

FREEDOM OF THE WILL AND THE CONCEPT OF A PERSON

Harry G. Frankfurt, *The Journal of Philosophy*, Vol. 68, No. 1, 1971.

(Edited by RJ)

... It is my view that one essential difference between persons and other creatures is to be found in the structure of a person's will. Human beings are not alone in having desires and motives, or in making choices. They share these things with the members of certain other species, some of whom even appear to engage in deliberation and to make decisions based upon prior thought. It seems to be peculiarly characteristic of humans, however, that they are able to form what I shall call "second-order desires" or "desires of the second order."

Besides wanting and choosing and being moved *to do* this or that, men may also want to have (or not to have) certain desires and motives. They are capable of wanting to be different, in their preferences and purposes, from what they are. Many animals appear to have the capacity for what I shall call "first-order desires" or "desires of the first order," which are simply desires to do or not to do one thing or another. No animal other than man, however, appears to have the capacity for reflective self-evaluation that is manifested in the formation of second-order desires.²

I

... Consider first those statements of the form "*A* wants to *X*" which identify first-order desires—that is, statements in which the term 'to *X*' refers to an action. A statement of this kind does not, by itself, indicate the relative strength of *A*'s desire to *X*. It does not make it clear whether this desire is at all likely to play a decisive role in what *A* actually does or tries to do. For it may correctly be said that *A* wants to *X* even when his desire to *X* is only one among his desires and when it is far from being paramount among them. Thus, it may be true that *A* wants to *X* when he strongly prefers to

do something else instead; and it may be true that he wants to *X* despite the fact that, when he acts, it is not the desire to *X* that motivates him to do what he does. On the other hand, someone who states that *A* wants to *X* may mean to convey that it is this desire that is motivating or moving *A* to do what he is actually doing or that *A* will in fact be moved by this desire (unless he changes his mind) when he acts.

It is only when it is used in the second of these ways that, given the special usage of 'will' that I propose to adopt, the statement identifies *A*'s will. To identify an agent's will is either to identify the desire (or desires) by which he is motivated in some action he performs or to identify the desire (or desires) by which he will or would be motivated when or if he acts. An agent's will, then, is identical with one or more of his first-order desires. But the notion of the will, as I am employing it, is not coextensive with the notion of first-order desires. It is not the notion of something that merely inclines an agent in some degree to act in a certain way. Rather, it is the notion of an *effective* desire—one that moves (or will or would move) a person all the way to action. Thus the notion of the will is not coextensive with the notion of what an agent intends to do. For even though someone may have a settled intention to do *X*, he may nonetheless do something else instead of doing *X* because, despite his intention, his desire to do *X* proves to be weaker or less effective than some conflicting desire.

Now consider those statements of the form "*A* wants to *X*" which identify second-order desires—that is, statements in which the term 'to *X*' refers to a desire of the first order. There are also 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.

Suppose that a physician engaged in psychotherapy with narcotics addicts believes that his ability to help his patients would be

enhanced if he understood better what it is like for them to desire the drug to which they are addicted. Suppose that he is led in this way to want to have a desire for the drug. If it is a genuine desire that he wants, then what he wants is not merely to feel the sensations that addicts characteristically feel when they are gripped by their desires for the drug. What the physician wants, insofar as he wants to have a desire, is to be inclined or moved to some extent to take the drug.

It is entirely possible, however, that, although he wants to be moved by a desire to take the drug, he does not want this desire to be effective. He may not want it to move him all the way to action. He need not be interested in finding out what it is like to take the drug. And insofar as he now wants only to *want* to take it, and not to *take* it, there is nothing in what he now wants that would be satisfied by the drug itself. He may now have, in fact, an altogether univocal desire *not* to take the drug; and he may prudently arrange to make it impossible for him to satisfy the desire he would have if his desire to want the drug should in time be satisfied.

It would thus be incorrect to infer, from the fact that the physician now wants to desire to take the drug, that he already does desire to take it. His second-order desire to be moved to take the drug does not entail that he has a first-order desire to take it. If the drug were now to be administered to him, this might satisfy no desire that is implicit in his desire to want to take it. While he wants to want to take the drug, he may have *no* desire to take it; it may be that *all* he wants is to taste the desire for it. That is, his desire to have a certain desire that he does not have may not be a desire that his will should be at all different than it is.

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. In such cases the statement means that *A* wants the desire to *X*

to be the desire that moves him effectively to act. 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.⁴

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. But the question of whether or not his second-order desire is fulfilled does not turn merely on whether the desire he wants is one of his desires. It turns on whether this desire is, as he wants it to be, his effective desire or will. If, when the chips are down, it is his desire to concentrate on his work that moves him to do what he does, then what he wants at that time is indeed (in the relevant sense) what he wants to want. If it is some other desire that actually moves him when he acts, on the other hand, then what he wants at that time is not (in the relevant sense) what he wants to want. This will be so despite the fact that the desire to concentrate on his work continues to be among his desires.

