
THE JOURNAL OF PHILOSOPHY

VOLUME LXXXV, NO. 8, AUGUST 1988

SANCTIFICATION, HARDENING OF THE HEART, AND FRANKFURT'S CONCEPT OF FREE WILL*

IN a much-discussed paper,¹ "Freedom of the Will and the Concept of a Person,"² Harry Frankfurt presents an analysis of the self in terms of hierarchically ordered desires, and he uses his analysis to argue for a certain notion of freedom. His paper has generated considerable debate among philosophers interested in the concept of freedom and the related concept of autonomy. My own interest in Frankfurt's paper is primarily in applying his analysis of the self to problems in the philosophy of religion, especially to puzzles raised by the doctrine of sanctification and the notion of God's

* I am indebted to the following people for useful suggestions: Don Adams, John Christman, Philip Quinn, Bruce Russell, John Tyson, Peter van Inwagen, and Allen Wood. I owe a special debt to Harry Frankfurt, who corrected some important misunderstandings of his account, and to William Alston, whose objections helped me to think out some essential distinctions. Finally, I am particularly grateful to Norman Kretzmann, who gave me many helpful comments on earlier drafts.

¹ For good summaries of the literature, see David Shatz, "Free Will and the Structure of Motivation," *Midwest Studies in Philosophy*, vol. x, Peter French, Theodore Uehling, Jr., and Howard Wettstein, eds. (Minneapolis: Minnesota UP, 1986), pp. 451-482; and Gary Watson, "Free Action and Free Will," *Mind*, XLVI (1987): 145-172. Besides the literature cited in those articles, I also found helpful the following articles (given in no particular order): Lawrence Haworth, *Autonomy: An Essay in Philosophical Psychology and Ethics* (New Haven: Yale, 1986), and "Autonomy and Utility," *Ethics*, xcv (1984): 5-19; William Rowe, "Two Criticisms of the Agency Theory," *Philosophical Studies*, XLII (1982): 363-378; Richard Arneson, "Freedom and Desire," *Canadian Journal of Philosophy*, xv (1985): 425-448; S. I. Benn, "Freedom, Autonomy and the Concept of a Person," *Proceedings of the Aristotelian Society*, LXXVI (1975/6): 109-131; Gerald Dworkin, "The Concept of Autonomy," in *Science and Ethics*, R. Haller, ed. (Amsterdam: Rodopi, 1981), pp. 203-213; Irving Thalberg, "Socialization and Autonomous Behavior," in *Studies in Action Theory*, Robert Whittlemore, ed. (New Orleans: Tulane, 1979), pp. 21-37; Marilyn Friedman, "Autonomy and the Split-level Self," *The Southern Journal of Philosophy*, xxiv (1986): 19-35; Susan Wolf, "The Importance of Free Will," *Mind*, xc (1981): 386-405.

² This JOURNAL, LXVIII, 1 (January 14, 1971): 5-20. Other papers by Frankfurt which have some bearing on the issues discussed here are the following (given in no

hardening hearts, although I think Frankfurt's analysis is remarkably fruitful for understanding a variety of religious claims and practices. In order to apply Frankfurt's views to issues in the philosophy of religion, however, it is important to reconsider his understanding of freedom and to refine his hierarchical analysis of the self. The resulting revised version of Frankfurt's account of freedom and the self is not vulnerable to the sorts of criticisms which have been leveled, quite correctly I think, against Frankfurt's original views. So in what follows I shall first consider Frankfurt's views of freedom and the self and suggest some revisions of them. Then I shall discuss the most important criticisms raised against Frankfurt's original position and argue that the revised Frankfurt account can be successfully defended against them. Finally, I shall show how that account can resolve some long-standing difficulties about sanctification and hardening of the heart.

I. FRANKFURT'S ACCOUNT

It is Frankfurt's view that the essence of a person is to be found in the structure of the will. He takes wants or desires to be the genus of acts of willing, or volitions, and he holds a volition to be an effective desire, which moves an agent all the way to action. According to Frankfurt, agents can have first-order desires and volitions—to do something—and also second-order desires and volitions—to have certain first-order desires. A person, on Frankfurt's view, is someone who has second-order desires and volitions. An agent who has no second-order volitions is "a wanton"; such an individual may be human but is not a person.

This analysis of the notion of a person is the basis for Frankfurt's account of freedom of will. The common conception of freedom as the ability to do what one wants to do, Frankfurt says, is best thought of as applying to freedom of action. Freedom of will can then be construed analogously as the ability to will what one wants to will, or the ability to have the sort of will one wants. On Frankfurt's view, in order to have freedom of will, an individual must meet the following

particular order): "Necessity and Desire," *Philosophy and Phenomenological Research*, XLV (1984): 1-13; "The Importance of What We Care About," *Synthese*, LIII (1982): 257-272; "The Problem of Action," *American Philosophical Quarterly*, xv (1978): 157-162; "Coercion and Moral Responsibility," in *Essays on Freedom of Action*, Ted Honderich, ed. (London: Routledge & Kegan Paul, 1973), pp. 65-86; "Three Concepts of Free Action II," in *Proceedings of the Aristotelian Society*, supp. vol. XLIX (1975): 113-125; "Alternate Possibilities and Moral Responsibility," this JOURNAL, LXVI, 23 (December 4, 1969): 828-839; "Identification and Externality," in *The Identities of Persons*, Amelie Rorty, ed. (Berkeley: California UP, 1976); "The Problem of Action," *American Philosophical Quarterly*, xv (1978): 157-162; and "Identification and Wholeheartedness," forthcoming (I am grateful to Frankfurt for letting me see this paper in typescript).

conditions: (1) he has second-order volitions, (2) he does not have first-order volitions that are discordant with those second-order volitions, and (3) he has the first-order volitions he has *because of* his second-order volitions (that is, his second-order volitions have, directly or indirectly, produced his first-order volitions; and if his second-order volitions had been different, he would have had different first-order volitions).³

This account might appear to imply that only a person who meets these strong criteria for freedom of the will is morally responsible for his actions, but Frankfurt is not committed to such a counter-intuitive view. He distinguishes between acting freely and having freedom of will when one acts. If a person has done what he wanted to do because he wanted to do it and the will by which he was moved when he did it was his own will, then he *acted* freely, even if he did not act with *freedom of will*. Assessments of moral responsibility, according to Frankfurt, should depend primarily on whether or not an agent acted freely, rather than on whether or not he acted with free will.

II. IMPLICATIONS OF FRANKFURT'S ACCOUNT

Following Frankfurt's lead, we can take the basic notion of freedom as the absence of obstacles to what one wants.⁴ The expressions 'free from' and 'free to' can then be seen as two branches of the same basic notion. Locutions involving the expression 'free from' specify which obstacles are absent, and locutions involving the expression 'free to' indicate the range of things available to the agent to do without obstacle. What Frankfurt and others call "freedom of action" is the absence of obstacles to doing what one wants to do; freedom of will is the absence of obstacles to willing what one wants to will.

As Frankfurt's work makes clear, obstacles to doing what one wants to do can arise in two ways. They may have their origin in something external to the agent, such as social institutions, or in something internal to the agent, such as psychoses. It is easy to show that there can also be external and internal obstacles to freedom of will by adapting the familiar example of a man who chooses to stay in a room, unaware that the door is locked.⁵ Suppose that the man's

³ I am grateful to both Frankfurt and Alston for suggestions regarding these conditions.

