https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

AUTONOMY, DURESS, AND COERCION*

By James Stacey Taylor

I. Introduction

For the past three decades philosophical discussions of both personal autonomy and what it is for a person to "identify" with her desires have been dominated by the "hierarchical" analyses of these concepts developed by Gerald Dworkin and Harry Frankfurt.[1] The longevity of these analyses is owed, in part, to the intuitive appeal of their shared claim that the concepts of autonomy and identification are to be analyzed in terms of hierarchies of desires, such that it is a necessary condition for a person to be autonomous with respect to (to identify with) a desire that moves her to act, that she desires that this desire so move her. (Conversely, on these analyses, a person will not be autonomous with respect to a desire that she is moved by, she will not identify with it, if she does not want to be so moved.) Despite the intuitive appeal of these analyses, however, Irving Thalberg has argued that they should be rejected.[2] This is because, he argues, a person who is forced to perform an action by being subjected to duress of a certain degree of harshness will desire to be moved by her desire to submit. Thus, he continues, the proponents of hierarchical analyses of autonomy and identification will be forced to hold that such a person acted willingly, and did not suffer from any impairment in her autonomy. This, Thalberg concludes, is so counterintuitive as to justify rejecting hierarchical analyses.

* I thank R. G. Frey, James W. Child, Fred D. Miller, Jr., Marina Oshana, Harry Frankfurt, Stefaan Cuypers, and Jim Gough for their helpful comments on a very early draft of this essay. Another draft of this essay was presented at the University of Texas at San Antonio Philosophy Symposium in 1999. I thank my commentator, John Hernandez, and my audience on that occasion (especially Michael Almeida) for their constructive criticisms. I especially thank Michael Bratman, Keith Lehrer, Bernard Berofsky, Jonathan Malino, and David Copp for their stimulating comments on this essay. Finally, I thank the editors of *Social Philosophy & Policy* for very helpful suggestions concerning the penultimate draft of this essay.

[1] Harry G. Frankfurt, "Freedom of the Will and the Concept of a Person," in Harry G. Frankfurt, ed., *The Importance of What We Care About* (Cambridge: Cambridge University Press, 1988), 11–25. Gerald Dworkin's early work on his hierarchical analysis of autonomy appears in two main sources: Gerald Dworkin, "Acting Freely," *Nous* 4, no. 4 (1970): 367–83; and Gerald Dworkin, "Autonomy and Behavior Control," in Thomas A. Mappes and Jane S. Zembaty, eds., *Biomedical Ethics* (New York: McGraw–Hill Book Company, 1981), 273–80.

[2] Irving Thalberg, "Hierarchical Analyses of Unfree Action," in John Christman, ed., *The Inner Citadel: Essays on Individual Autonomy* (Oxford: Oxford University Press, 1989), 123–36.

© 2003 Social Philosophy & Policy Foundation. Printed in the USA.

Thalberg's objections to the hierarchical analyses are striking, for they are the only objections that defenders of these analyses have not sought to counter.[3] Instead, it has come to be generally accepted by both defenders and critics of these analyses that Thalberg's objections show that these analyses must either be rejected outright or severely modified.[4] However, that Thalberg's objections are so widely accepted is unfortunate, for (as I will argue in this essay) they are based on a fundamental misunderstanding of Frankfurt's and Dworkin's hierarchical analyses—a misunderstanding that has been consistently overlooked.

Alas, any relief that my argument might provide to the defenders of these hierarchical analyses will be short-lived. Even when these analyses are properly understood, they *still* have counterintuitive implications when they are applied to a situation in which a person acts solely to avoid a threatened harm. However, the arguments in this essay do not lead only to negative conclusions, nor are they only of parochial concern to persons interested in hierarchical analyses of autonomy and identification. Rather, I will argue that, even though hierarchical analyses of autonomy and identification are seriously flawed, they can still provide the key to answering one of the perennial questions in the philosophy of action: Does a person who is forced to perform an action solely in order to avoid a threatened penalty thereby suffer from impaired autonomy?