II

Someone has a desire of the second order either when he wants simply to have a certain desire or when he wants a certain desire to be his will. In situations of the latter kind, I shall call his second-order desires "second-order volitions" or "volitions of the second order." Now it is having second-order volitions, and not having second-order desires generally, that I regard as essential to being a person. It is logically possible, however unlikely, that there should be an agent with second-order desires but with no volitions of the

second order. Such a creature, in my view, would not be a person. I shall use the term 'wanton' to refer to agents who have first-order desires but who are not persons because, whether or not they have desires of the second order, they have no second-order volitions.⁵

The essential characteristic of a wanton is that he does not care about his will. His desires move him to do certain things, without its being true of him either that he wants to be moved by those desires or that he prefers to be moved by other desires. The class of wantons includes all nonhuman animals that have desires and all very young children. Perhaps it also includes some adult human beings as well. In any case, adult humans may be more or less wanton; they may act wantonly, in response to first-order desires concerning which they have no volitions of the second order, more or less frequently.

The fact that a wanton has no second-order volitions does not mean that each of his first-order desires is translated heedlessly and at once into action. He may have no opportunity to act in accordance with some of his desires. Moreover, the translation of his desires into action may be delayed or precluded either by conflicting desires of the first order or by the intervention of deliberation. For a wanton may possess and employ rational faculties of a high order. Nothing in the concept of a wanton implies that he cannot reason or that he cannot deliberate concerning how to do what he wants to do. What distinguishes the rational wanton from other rational agents is that he is not concerned with the desirability of his desires themselves. He ignores the question of what his will is to be. Not only does he pursue whatever course of action he is most strongly inclined to pursue, but he does not care which of his inclinations is the strongest.

Thus a rational creature, who reflects upon the suitability to his desires of one course of action or another, may nonetheless be a wanton. In maintaining that the essence of being a person lies not in reason but in will, I am far from suggesting that a creature without reason may be a person. For it is only in virtue of his

rational capacities that a person is capable of becoming critically aware of his own will and of forming volitions of the second order. The structure of a person's will presupposes, accordingly, that he is a rational being.

The distinction between a person and a wanton may be illustrated by the difference between two narcotics addicts. Let us suppose that the physiological condition accounting for the addiction is the same in both men, and that both succumb inevitably to their periodic desires for the drug to which they are addicted. One of the addicts hates his addiction and always struggles desperately, although to no avail, against its thrust. He tries everything that he thinks might enable him to overcome his desires for the drug. But these desires are too powerful for him to withstand, and invariably, in the end, they conquer him. He is an unwilling addict, helplessly violated by his own desires.

The unwilling addict has conflicting first-order desires: he wants to take the drug, and he also wants to refrain from taking it. In addition to these first-order desires, however, he has a volition of the second order. He is not neutral with regard to the conflict between his desire to take the drug and his desire to refrain from taking it. It is the latter desire, and not the former, that he wants to constitute his will; it is the latter desire, rather than the former, that he wants to be effective and to provide the purpose that he will seek to realize in what he actually does.

The other addict is a wanton. His actions reflect the economy of his first-order desires, without his being concerned whether the desires that move him to act are desires by which he wants to be moved to act. If he encounters problems in obtaining the drug or in administering it to himself, his responses to his urges to take it may involve deliberation. But 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. ...

III

There is a very close relationship between the capacity for forming second-order volitions and another capacity that is essential to persons—one that has often been considered a distinguishing mark of the human condition. It is only because a person has volitions of the second order that he is capable both of enjoying and of lacking freedom of the will. The concept of a person is not only, then, the concept of a type of entity that has both first-order desires and volitions of the second order. It can also be construed as the concept of a type of entity for whom the freedom of its will may be a problem. This concept excludes all wantons, both infrahuman and human, since they fail to satisfy an essential condition for the enjoyment of freedom of the will. And it excludes those suprahuman beings, if any, whose wills are necessarily free.

Just what kind of freedom is the freedom of the will? This question calls for an identification of the special area of human experience to which the concept of freedom of the will, as distinct from the concepts of other sorts of freedom, is particularly germane. In dealing with it, my aim will be primarily to locate the problem with which a person is most immediately concerned when he is concerned with the freedom of his will.

According to one familiar philosophical tradition, being free is fundamentally a matter of doing what one wants to do. Now the notion of an agent who does what he wants to do is by no means an altogether clear one: both the doing and the wanting, and the appropriate relation between them as well, require elucidation. But although its focus needs to be sharpened and its formulation refined, I believe that this notion does capture at least part of what is implicit in the idea of an agent who *acts* freely. It misses entirely, however, the peculiar content of the quite different idea of an agent whose *will* is free.