⁴ Making precise the rather vague intuition behind this claim would take some doing and is beyond the scope of this paper. Perhaps it is enough for present purposes to say that by 'obstacle' here is surely meant something like "an obstacle which is, for all practical purposes, at the moment, physically insuperable."

⁵ Discussed in Frankfurt, "Alternate Possibilities and Moral Responsibility," *op. cit.*; cf. John Martin Fischer, "Responsibility and Control," this JOURNAL, LXXIX, 1 (January 1982): 24-40.

volition to stay in the room is produced in him by means of some device implanted in his brain and controlled by scientists who want him to stay in the room. In this case what the agent wills is not what he himself wants to will; he wills what the scientists want him to will. There is thus an external obstacle to the agent's willing what he wants to will, and so he does not act with free will in staying in the room.

Ordinarily when we ask whether a person acts with free will, what we are asking is in effect whether there is an external constraint of this sort on his will. The issue between compatibilists and libertarians, for example, can be understood at least in part as a dispute over whether the causal influences that compatibilists claim operate on a person constitute an external constraint on his will. Frankfurt's definition of freedom of will is a strong one, because, in order to have free will in his sense of the term, something more is required than the simple absence of external obstacles to willing what one wants. What else is needed can be seen by considering a revised version of the example of the man in the room. This time suppose that the man does want to leave the room—say, in order to get to his classroom to teach—and that the door is not locked, but that there is a black cat asleep on the lintel over the door and he is superstitious about black cats. He struggles with his superstitious fear, but finally his desire not to have to walk past the black cat gets the better of him; he gives up the struggle and wills to stay in the room. As before, he stays willingly, but he does not stay with freedom of will, in Frankfurt's sense of the term. In this case, however, the obstacle to his willing what he wants to will is internal rather than external, because it is his own desires which are the impediment to his freedom of will. Frankfurt's sense of 'freedom of will' is thus different from the ordinary sense of the term which we have in mind when we ask, for example, whether Aquinas believes in free will rather than theological determinism. It seems to me more nearly the sense of 'free' in the theological claim "You shall know the truth and the truth shall make you free." If it were not for the clumsy locution, we might call Frankfurt's sense 'complete freedom of will' since it encompasses and exceeds the ordinary sense of free will as absence of external obstacles to willing what one wants.⁶

There are, then, four basic sorts of obstacles to what one wants and four corresponding basic senses in which someone can be free in virtue of the absence of one or another of those obstacles:

⁶ Frankfurt's account is usually employed by those who want to defend compatibilism, but both the revised account and Frankfurt's original views seem to me perfectly consistent also with incompatibilism.

- (1) having no external obstacles to doing what one wants to do,
- (2) having no internal obstacles to doing what one wants to do,
- (3) having no external obstacles to willing what one wants to will, and
- (4) having no internal obstacles to willing what one wants to will.

When an agent lacks freedom in one of the first three senses, we generally do not hold him morally responsible for what he does or fails to do. We would not blame a student for failing to complete his studies if he could not attend the university for lack of funds or if psychosis rendered him incompetent to study. Similarly, we would not hold the man responsible for failing to try to leave the room if his willing to stay in the room were technologically induced by scientists.⁷ But most of us would be inclined to blame the faculty member whose failure to meet his class stemmed from a superstitious fear of black cats, and the fact that we would is instructive. The division between freedom in the first three senses and freedom in the fourth sense is similar to Frankfurt's distinction between acting freely and acting with freedom of will. Someone who has freedom in the first three senses, even if he lacks it in the fourth sense, acts freely and so is responsible for his action. Only someone who also has freedom in the fourth sense, however, acts with freedom of will, in Frankfurt's strong sense of free will.

III. THE REVISED FRANKFURT ACCOUNT

On Frankfurt's analysis, the concept of a person is marked by a single hierarchical distinction, between an agent's first-order desires and volitions and his higher-order desires and volitions. But traditionally philosophers have analyzed a human person into three parts: desires or passions, will, and intellect. If we revise Frankfurt's view to take account of the role of intellect, I think we strengthen it without losing any of its explanatory power. Aquinas held that an agent wills to do some action *p* (or bring about some state of affairs *q*) only if the agent's intellect at the time of the action represents *p* (or *q*), under some description, as the good to be pursued.⁸ Here it is crucial to understand that, in this context, 'an agent's intellect' or 'an agent's reasoning' does not refer to something which is solely rational. By 'an

⁷ There are, of course, complicated cases in which the agent's lack of freedom in one of these respects is ultimately attributable to the agent himself. The student might be impoverished because he lost all his money gambling, or his psychosis might be one he brought on himself as a result of his use of certain drugs known to cause psychosis. In such cases, the agent's inability to do what he wants is a direct result of something the agent does want. To the extent to which the inability is tied to what the agent wants, the inability does not diminish moral responsibility even if it does diminish the agent's freedom at the time of his inability.

⁸ Aquinas's understanding of the relation of will to intellect is complicated and cannot be adequately discussed here. Basically, his view is that the intellect moves

agent's intellect' I mean just the computing faculty of an agent. So understood, an agent's intellect may formulate a reason for an action in a manner that is hasty, thoughtless, ill-informed, invalid, or in any other way irrational.

Furthermore, it is important to understand that an agent's reason for an action may also be only implicit and not an explicit or conscious feature of his thought. In a recent paper, Robert Audi⁹ has argued cogently that x may count as the reason for an agent's action even when the agent has not consciously formulated some reason x as a reason for his action but would nonetheless give x as the reason for what he did if asked for an explanation. On this view, then, it is possible that an agent's intellect have gone through some process which contributes to a certain action on the agent's part, without the agent's being aware of that process as it is occurring. So to hold, as Aquinas does, that an agent wills to do some action p only if his intellect represents p as the good to be pursued does not entail that an agent does an action willingly only in case he first engages in a conscious process of reasoning about the action.¹⁰ Aquinas's view requires only that some chain of reasoning (even if invalid and irrational reasoning) representing p as the good to be pursued would figure in the agent's own explanation of his action. In what follows, discussions of an agent's reasoning should be understood in light of these caveats: the reasoning in question need not be either rational or conscious.

In the spirit of Aquinas's view of intellect's direction of the will, we can make the following first revision of Frankfurt's account of the self. An agent has a second-order volition V2 to bring about some first-order volition V1 in himself only if the agent's intellect at the time of the willing represents V1, under some description, as the good to be pursued. A second-order volition, then, is a volition formed as a result of some reasoning (even when the reasoning is neither rational nor conscious) about one's first-order desires.

It will be helpful to make one other revision in Frankfurt's account. His use of the term 'second-order volition' is ambiguous between an agent's second-order desire that is effective in moving

the will by presenting it with an understanding of the good. In so moving the will, the intellect acts not as an efficient cause but rather as a final cause. The will is a natural inclination or appetite for the good; and the intellect moves the will, without coercion, by showing it what the good to be pursued is in a particular set of circumstances. But the will also moves the intellect directly, as an efficient cause—for example, by directing it to consider certain things and to neglect others. Cf., e.g., ST Ia, q.82, a.4.

