings. In all of these instances, agency is more readily attained because there is no reflectiveness that needs to be overcome.

Yet we humans cannot avoid exercising our reflective essence on the way to transcending it. Butcher Ding has gone “beyond technique”: that is, he cultivated a skill and then left it behind. So it must be with all of the artisans. They must have undertaken many years of self-critical practice in order to attain their capacity for self-forgetful spontaneity. Hence, the spiritual ideal of the Zhuangzi is not an alternative to that of reflective agency: it is the next step.

~

When a human being “finds flow” in the exercise of a skill, does he instantiate agency, as Frankfurt conceives it, or does he instantiate wantonness instead? Or is this case rather a challenge to the categories of agent and wanton altogether?

When we are in flow, we are not actuated by the self-concern that is central to agency in Frankfurt’s conception. We are not “concerned with our own motives”, and certainly not “ceaselessly alert to the danger that there may be discrepancies between what we wish to be (or what we wish to seem to be) and how we actually appear to others and to ourselves.”[52] Such self-concern is precisely what must be left behind if flow is to be attained. When a person is in flow, “it never occurs to him to consider whether he wants the relations among his desires to result in his having the will he has”—which is how Frankfurt describes the wanton.[53] The spontaneous activity celebrated by the Zhuangzi and Csikszentmihalyi thus seems to resemble a manifestation of wantonness.

But surely the actors described by the Zhuangzi and Csikszentmihalyi are “in touch” with the mechanisms guiding their behavior; they are anything but “helpless bystanders” to their activities.[54] And having set aside their human essence, they would seem entitled to be judged by the conditions of agency as they apply to animals, in whom reflection and higher-order attitudes aren’t required.[55] So can these actors be wantons and agents simultaneously?

The way to answer this question, I think, is to consider the sense in which these actors are “beyond
technique”. They have acquired their skills through training that involved self-scrutiny, self-criticism, and self-correction. The reflective capacity required for this training—the capacity to monitor their own performance, to consider how it falls short of an ideal, and to correct it accordingly—is no longer exercised after they have perfected their skill; but it is merely dormant, ready to be reactivated by the first misstep. Though Woodworker Qing loses himself in his work, forgetting even that he has four limbs, a single false stroke will recall him to self-awareness and re-engage his capacity for self-criticism and self-correction. He is beyond technique not because he has lost it but because his spontaneous activity now accords so well with the technique that it can lie dormant.

Actors in flow have thus achieved a higher wantonness. They act wantonly in the sense that they have dispensed with self-regulation. But they have dispensed with self-regulation only because it has been so effective as to render itself unnecessary. And their capacity for self-regulation remains in reserve in case it is needed. Hence, their wantonness is also a consummate example of agency.

NOTES

1. This chapter was written while I was a visiting professor in the Philosophy Department at the University of North Carolina, Chapel Hill, with the support of a grant made to Susan Wolf by the Mellon Foundation: my thanks go to the Foundation, the Department, and especially to Susan. I am indebted to Eric Hutton for extensive assistance with the Chinese materials discussed at the end of the chapter, and to Kim Atkins, Sarah Buss, P.J. Ivanhoe, and Catriona Mackenzie for comments. This version of the paper differs from the original in a few substantive respects.


7. Frankfurt, “The Problem of Action”, in The Importance of What We Care About, 74; hereafter cited as 1988c. This paper was originally published in American Philosophical Quarterly 15 (1978) 157–62. Although I will draw material from this paper, I think that its main thesis must be disregarded. Its thesis is that causal theories of action are wrong because they distinguish between actions and mere bodily movements solely in terms of their antecedent causes, thus treating them as “intrinsically” indistinguishable. The paper then argues that the distinction should be drawn in terms of how behavior is governed as it progresses. But the standard causal theories do not specify that the causes distinctive of action are ballistic causes, which launch action without going on to guide its progress; the standard causal theories are compatible with the assumption that the causes distinctive of action are cybernetic causes, in the sense that they exert ongoing guidance. Frankfurt’s arguments for this assumption, though valid, are wrongly targeted at causal theories of action. And as I interpret Frankfurt, his own theory of action falls squarely in the causalist tradition. (Frankfurt also objects to causal theories on the grounds that an agent may guide his behavior without actively intervening in its progress, if it progresses appropriately on its own. But, of course, omissions can be causes too.)