We do not suppose that animals enjoy freedom of the will, although we recognize that an animal may be free to run in

whatever direction it wants. Thus, having the freedom to do what one wants to do is not a sufficient condition of having a free will. It is not a necessary condition either. For to deprive someone of his freedom of action is not necessarily to undermine the freedom of his will. When an agent is aware that there are certain things he is not free to do, this doubtless affects his desires and limits the range of choices he can make. But suppose that someone, without being aware of it, has in fact lost or been deprived of his freedom of action. Even though he is no longer free to do what he wants to do, his will may remain as free as it was before. Despite the fact that he is not free to translate his desires into actions or to act according to the determinations of his will, he may still form those desires and make those determinations as freely as if his freedom of action had not been impaired.

When we ask whether a person's will is free we are not asking whether he is in a position to translate his first-order desires into actions. That is the question of whether he is free to do as he pleases. The question of the freedom of his will does not concern the relation between what he does and what he wants to do. Rather, it concerns his desires themselves. But what question about them is it?

It seems to me both natural and useful to construe the question of whether a person's will is free in close analogy to the question of whether an agent enjoys freedom of action. Now freedom of action is (roughly, at least) the freedom to do what one wants to do. Analogously, then, the statement that a person enjoys freedom of the will means (also roughly) that he is free to want what he wants to want. More precisely, it means that he is free to will what he wants to will, or to have the will he wants. Just as the question about the freedom of an agent's action has to do with whether it is the action he wants to perform, so the question about the freedom of his will has to do with whether it is the will he wants to have.

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 in 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. The unwilling addict's will is not free. This is shown by the fact that it is not the will he wants. It is also true, though in a different way, that the will of the wanton addict is not free. The wanton addict neither has the will he wants nor has a will that differs from the will he wants. Since he has no volitions of the second order, the freedom of his will cannot be a problem for him. He lacks it, so to speak, by default. ...

IV

My theory concerning the freedom of the will accounts easily for our disinclination to allow that this freedom is enjoyed by the members of any species inferior to our own. It also satisfies another condition that must be met by any such theory, by making it apparent why the freedom of the will should be regarded as desirable. The enjoyment of a free will means the satisfaction of certain desires—desires of the second or of higher orders—whereas its absence means their frustration. The satisfactions at stake are those which accrue to a person of whom it may be said that his will is his own. The corresponding frustrations are those suffered by a person of whom it may be said that he is estranged from himself, or that he finds himself a helpless or a passive bystander to the forces that move him.

A person who is free to do what he wants to do may yet not be in a position to have the will he wants. Suppose, however, that he enjoys both freedom of action and freedom of the will. Then he is not only free to do what he wants to do; he is also free to want what he wants to want. It seems to me that he has, in that case, all the freedom it is possible to desire or to conceive. There are other good things in life, and he may not possess some of them. But there is nothing in the way of freedom that he lacks.

It is far from clear that certain other theories of the freedom of the will meet these elementary but essential conditions: that it be

understandable why we desire this freedom and why we refuse to ascribe it to animals. Consider, for example, Roderick Chisholm's quaint version of the doctrine that human freedom entails an absence of causal determination. 7 Whenever a person performs a free action, according to Chisholm, it's a miracle. The motion of a person's hand, when the person moves it, is the outcome of a series of physical causes; but some event in this series, "and presumably one of those that took place within the brain, was caused by the agent and not by any other events" (18). A free agent has, therefore, "a prerogative which some would attribute only to God: each of us, when we act, is a prime mover unmoved" (23).

This account fails to provide any basis for doubting that animals of subhuman species enjoy the freedom it defines. Chisholm says nothing that makes it seem less likely that a rabbit performs a miracle when it moves its leg than that a man does so when he moves his hand. But why, in any case, should anyone *care* whether he can interrupt the natural order of causes in the way Chisholm describes? Chisholm offers no reason for believing that there is a discernible difference between the experience of a man who miraculously initiates a series of causes when he moves his hand and a man who moves his hand without any such breach of the normal causal sequence. There appears to be no concrete basis for preferring to be involved in the one state of affairs rather than in the other. ...

My conception of the freedom of the will appears to be neutral with regard to the problem of determinism. It seems conceivable that it should be causally determined that a person is free to want what he wants to want. If this is conceivable, then it might be causally determined that a person enjoys a free will. There is no more than an innocuous appearance of paradox in the proposition that it is determined, ineluctably and by forces beyond their control, that certain people have free wills and that others do not. There is no incoherence in the proposition that some agency other than a person's own is responsible (even *morally* responsible) for the fact that he enjoys or fails to enjoy freedom of the will. ...