⁹ "Acting for Reasons," *The Philosophical Review*, xcv (1986): 511–546.

¹⁰ I am grateful to Bruce Russell for raising this issue in correspondence.

him to make the corresponding first-order desire his will and a second-order desire that is not effective in that way. On Frankfurt's account, an unwilling addict who wants not to have the desire for heroin be his will but who nonetheless succumbs to his heroin addiction has a second-order volition for a desire not to take heroin. So, however, does the reformed addict who wants not to have the desire for heroin be his will and who has succeeded in that endeavor. When discussing first-order desires, Frankfurt identifies an agent's first-order volition as a first-order desire which the agent makes his will and on which he acts. Consequently, although an agent may simultaneously have conflicting first-order desires, he cannot simultaneously have conflicting first-order volitions. For the sake of clarity in what follows, I want to make a second revision of Frankfurt's account by disambiguating the sense of 'second-order volition' along the lines of Frankfurt's distinction between first-order desires and first-order volitions.

If an effective desire is one which moves the agent all the way to action, then an effective second-order desire is one which moves the agent all the way to the action of making the corresponding first-order desire his will. So a second-order desire constitutes a second-order volition only if it is an effective desire and the agent has a first-order volition corresponding to it. On this usage, the reformed addict has a second-order volition not to have the desire for heroin be his will, but the unwilling addict who succumbs to his addiction does not. He has a second-order desire not to have the desire for heroin be his will; but because the second-order desire is not an effective desire, it does not constitute a second-order volition. To express Frankfurt's concept of freedom using this revised understanding of second-order desires and volitions, we should say that an individual has freedom of the will just in case he has second-order desires, his first-order volitions are not discordant with his second-order desires, and he has the first-order volitions he has because of his second-order volitions.

As is the case with first-order volitions, it is not possible for an agent to have conflicting second-order volitions, but it is possible for him to have conflicting second-order desires. We might suppose, for example, that Verkhovensky in Dostoyevsky's *The Possessed* has both a second-order desire to have a desire for gambling, because the desire for gambling will make him well-liked by his friends, and a second-order desire not to have a desire to gamble, because stamping out the desire for gambling will win the admiration of Mrs. Stavrogin. It is worth noticing that where second-order desires conflict, it will not be possible for an agent to act on the corresponding first-order desires with freedom of will. Whether Verkhovensky has a

first-order volition to gamble or a first-order volition not to gamble, his will with regard to that volition is not free. In either case his first-order volition is opposed to one of his second-order desires, and so he fails to meet one of the criteria for freedom of will.¹¹

IV. CRITICISMS OF FRANKFURT & THE REVISED FRANKFURT ACCOUNT
Of the many criticisms that have been raised against Frankfurt's account, three appear recurrently in the literature and seem especially worth considering. Various critics have charged that (1) on Frankfurt's account, an agent would be free with respect to a certain volition that was in accord with some higher-order volition even if that higher-order volition were directly produced somehow by someone else. But this is surely a counterintuitive result, because such an agent seems to be just as much the puppet of his manipulator as the agent whose first-order volitions are directly produced by someone else, and no one would want to say that the agent in the latter case has free will.¹² Another recurrent complaint is that (2) Frankfurt's account of free will leaves us with an infinite regress of volitions. For an agent to act with free will with respect to some first-order volition

¹¹ The revised Frankfurt account has some affinities with Wright Neely's account of freedom, which holds that "an agent is free with respect to some action which he performed only if it is true that, if he had been given what he took to be good and sufficient reason for not doing what he did, he would not have done it" [Neely, "Freedom and Desire," *The Philosophical Review*, LXXXIII (1974):48]. On the revised Frankfurt account, an agent may have what he would agree, if asked, is good and sufficient reason for not doing some action *x* and yet still do *x*, because, under the sway of the passions (to take just one example), he interprets doing *x* at this time under some description which makes it mistakenly seem as if the good and sufficient reason does not apply to this particular action. But he would not be free with respect to *x*, even though he might be morally responsible for doing it, since in having what he takes to be good and sufficient reason against doing *x*, he also has a second-order desire not to do *x*. Neely has sometimes been interpreted as holding that what an agent really desires is what he desires when he is thinking rationally, i.e., without epistemic error. If this interpretation were correct (and I think it misses the force of the phrase 'what he took to be' in the preceding quotation from Neely's article), then the revised Frankfurt account would be opposed to Neely's, rather than similar to it, since on the revised Frankfurt account what an agent really desires can be based on a process of reasoning full of epistemic error. The revised Frankfurt account is also in some respects similar to the account of freedom given by Watson in his insightful paper "Free Agency," this JOURNAL, LXXII, 8 (April 24, 1975): 205–220. Watson objects to Frankfurt's account on the grounds that Frankfurt has given us no reason to suppose that second-order desires represent what the agent himself wants just in virtue of being second-order. (I shall defend the revised Frankfurt account against this sort of objection in section IV below.) Watson substitutes an account based on a distinction between what an agent desires and what he values, and he takes an agent's values to be those principles and ends which the agent desires when he is being rational. Watson is right to introduce intellect into the hierarchical notion of the self, but I think it is a mistake to suppose that the valuing which play a role in free will are just the agent's rational principles and ends. In "Free Action and Free Will" (*op. cit.*), Watson expresses dissatisfaction with his earlier view for the same reason.

¹² See, e.g., Shatz, *op. cit.*, pp. 468/9.

V1 apparently requires that he have a higher-order volition V2 with which V1 is in accord. But then it seems that V2 itself must be freely willed in order for the agent to be acting with free will; for V2 to be freely willed, however, requires a higher-order volition V3 with which V2 is in accord, and V3 itself will require a higher-order volition V4, and so on.¹³ Finally, some critics have objected to Frankfurt's account on the grounds that (3) it rests on an unwarranted notion of what counts as "the real you" and on a false theory of what counts as external or alien to a person. When we say that freedom is basically a matter of doing what one wants and that what one really wants is determined by considering one's higher-order volitions, we clearly presuppose that the first-order desires not in accord with the second-order volitions are not something the agent himself really wants, that these repudiated first-order desires are not part of the agent's self but somehow external to it. But such an understanding of the real self is at best controversial. Why should we identify an agent's self only with his higher-order volitions? Why should we suppose that these higher-order volitions represent what the agent really wants? Has psychology not made us aware that the darker sides of our nature, the repudiated or repressed first-order desires, are just as much part of our selves as our higher-order desires and the first-order desires we approve of?¹⁴ The revised Frankfurt account is, I think, not vulnerable to any of these criticisms.

On the revised account,¹⁵ an agent forms a second-order desire by reasoning (rationally or otherwise, consciously or not) about his first-order desires; and a second-order desire is a direct result of an agent's intellect representing a certain first-order desire as the good to be pursued. Given this connection between intellect and second-order desires, an agent cannot be a passive bystander to his second-order volitions. To be a second-order volition, a volition must be the result of reasoning on the agent's part. Even if it were coherent to suppose that one agent, say, Verkhovensky, could directly produce some reasoning in the mind of another, such as Stavrogin, that reasoning would not be Stavrogin's but rather Verkhovensky's (or at any rate a product of Verkhovensky's reasoning). If Verkhovensky con-

¹³ See, e.g., David Zimmerman, "Hierarchical Motivation and Freedom of the Will," *Pacific Philosophical Quarterly*, LXII (1981): 358ff.