9. Ibid., 168.

10. Although Frankfurt never discussed the operation of this motive, it must operate sub-agentially, since evaluating one’s motives is not in the normal case a full-blooded action. Deliberate acts of self-evaluation tend to interrupt one’s other endeavors and can hardly be required to produce the reflective endorsements constituting those ordinary endeavors as actions. Hence the second-order volitions that transform mere motivated behavior into full-blooded action cannot, in the ordinary case, result from full-blooded acts of self-evaluation. One’s characteristically human concern with one’s motives must typically prompt self-evaluation that is not an action in its own right.


13. Frankfurt sometimes speaks as if the agent himself causes a first-order desire to become motivationally effective in accordance with his second-order volition:

It is in securing the conformity of his will to his second-order volitions, then, that a person exercises freedom of the will. And it is in the discrepancy between his will and his second-order volitions, or his awareness that their coincidence is not his own doing but only a happy chance, that a person who does not have this freedom feels its lack. (Frankfurt 1988a, 20–21)

To my knowledge, however, Frankfurt does not offer an account of how the person brings about this coincidence. And in other passages, he suggests that it is usually brought about without agential intervention:

[T]he conformity of a person’s will to his higher-order volitions may be far more thoughtless and spontaneous than this. Some people are naturally moved by kindness when they want to be kind, and by nastiness when they want to be nasty, without any explicit forethought and without any need for energetic self-control. (Frankfurt 1988a, 22)

See also the following passage:

Suppose that a person has done what he wanted to do, that he did it because he wanted to do it, and that the will by which he was moved when he did it was his will because it was the will he wanted. Then he did it freely and of his own free will . . . [S]ince the will that moved him when he acted was his will because he wanted it to be, he cannot claim that his will was forced upon him or that he was a passive bystander to its constitution. (Frankfurt 1988a, 24)


15. Admittedly, my interpretation is not favored by the footnote to this passage:

It is not so clear that the entailment relation described here holds in certain kinds of cases, which I think may fairly be regarded as nonstandard, where the essential difference between the standard and the nonstandard cases lies in the kind of description by which the first-order desire in question is specified. Thus, suppose that A admires B so fulsomely that, even though he does not know what B wants to do, he wants to be effectively moved by whatever desire effectively moves B; without knowing what B’s will is, in other words, A wants his own will to be the same. It certainly does not follow that A already has, among his desires, a desire like the one that constitutes B’s will. [Frankfurt 1988a, 15n.]

In this case, the desire at issue — wanting to want to X — is de re with respect to its object: A wants to have a desire that is, in fact, a desire to X, though he doesn’t want it under that description. The “description by which the desire in question is specified” is ‘whatever B wants’. Having this desire to want to X fails to entail having a desire to X for two reasons, corresponding to the two readings of Frankfurt’s main text. One reason is that although this second-order desire has a first-order desire to X
as its object, that desire is B's rather than A's. The other reason is that because the second-order desire doesn't specify the object of B's desire — namely, to X — it cannot be a desire for that object. In Frankfurt's parlance, however, the desire that "constitutes B's will" is his first-order desire. So the most natural reading is that the desire entailed in the "standard" case is a first-order desire, as presupposed in the first reading of the main passage rather than the reading that I prefer.

16. For a similar statement, see Frankfurt 1988a, 16: "Suppose a man wants to be motivated in what he does by the desire to concentrate on his work. It is necessarily true, if this supposition is correct, that he already wants to concentrate on his work. This desire is now among his desires." Why "now"? The answer, I will suggest, is that the desire to concentrate on his work consists in his second-order desire for such a desire to motivate what he does.

17. Gary Watson makes a related (though somewhat different) point: "the same considerations that constitute one's on-balance reasons for doing some action, a, are reasons for wanting the 'desire' to do a to be effective in action, and for wanting contrary desires to be ineffective" (Gary Watson, "Free Agency", Journal of Philosophy 75 [1975] 205–20, 219). I discuss Watson's paper in Chap. 5, "What Happens When Someone Acts".


21. Ibid., 78.

22. Ibid.

23. Ibid., 78–79.

24. Ibid., 74–75.

25. See also Frankfurt 1988c, 71, where Frankfurt says that "the most salient differentiating characteristic of action" is that "during the time a person is performing an action he is necessarily in touch with the movements of his body in a certain way, whereas he is necessarily not in touch with them in that way when movements of his body are occurring without his making them."