¹⁴ See, e.g., Thalberg, "Hierarchical Analyses of Unfree Action," *Canadian Journal of Philosophy*, VIII (1978): 211-225.

¹⁵ For Frankfurt's own response, see "Three Concepts of Free Action II," pp. 121/2.

tinuously produced thoughts in Stavrogin, then Stavrogin would have ceased to be a person and would instead be something like Verkhovensky's puppet. On the other hand, suppose Verkhovensky produced thoughts in Stavrogin's mind only occasionally, so that Stavrogin remained a person. In the computations leading to an action, Stavrogin's own intellect would take cognizance of the thought Verkhovensky had produced in Stavrogin's mind, and Stavrogin would then either accept or reject Verkhovensky's thought as a result of Stavrogin's own reasoning (however tacit or irrational that reasoning may be). As Stavrogin acts, then, the first-order volition stemming from his reasoning and the accompanying second-order desire will be Stavrogin's, not Verkhovensky's. Either way, Stavrogin would not have any second-order volitions produced by Verkhovensky. So, on the revised Frankfurt account, an agent's second-order volitions cannot be produced by someone else.

As for objection (2), Frankfurt himself believes that there is no theoretical limit to the levels of higher-order desires which a person may have, and that in general only common sense and fatigue keep a person from entertaining ever higher levels of higher-order desires.¹⁶ On the revised Frankfurt account, however, the claim that the levels of higher-order desires may be infinite does not hold. In formulating a second-order volition an agent is bringing reason to bear on a state of his will and either approving or rejecting it. But in forming a third-order volition, the agent is not reiterating the process gone through to formulate a second-order volition. On the model the revised Frankfurt account gives of forming second-order volitions, forming a third-order volition consists in reasoning about and either accepting or rejecting a second-order volition. So an agent has a third-order volition V3 to bring about some second-order volition V2 in himself only if his intellect at the time of the willing represents V2, under some description, as the good to be pursued. But since V2 is a desire for a first-order volition V1 generated by reason's representing V1 (at that time) as the good to be pursued, V3 will consist just in reaffirming the original reasoning about V1 which led to V2. In forming a third-order volition and considering whether he wants to have the relevant second-order volition, the agent will consider whether a desire for a desire for

¹⁶ See "Freedom of the Will and the Concept of a Person," *op. cit.*, p. 16. Frankfurt returns to this problem in his forthcoming paper, "Identification and Wholeheartedness," where he analyzes the notion of wholeheartedness and uses it as the basis for a response to this problem.

some action p (or state of affairs q) is the good to be pursued. But a desire for a desire for p (or q) will be a good to be pursued just in case the desire for p (or q) is a good to be pursued, and that in turn will depend on whether the agent considers p (or q), under some description, at that time, a good to be pursued. So a third-order volition that supports a currently held second-order volition is in effect just the expression of a reevaluating and affirming of the reasoning that originally led to VI.¹⁷ And, in the same way, a third-order volition that rejects a currently held second-order volition will just be an expression of the reevaluation and rejection of the reasoning that led to the second-order volition. A third-order volition, then, is a result of a recalculation of the reasoning that originally underlay a second-order volition.

I do not mean to suggest that third-order volitions must always collapse into second-order volitions. It is possible for an agent to have third-order desires that are not only distinct from, but even discordant with, his second-order desires. This is so because volitions and desires, like emotions, are not always *immediately* responsive to reasoning. Even after we are quite sure that a danger is entirely past, we may nonetheless continue to feel some fear; and for a while, until the emotion subsides, we may need to remind ourselves recurrently that there is no cause for fear. Similarly, a second-order desire may take time to fade even when the agent has repudiated the reasoning that generated it. Consider someone whose childhood among Southern Baptists has left him with a desire to avoid alcohol and a second-order desire to have such a desire. Suppose that this person subsequently repudiates his childhood religion and, among other things, joins a sailing club where beer is regularly served after races. He will then have a first-order desire to drink beer, in order to fit in and be companionable and perhaps also to indulge a newly acquired taste for beer; but he may also notice in himself, as a surviving trace of childhood inclination, guilt at drinking and a second-order desire to have the sort of will which wills not to drink. Then he will remind himself that he has repudiated his Baptist background and that there is no harm and even some positive good in drinking a few beers. In this case, then, he will have a third-order desire not to have the second-order desire bequeathed him by his upbringing. His third-

¹⁷ It is, of course, also possible that, in a case of the sort I have been describing, an agent might reevaluate his original reasoning, reject it as unsatisfactory, and yet adopt the same second-order volition as before, although for different reasons. In such a case, the third-order volition is a result of reevaluating the reasoning and reaffirming not the original reasoning but rather just its conclusion.

order desire is thus a temporary measure to bring his second-order desire into line with a change in his reasoning, when he has repudiated his former reasoning and found his second-order desires slow in adjusting to the change.¹⁸

Apart from such cases, however, a third-order (or any higher order) desire or volition will collapse into a second-order desire or volition. Even in the case of the beer-drinking lapsed Baptist, the third-order desire will always be just a reflection of a certain second-order desire. The lapsed Baptist's third-order desire to have a second-order volition that wills to will drinking is just a result of his reasoning that (at the time of the willing, under some description) drinking is a good to be pursued, that a desire for drinking is consequently also a good to be pursued, and that therefore his second-order desire for a will that wills not to drink is a desire to be repudiated. If this Baptist now returns to his childhood convictions and again becomes convinced that drinking is wrong, he will not now form a fourth-order volition not to have his previous third-order volition. Instead, he will have once more reevaluated the reasoning behind the desire to drink and this time rejected it. He will consequently change his second-order volition to accommodate this new alteration in his reasoning, but he will not form a fourth-order volition. The third-order volition stemmed from his efforts to repress the second-order desire habitual from his youth, and his change of beliefs will carry with it the cessation of those efforts. Any attempt, then, to describe his state in terms of a fourth-order (or even higher-order) volition will collapse into the formulation of a second-order volition. So, on the revised Frankfurt account, the number of levels of higher-order desires is not infinite but is rather limited to two or three.

¹⁸ There are also cases in which an agent's reasoning is confused and warrants conflicting second-order desires. An agent who notices such a conflict in his second-order desires and who reflects on it may then sort out the confusion in his reasoning and form a third-order volition in consequence. Nothing in this account entails that the agent's sorting out of his reasoning and accepting only one side of his divided second-order desires will be stable or permanent. In Dostoevsky's novel *The Possessed*, Verkhovensky swings from being contemptuous of Mrs. Stavrogin and rejecting the desires designed to please her to being infatuated with her and yearning to have the desires that will win her admiration. In fact, the revised Frankfurt account and Dostoevsky's portrayal of the vacillating Verkhovensky give a vivid explanation of the line "a double-minded man is unstable in all his ways." We might add that such a man is also unfree. Because he has conflicting second-order desires, whatever his corresponding first-order volition is, it will be discordant with one of his second-order desires.

Freedom of the will on the revised Frankfurt account should consequently not be taken as entailing that the will is free only if an act of willing is in accord with some higher-order volition. On this view, the will is free with respect to a volition *V* just in case *V* is accepted by the agent because his intellect approves of *V* (at that time, under some description) as the good to be pursued, and there is no higher-order desire of the agent's with which *V* is discordant.¹⁹ Second-order desires can fit this definition of freedom without postulating higher-order desires over them. A second-order desire is itself an expression of the agent's reasoning and therefore *eo ipso* accepted by the agent as approved by his reasoning.²⁰ If an agent's reasoning approves a desire as the good to be pursued, it must also in the very same process approve the desire for that desire. An agent who at one and the same time unambiguously considered a certain desire as a good to be pursued and also rejected the desire for that desire as not good would not be sane. So, on the revised Frankfurt account, a second-order volition may be a free volition without itself being the object of some higher-order volition.

Objection (3) has been given its most forceful presentation by Irving Thalberg.²¹ Why, he asks, should we identify ourselves with our higher-order volitions? Have psychologists (and, in particular, Freud) not shown us that the "darker, savage, and nonrational aspects [of ourselves] are equally—if not more—important"? (*ibid.*, p. 224). And if an agent is to be identified with certain of his darker first-order desires, then, in making such desires his will, he would be willing as he himself wants to will. Consequently, Frankfurt is wrong in holding that an agent's freedom depends on his having second-order volitions and governing his first-order desires so that they are not discordant with those second-order volitions. An agent may be doing just what he himself really wants, and so be free, when he acts against his second-order volitions and follows his savage or non-rational desires.

¹⁹ From this definition of free will together with the description of the relation between intellect and will sketched above, it is easy to see that we can generate the three Frankfurt criteria for freedom of the will. When an agent's reasoning approves a first-order volition *V*1 as the good to be pursued at the time of willing, under some description, then the agent's will also forms a second-order desire for that first-order volition; and the intellect's (rational or irrational) approval of *V*1, manifested by the second-order desire, is at least part of what makes *V*1 a volition rather than an ineffective first-order desire.

²⁰ Except for the special sort of case sketched in the example of the beer-drinking lapsed Baptist, where third-order desires fill this role.

²¹ "Hierarchical Analyses of Unfree Action," *op. cit.*

On the revised Frankfurt account, it will not be quite right to say that an agent is to be identified with his second-order volitions. An agent wills what he really wants and is thus free when his first-order volitions are not discordant with his second-order desires, not because the agent is simply declared to be more truly identified with his second-order than with his first-order desires, but rather because the agent's second-order desires are the expressions of his intellect's reflection on his will, and the agent is to be identified with his intellect. It is, of course, possible to recast Thalberg's criticism so that it is directed against the revised version of Frankfurt's account. Why, one could ask, should an agent identify himself with his reasoning faculty rather than with his first-order desires (or his emotions, or subconscious, or any other element of his nature)? But this version of objection (3) is based on a confusion; it depends for its plausibility on a failure to distinguish two different senses of identification.

To see that this is so, consider a fictional variation on the biblical story of Tamar and Amnon (2 Sam. 13:1–20). According to the biblical story, David's son Amnon fell in love with his half-sister Tamar, tried to seduce her, raped her, and then rejected her with hatred. Suppose (contrary to the story in Samuel) that Tamar became pregnant as a result of the rape and bore a son, who quickly grew to look just like his father Amnon, and that no acceptable provisions for the child's care would be available if Tamar rejected him. In such circumstances Tamar would no doubt be torn between conflicting attitudes toward the child. On the one hand, she will recognize that the child is not identical with his father but rather an independent person, who is entirely innocent of the crime that resulted in his conception, and in this spirit she will want to cherish the child and be a good mother to him. On the other hand, when she looks at the child, she will see in him the hated face of the man who raped her; and if she is an ordinary human being, feelings of hatred and revulsion toward the child will rise in her as she recognizes signs of the father in the son. Now, suppose that on one particularly bad day when the child comes running to her, instead of welcoming him she flares up at him for no reason and hits him, because on that occasion the revulsion toward him has gotten the upper hand in her.

As she attempts to sort out her thoughts and feelings after this event, two interpretations of her action are open to her. She could say to herself: "I love my son. How could I treat him in that way? I've lost control of myself." Or she could say: "I can't stand that Amnon-faced child; I hate his father, and the sad truth is that I hate him, too." Tamar has been divided between her first-order desire to

cherish the child and her first-order desire to reject the child; and as she reflects on what she has done, she is in effect asking herself with which half of her divided will she identifies herself. We might be inclined to say that she is not to be identified with either side of her conflicting first-order desires; rather, what she is is a person struggling with a divided will. Such an understanding of Tamar may be correct as regards her past state, but it ceases to be viable once she has attacked the child, because she must then decide how to react to what she has done; and which of the two interpretations Tamar places on her action will make a great difference to the way in which life goes on between her and her son. She will, for example, have a very different sort of relationship with the child if she identifies herself in her own mind as a hater of the child. In that case she is in effect assenting to her rejection of him, and she will consequently not repent but excuse her action to herself, thinking: "I couldn't help it; I hate him for what his father did to me."

So Thalberg is certainly right in holding that, on some occasions, an agent is in fact to be identified with her darker first-order desires, as in the fictional example of the unrepentant Tamar. Notice, however, that, even in such a case, the reason for postulating such an identification is a second-order volition on the agent's part. In identifying herself with her first-order desire to reject the child, Tamar is evidently assenting to that desire and thereby ceasing to assent to the first-order desire to cherish him. But to say she assents to the first-order desire to reject the child is to say that she has a second-order volition to have a first-order volition to reject the child. On the view of the relation between will and intellect sketched above, what has happened is that her reasoning faculty has (at that time) rationalized her rejecting the child and found it acceptable, and she has consequently formed a second-order volition to make her first-order desire to reject the child her will.

In this sense of identification, then, for an agent to identify herself with some part of herself, such as certain of her first-order desires, is for her to form a second-order volition that accepts or assents to that part of herself. On this sense of identification, it is clear that an agent may identify herself with any of her first-order desires, no matter how savage or irrational they may in fact be; and what an agent identifies herself with is clearly up to her and depends on her reason and will. But it is important to see that, contrary to what Thalberg supposes, this conclusion in no way undermines the hierarchical account of the self. For what distinguishes the Tamar who identifies herself with her darker desires from the Tamar who repents them is

not the presence of a first-order desire to reject the child (for that desire is present in both), but rather the presence in the unrepentant Tamar of a second-order volition assenting to those darker first-order desires. This second-order volition the repentant Tamar, who struggles against her desire to reject the child, clearly lacks.²²

But there is also another sense in which an agent can be identified with some part of her character. However erring and faulty it may be, Tamar's reasoning faculty is essential to her, as no other part of her character is. If she were incapable of emotion, or if she were to become apathetic through depression, it would still be possible to consider Tamar a person.²³ But if her reasoning faculty were destroyed, she could no longer be counted a person; certainly she could not enter into any personal relationships, and the ability to do so seems a hallmark of a person.