27. Frankfurt 1988a, 21.

28. What I have called the standard interpretation is probably correct in application to the view expressed in Harry Frankfurt, "The Faintest Passion", in Necessity, Volition, and Love (Cambridge: Cambridge University Press, 1999); hereafter cited as Frankfurt 1999a. (This paper was originally presented as Frankfurt's APA Presidential address, and published in Proceedings and Addresses of the American Philosophical Association 66 "1992" 5–16.) But that paper represents a departure from the earlier papers under discussion here. One measure of the departure is that the state of identification described in "The Faintest Passion" closely resembles the state of wantonness as defined in the earlier papers.

29. Frankfurt 1988c, 78–79.


32. Davidson, "How is Weakness of the Will Possible?", 35–36.

33. See "On the Necessity of Ideals" and "Autonomy, Necessity, and Love" (Cambridge: Cambridge
University Press, 1999); in Necessity, Volition, and Love. The latter paper was originally published in Hans Friedrich Fulda and Rolf-Peter Horstmann (eds.), Vernunftbegriffe der Moderne (Stuttgart: Klett-Cotta, 1994).


35. One possible difference between the Zhuangzi and Frankfurt is that the Daoist spiritual ideal can be interpreted as having an ethical dimension. For Frankfurt, overcoming reflective alienation does not entail conformity to any ethical ideal.


38. See also Chuang-Tzu: The Inner Chapters, 135; Slingerland, Effortless Action, 186.


40. Chuang-Tzu: The Inner Chapters, 91

41. Ibid., 135. Here is Graham’s translation: “Let your heart roam in the flavourless, blend your energies with the featureless, in the spontaneity of your accord with other things leave no room for selfishness, and the Empire will be in order” (95). As I have mentioned in note 35, spontaneous action may be more than effective, since the Way (dao) may have an ethical dimension as well.

42. Chuang-Tzu: The Inner Chapters, 92. See also The Complete Works of Chuang Tzu, 90–91.


44. Csikszentmihalyi 1990, 53.

45. Ibid., 54.

46. Ibid., 53.

47. Ibid., 64. David Satran has directed me to a wonderful description of flow in Anna Karenina. In Part III, chapters 4–5, Levin takes up a scythe to mow grass alongside the peasants:

They came to the end of another swath. They went on mowing long and short rows, good and poor grass. Levin had lost all count of time and had really no idea whether it was late or early.
His work was undergoing a change which gave him intense pleasure. While working he sometimes forgot for some minutes what he was about, and felt quite at ease; then his mowing was nearly as even as that of Titus. But as soon as he began thinking about it and trying to work better, he at once felt how hard the task was and mowed badly.

Now, in the hottest part of the day, the work did not seem so hard to him. The perspiration in which he was bathed was cooling, and the sun which burnt his back, his head and his arm—bare to the elbow—added to his strength and perseverance in his task, and those unconscious intervals when it became possible not to think of what he was doing recurred more and more often. The scythe seemed to mow of itself. Those were happy moments.

The longer Levin went on mowing, the oftener he experienced those moments of oblivion when his arms no longer seemed to swing the scythe, but the scythe itself his whole body, so conscious and full of life; and as if by magic, regularly and definitely without a thought being given to it, the work accomplished itself of its own accord. These were blessed moments.

It was difficult only at those times when it was necessary to interrupt the movement which had become unconscious, and to think in order to mow around a molehill or a space where the hard sorrel stalks had not been weeded out.

Levin did not notice how time passed. Had he been asked how long he had been mowing, he would have answered “half an hour”, although it was nearly noon.

Leo Tolstoy, Anna Karenina, the Maude Translation, revised by George Gibian (New York: W.W. Norton & Company, 1995), 227–30.


49. Chuang-Tzu: The Inner Chapters, 82.

50. Frankfurt 1988b, 163.


52. Frankfurt 1988b, 163.


55. These considerations raise the question whether animals can be wantons. The passage quoted earlier from Frankfurt 1988a, 18, appears to suggest that they can. But I suspect that Frankfurt generally reserves the term ‘wanton’ for creatures who are capable but fall short of human agency. The wanton is a human being whose activity would qualify as agency in animals but doesn’t qualify as agency in humans.

Originally published at http://dx.doi.org/10.3998/maize.13240734.0001.001