There is a second sense of identification, then, in which it is correct to say that any agent is always to be identified with her reasoning faculty (whether it functions well or badly). This is not the same as the sense of identification in which the unrepentant Tamar identifies with her first-order desire to reject the child. An agent does not have a reasoning faculty in virtue of some second-order volition assenting to it;²⁴ and, in this second sense of identification, it is not up to an agent to decide what she identifies herself with. Rather, in this sense of identification, an agent is to be identified with what is essential to her as a person, namely, her reasoning faculty. So, on the revised Frankfurt account, we are not simply trading Frankfurt's assumption that agents are to be identified with their second-order volitions for the new assumption that agents are to be identified with their intellects. What we have done instead is show that an agent's reasoning faculty is integral to her existence as a person. If we look to know what *she herself* really wants, we must consider what her reasoning faculty (at a certain time, under a certain description) assents to.

²² The same sort of analysis could be given of an agent who identified herself with her baser first-order desires without the sort of internal struggle I have postulated for Tamar.

²³ Even on the revised Frankfurt account, an apathetic Tamar would count as a person if she had second-order desires, since, without first-order desires, there cannot be in her any discord between first- and second-order desires. And even very apathetic or severely depressed people can have the second-order desire not to be depressed any more.

²⁴ Although it is, of course, open to the agent to form a second-order volition regarding a first-order desire to destroy his reasoning faculty. We can imagine a contemporary version of Dostoevsky's Kirilov, wanting to have a first-order volition to lobotomize himself through drugs or surgery in order to put a stop to the torments his reasoning always brings him.

(And, as the story of Tamar shows, what her reasoning faculty assents to may be something that is, objectively considered, quite irrational.)

Second-order desires represent an agent's reasoning since they stem from the reflection of an agent's intellect on her state of will. Therefore, an agent is to be identified with her second-order desires as much as with her reasoning; her second-order desires represent what her intellect assents to (and so what *she* assents to) among her first-order desires. Consequently, on the revised Frankfurt account, it is correct, contrary to objection (3), to hold that second-order desires represent what an agent really wants. In fact, this second sense of identification is presupposed by the first. We can agree with Thalberg that an agent such as the unrepentant Tamar identifies with her base first-order desires because her identifying with them consists in her forming a second-order volition assenting to them. But the reason why the presence of such a second-order volition suffices for supposing Tamar to be identified with her desire to reject the child is that the second-order volition stems from Tamar's reasoning faculty and, in virtue of that connection, indicates what Tamar herself really wants. So the sorts of considerations Thalberg raises against Frankfurt's account, when they are properly understood, not only do not undermine, but instead actually support, Frankfurt's hierarchical analysis of the self.

V. APPLICATIONS TO ISSUES IN THE PHILOSOPHY OF RELIGION

The revised Frankfurt account of a person and of free will is very fruitful for understanding a variety of religious practices and doctrines, such as the practices of adult baptism, confirmation, asceticism, the doctrine of justification by faith, the nature of heaven and hell, and the point of Romans 7.²⁵ But in the remainder of this paper I shall concentrate on the application of that account to just two Christian doctrines, namely, that God sanctifies some people and that he hardens the hearts of others.

The doctrine of sanctification includes as a central component the claim that God intervenes in the minds of some people in order to make them morally better than they would otherwise be. Many Christian theologians, however, also hold that human beings have free will, where 'free will' is to be understood in an incompatibilist sense. From these two views a number of puzzles arise. First, to be morally good one must freely will some moral good. But then it seems that it is not possible even for an omnipotent God to make anyone morally good, since it is not possible for anyone to cause an

²⁵ I pursue some of these topics in forthcoming papers.

individual freely to will anything (where freedom of will is understood in an incompatibilist sense). Second, even if it were possible for God to make an agent freely will a good he would not have willed otherwise, it does not seem as if his doing so could count as making that agent morally better. If God causes the agent to will some moral good, then we might attribute some moral goodness to God in consequence, but why would we attribute moral goodness to the agent, who is nothing but a puppet of God's will? Finally, if God could in fact make a person morally good, why would he not do so for all persons? How could a good God fail to impart such a benefit to all human beings, so that there would never be any moral evil on earth and no one would ever be brought to hell?

To see the appropriate resolution of these puzzles it is important to understand that, on the Christian doctrine of sanctification, those whom God sanctifies are Christians who are still struggling with moral evil in themselves. Besides holding the traditional Christian beliefs about God, such a person will also believe both that certain things (such as beating one's wife, for example) are wrong and should not be done and that he himself is engaged in some of these morally wrong practices. Consider, for instance, some Christian Patricius who beats his wife Monica.²⁶ On the revised Frankfurt account, we will say that, because Patricius believes it is wrong for him to beat Monica, he forms a second-order desire to make the first-order desire not to beat his wife his will. But when the fit of wrath is on him, he acts on his first-order desire to beat her. When the fit has passed, he laments his action and recognizes that by his own lights he should have acted on his general prohibition to himself not to beat her. Patricius does not have control of himself; he does not have the strength of will to make his first-order desires conform to his second-order desires, and he is not able to make himself have the will he wants to have.

Suppose that Patricius also recognizes that this is his state and prays to God for help. Patricius has reasoned that his beating Monica is an evil but that he is a failure at his efforts to stop it, and that he needs God's help to be the sort of man he himself can approve of. Patricius's prayer for help expresses a second-order volition for God to alter Patricius's first-order will. What *Patricius* wants is for God to change his will in such a way that he no longer wills to beat his wife. If

²⁶ Of course, the real wife-beating Patricius was pagan, according to the account left us by his son Augustine; and no doubt when the patience of his wife Monica had won his conversion to Christianity, he abandoned his practice of wife-beating along with his paganism.

God were so to alter Patricius's will, Patricius's first-order volitions would be in accord with his second-order desires; and, on the revised Frankfurt account, Patricius's will would consequently be free. In giving Patricius a first-order volition not to beat his wife, then, God would not be destroying Patricius's freedom of the will but actually establishing it, since while Patricius's first-order volitions are discordant with his second-order desires, he does not have free will, however free his action of wife-beating may be.

Of course, the strength of second-order desires may vary. As Augustine tells the story, when he prayed to God to give him a will for sexual continence, he made the mental reservation "But not yet." How exactly to characterize Augustine's second-order desire in this case is not certain; but it is clear that, if God had given Augustine a will for sexual continence on that occasion in response to such a prayer, he would have been acting against Augustine's own second-order desires. The result would have been not to evoke or enhance Augustine's free will but to undermine it, because the consequent first-order volition for sexual continence would have been against Augustine's second-order desire to have continence "but not yet." So, in general, sanctification will be a slow process. In response to such half-hearted prayers for help as Augustine's, God can produce some alteration in an individual's first-order will, by strengthening, to the degree warranted by the prayer for help, those first-order desires which are in accord with that individual's second-order desires. But if he is not to destroy freedom of the will, God will not be able to produce a first-order volition unless the second-order desire in the prayer for help is like the whole-hearted turning of the will experienced by Paul on the road to Damascus. Even with a second-order desire for God to alter his will, then, Patricius may find that his struggle against the habit of wife-beating takes some time to win.

On this understanding of sanctification, it is clear that God is not violating an individual's free will in sanctifying him, even when free will is understood in an incompatibilist sense. What God is doing in sanctification is altering an agent's first-order desires to bring them into accord with that agent's own second-order desires, so that, in sanctifying an agent, God is producing or enhancing the agent's freedom of will. Furthermore, it is also clear that the alteration of will God effects in sanctification really does produce moral goodness on the part of the agent. In being sanctified, the agent does not become God's puppet, a simple adjunct to God's will; on the contrary, in sanctifying him, God is helping that agent to have the will the agent himself wants to have. The consequent moral goodness has

its origin in the agent's own volitions, not just in God's, as the objection to the doctrine of sanctification had supposed. Finally, on this view of sanctification, it is clear why it is not possible for God simply to sanctify everyone. The process of God's sanctifying a person consists in God's bringing an agent's first-order desires into line with his second-order desires in response to the agent's second-order volition that God do so. Where the requisite second-order desires and volitions are absent, God cannot alter the first-order desires without undermining or destroying freedom of will.²⁷ And, as the objections to the doctrine of sanctification indicate, it is not possible for an alteration of an agent's will which undermines the will's freedom to result in moral goodness on that agent's part. God cannot make human beings morally better unless they will that he do so.

Someone might think that such an account of sanctification is guilty of Pelagianism. But it does not entail the claim that an agent can achieve sanctification primarily by the exertions of his own will. Instead, this account holds that all the work resulting in moral improvement is done by God in response to the agent's recognition that he needs God's help and his willing to have it. (Of course, a person's willing of God's help is itself a response to God's action in that person's life; but the relation of God's action to such willing is part of the doctrine of justification, and what is at issue here is just the doctrine of sanctification—and not even the whole doctrine of sanctification, since faith, hope, and love are left out of account.) It is true that, on this account of sanctification, an act of free will on the agent's part is necessary for God's work of sanctification, but such a view was also held by Augustine and Aquinas, who are scarcely noted for their adherence to Pelagianism.

As for the doctrine that God sometimes hardens hearts, the classic text is in Exodus 7–14, where God is said to harden Pharaoh's heart so that Pharaoh does not weaken and allow the Israelites to leave Egypt, although the exodus of the Israelites from Egypt is the ultimate object of God's actions in the story, and God punishes Pharaoh for his resistance. That this is not a unique or isolated instance of God's hardening a heart is made explicit in Paul's epistle to the Romans, where Paul generalizes on the story of Pharaoh and concludes that "God hardens whom he will." This doctrine raises two

²⁷ If an individual were a wanton, in Frankfurt's sense, it would be possible for God to alter his first-order volitions without undermining any freedom of his, because such an individual by definition has no second-order desires and so fails to meet one of the conditions for freedom of will. If there are any human beings who are wantons, they will then be exceptions to the claim I make here.

questions concerning God's goodness. First, how could a good God ever intervene in an agent's willing in such a way as to make that agent morally worse? Willing that a human being become morally worse seems the essence of malice, and bringing about a person's moral deterioration is the sort of action typically attributed to Satan. At any rate, it seems utterly incompatible with perfect goodness. Second, even if it were somehow compatible with God's goodness to make an agent morally worse, how could a good God punish that agent for the moral failing God himself has induced in him? On the contrary, since the will that initiated and produced the moral failing is God's, any punishment appropriate for that moral failing seems to be deserved by God, not by the agent manipulated by God. But, on the revised Frankfurt account, there is a way to understand the doctrine that God hardens hearts which satisfactorily answers these questions and which shows this doctrine to be the mirror image of the doctrine of sanctification.

At one point in his diaries, Goebbels pauses to reflect on his own reviewing of German newsreel footage that shows the incredible devastation of Poland wrought by the German armies. Insofar as Goebbels has any morals at all, he seems to hold a primitive sort of divine-command ethics with Hitler fulfilling the role usually assigned to God. So he has no moral compunctions about Poland's devastation; on the contrary, his reason seems to regard it as morally approvable in virtue of having been commanded by Hitler. In addition, Goebbels is determined to further the German war effort in any way he can; and if the war requires inflicting more suffering of the sort visited on Poland, Goebbels is calmly resolute in his willingness to inflict it. In short, Goebbels has first-order desires to wreak the sort of devastation Poland suffered whenever doing so serves Germany's interests, and he also has second-order desires accepting and assenting to those first-order desires.

Nonetheless, as he reviews the newsreel from Poland, those first-order desires begin to slip. It may be that successful efforts at repressing normal human compassion consume a great deal of psychic energy; and since Goebbels's complaint of fatigue is one of the most frequent themes of the diaries, we might suppose that sheer exhaustion is sapping his ability to make his first-order desires conform to his second-order ones. At any rate, Goebbels notices those first-order desires slipping, and he addresses an exhortation to himself: "be hard, my heart, be hard." This exhortation to himself expresses, in effect, a second-order desire. He wants his heart to be hard; that is, he wants to have a first-order will which assents to the sufferings of

the Poles and which is consequently in accord with his second-order desires. Patricius's prayer to God for help expresses a second-order volition that God govern Patricius's first-order will for him, and it shows that Patricius despairs of governing his first-order desires himself. Goebbels's exhortation to himself, on the other hand, indicates an attempt on Goebbels's part to reassert his own control over his first-order desires, and it expresses a second-order desire that Goebbels's first-order desires be in conformity with his second-order desires to make his will serve the German war effort.

Now suppose that because of fatigue or some similar reason Goebbels is unable to reassert control over his first-order desires. Then his first-order desires to inflict suffering on the Poles will weaken, and some first-order desires to relieve their suffering may even appear. In his moment of weakness, he might even (for example) order a shipment of food sent to Poland. A first-order volition to relieve the Poles is clearly morally superior to a first-order volition to cause their suffering. So, if Goebbels weakens in this way, we might suppose that he will in consequence become morally better in some sense. And yet this supposed moral improvement will occur by a sort of accident, and it will certainly be against Goebbels's will, since what Goebbels himself really wants is represented by his second-order desires. In fact, insofar as his first-order desires to ravage the Poles weaken and the opposing first-order desires intensify, Goebbels's free will is undermined, because the new first-order desires are an obstacle to Goebbels's willing what he wants to will and thus constitute a hindrance to Goebbels's freedom of will. Suppose that God were to respond to Goebbels's exhortation to himself to be hard as if it were the atheist's analogue to a prayer, and suppose that God hardened Goebbels's heart for him, supplying the strength of will Goebbels in his fatigue was missing. Such an action on God's part would in no way violate Goebbels's free will. Rather, God would in that case be giving Goebbels what Goebbels himself wants. He would, in effect, be preserving Goebbels's free will from being undermined or destroyed, because in strengthening Goebbels's first-order desires for the ravaging of the Poles, God keeps Goebbels's first-order desires from being discordant with his second-order desires and so helps Goebbels to fulfill one of the requirements for freedom of will.

Finally, in hardening Goebbels, God would, I think, not be making Goebbels morally worse than he would have been without God's intervention. One might suppose that, if the sort of strengthening God grants Patricius makes him morally better, then analogously the sort of strengthening given Goebbels must make him morally worse.

But the two cases are disanalogous in a way which vitiates this inference. Consider Patricius when he fails to act in accordance with his second-order desire assenting to the first-order desire not to beat his wife. In this case as I have constructed it, Patricius is engaged in a struggle within himself; in beating his wife, he acts against what he himself believes to be good. If it is right that an erring conscience binds (and I think it is), then, even if wife-beating were not in itself a moral evil, Patricius would be made morally worse by beating his wife in virtue of putting into action what *he* believes is morally bad. For the case of Goebbels to be analogous in a way that would warrant saying that God's intervention makes Goebbels morally worse than he would have been without God's intervention, Goebbels's state as he acts to help Poland would have to be a mirror image of Patricius's state as he beats his wife: Goebbels would have to believe that what he was doing in helping Poland was a good thing, and he would thus have to be acting in accordance with his moral beliefs. If this were Goebbels's state, it seems clear that hardening his heart would make him morally worse than he would otherwise have been. It would also violate his free will on the revised Frankfurt account. Given both considerations, a good God would never harden anyone's heart in such a case.

But the actual case of Goebbels under consideration here is different. In fact, Goebbels believes that helping Poland is a bad thing; his intellect and his second-order desires are against doing so. Consequently, his state in helping Poland is not the mirror image of Patricius's but rather just the same: in helping Poland he would be putting into practice what he himself believes morally bad. So he does not become morally better by helping Poland in this frame of mind; and, if an erring conscience binds, then there is even a sense in which he becomes morally worse by doing so when he believes it wrong. At any rate, it is not true that, by hardening Goebbels's heart in these circumstances, God makes him morally worse by preventing an action that would contribute to Goebbels's moral improvement. In such a case, and only in such a case, God can harden a person's heart without making him morally worse than he would have been otherwise.

One might suppose that, even so, Goebbels is made worse just by having his evil first-order desires strengthened. If we suppose that Patricius is made morally better in virtue of having his good first-order desires strengthened even when they are not strengthened enough to constitute an effective desire, a volition on which he acts, then the mere fact that Goebbels's evil first-order desires are

strengthened makes him morally worse than he was before. I am willing to concede this claim. The problem with the radical sort of evil in which a moral monster such as Goebbels is sunk is that it blinds and distorts the conscience,²⁸ and even an erring conscience binds. Consequently, Goebbels is made morally worse than he was before whether he acts to afflict the Poles or refrains. If he refrains from destruction of the Poles, he undermines the German war effort in Poland and so betrays his chosen master, Hitler. He thus knowingly does what he believes to be wicked and is consequently made morally worse. On the other hand, since the destruction of Poland is in reality an objective evil, the desire for contributing to that destruction (apart from mitigating circumstances, absent here) is also evil; and the intensification of an evil desire does perhaps imply intensified evil in the desirer. If that is right, then, because Goebbels's conscience is distorted, his moral evil will increase whether God hardens his heart or leaves him alone.

The question, then, is which of the options open to God, hardening Goebbels or leaving him alone, produces the lesser of two evils; and there are several reasons for supposing the answer is hardening Goebbels's heart. As I have been at pains to show, by hardening Goebbels, God preserves a certain freedom of will for Goebbels which would be undermined if Goebbels acted against his moral beliefs. Second, it seems to me arguable that there would be some moral appropriateness in God's fulfilling Goebbels's quasi prayer for hardness of heart. There is something at least morally dubious about a villain such as Goebbels falling into some creditable action simply through fatigue. And, just as it seems right for God to answer Patricius's prayer for help by granting him the strength of will he wants to have, so it seems that Goebbels is getting what he deserves if in answer to his quasi prayer God grants him the strength to persist in the evil he has resolved on. Finally, it may be that in the case of a desperately evil man, such as Goebbels, giving him the strength to have as wicked a will as he wants is hazarding a last shot at reforming him. In hardening him, on my account, God acts only on Goebbels's first-order desires, strengthening them to bring them into conformity with his evil second-order desires; but Goebbels's intellect and second-order desires are in no way cemented in their evil when God hardens Goebbels's heart. In giving Goebbels the first-order desires

²⁸ What makes Goebbels a moral monster is that his intellect and second-order volitions are on the side of a major evil *and* he has no dissenting second-order desires. If he had dissenting second-order desires, then it would not be the case that God enhanced his freedom of will by hardening his heart.

he wishes, God may be providing Goebbels with a mirror, in his character and its consequences, to show him the evil of his wish; and Goebbels's understanding the evil of his second-order desire is the requisite first step to straightening his distorted conscience, to reforming his reason and second-order desires, and thus to beginning a moral rebirth.

In consequence, I think, we have our answers to the two questions raised above concerning the doctrine of hardening the heart. On this understanding of the doctrine, God hardens a heart when he strengthens evil first-order desires so that a second-order will bent on such evil may maintain its control over those first-order desires when fatigue or some other nonmoral accident might have caused the control to weaken. In doing so, God is not violating the agent's free will, and he is also not causing the agent to become morally worse than he would otherwise have been. And if God were to assign blame and punishment to an agent for what that agent did in consequence of God's hardening his heart, God would in no way be unjust to him, for the agent's own will is the source and origin of the evil the agent does. So, when God is said in Exodus to harden Pharaoh's heart, we should understand the text as claiming that God is doing for Pharaoh just what Pharaoh wants and lacks the strength to do for himself, namely, making Pharaoh's first-order volitions correspond to his evil second-order desire. In fact, this interpretation of the hardening of Pharaoh's heart helps to explain an otherwise perplexing feature of the story, namely, that God's hardening Pharaoh's heart alternates with Pharaoh's own hardening of his heart.

Someone might object to this account that it heartlessly neglects the welfare of the Poles, because whatever we might want to say about the merits of allowing Goebbels to fall into some momentary first-order desires to relieve Poland when we are considering only Goebbels's state of character, it is apparently undeniable that such a moment of weakness would have afforded the Poles relief. Therefore, just for the sake of the Poles, a good God would not harden Goebbels's heart. But this objection rests on the mistaken assumption that the fate of the Poles depends solely on the state of Goebbels's will. On the contrary, it is important to see that, if there is a God, there is *no* sort of inviolable connection between the state of Goebbels's will and the welfare of the Poles; and, from God's point of view, the fate of the Poles and the condition of Goebbels's will constitute two entirely separate issues. If Goebbels failed to have a moment of compassion for Poland's sufferings, it was still open to God to aid the Poles in some other way; and if in consequence of

God's hardening Goebbels's heart, Goebbels planned some new suffering for them, it would nonetheless be possible for God to interfere (in any number of ways) to prevent Goebbels from successfully accomplishing what he willed. If God does allow harm to Poland in consequence of some act of will on Goebbels's part, it is because of a separate decision on God's part about the Poles. What sort of decision that might be or how it might be justified are no doubt hard to explain, but no harder than any other instance of the problem of evil; and that problem lies just outside the scope of this paper.

In this way, then, the revised Frankfurt account, which gives a cogent and illuminating analysis of the nature of freedom and the concept of a person, shows divine sanctification and hardening of the heart to be mirror images of each other. In each case, God responds to an agent's desires by giving that agent the first-order volition he wants. When he hardens an agent's heart, he strengthens evil first-order desires in conformity to a second-order desire bent on that evil; and when he sanctifies an agent, he strengthens the first-order desires that the agent's second-order desires want as the good for that agent. In neither case does God's affecting an agent's first-order willing interfere with that agent's free will, on the revised Frankfurt account of freedom of will.

Frankfurt's basic idea of the will as commanding itself seems to me to have great explanatory power in more than one area of philosophy. That it should be particularly fruitful for philosophy of religion is perhaps not surprising given the central role of will in religion.

ELEONORE STUMP

Virginia Polytechnic Institute and State University