[3] The three standard objections to hierarchical analyses of autonomy that defenders of hierarchical analyses have sought to counter are the regress *cum* incompleteness objection, the problem of authority, and the objection from manipulation. A concise statement of the regress *cum* incompleteness objection can be found in John Christman, "Introduction," in Christman, ed., *The Inner Citadel*, 8–12. The first outline of the problem of authority can be found in Gary Watson, "Free Agency," *Journal of Philosophy* 72, no. 8 (1975): 205–20. The problem of manipulation is outlined in Michael Slote, "Understanding Free Will," in John Martin Fischer, ed., *Moral Responsibility* (Ithaca, NY: Cornell University Press, 1986), 124–39. A response to the regress *cum* incompleteness objection can be found in Keith Lehrer, "Freedom, Preference, and Autonomy," *Journal of Ethics* 1, no. 1 (1997): 3–25. A response to both the problem of authority and the regress *cum* incompleteness problem is offered in Stefaan E. Cuypers, "Autonomy beyond Voluntarism: In Defense of Hierarchy," *Canadian Journal of Philosophy* 30, no. 2 (2000): 225–56. Finally, difficulties that undermine the objection from manipulation are pressed in Richard Double, "Puppeteers, Hypnotists, and Neurosurgeons," *Philosophical Studies* 56 (1989): 163–73.

[4] Thalberg, for example, holds that these analyses should be rejected outright; see his "Hierarchical Analyses," 135. By contrast, while both Christman and Slote take Thalberg's objections to the hierarchical analyses of autonomy at their face value, these authors hold that, rather than showing that these analyses should be rejected completely, Thalberg's objections only demonstrate that they require some serious modifications. See Christman, "Introduction," 8–9; and Slote, "Understanding Free Will," 127. Similarly, in his most recent account of autonomy, Dworkin acknowledges that his analysis of autonomy requires modification (or, at least, clarification) in order to meet Thalberg's objections. See Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), 18–20.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

II. Hierarchical Analyses of Autonomy and Thalberg's Objections

A. Frankfurt's and Dworkin's hierarchical analyses of autonomy

As I noted above, the core feature shared by Frankfurt's and Dworkin's analyses of autonomy and identification is that these concepts are to be analyzed in terms of hierarchies of desire. (For the sake of clarity, from now on I will take the phrase "is autonomous with respect to her desire *x*" to be synonymous with the phrase "identifies with her desire that *x*,"[5] and the phrase "acts freely" to be synonymous with the phrase "is autonomous with respect to her actions.")[6] More specifically, on Frankfurt's original analysis of identification, a person identifies with a first-order desire that *x* that actually moves him to act if he *volitionally endorses* his desire that *x*. That is to say, a person identifies with (i.e., is autono-

[5] Although it is becoming standard to take Frankfurt's analysis of what it is for a person to identify with her desires to be an analysis of what it is for her to be autonomous with respect to them, to do so still requires some justification, for Frankfurt never explicitly states that his hierarchical analysis of identification is to be understood as an analysis of autonomy. Support for this understanding of his work, however, can be drawn from his "Autonomy, Necessity, and Love," in Harry G. Frankfurt, ed., *Necessity, Volition, and Love* (Cambridge: Cambridge University Press, 1999), 130–38. Here, Frankfurt argues, "The distinction between heteronomy and autonomy coincides . . . with the distinction between being passive and being active" (ibid., 133). In addition to this, in an earlier essay Frankfurt had argued that a person could be said to identify with his desires only if he is "active" with respect to them. (Harry G. Frankfurt, "Identification and Externality," in Frankfurt, ed., *The Importance of What We Care About*, 58–68.) Putting these two claims together, then, it can be seen that it would not be mistaken to hold that, in Frankfurt's view, for a person to "identify with" her desires is *also* for her to be "autonomous with respect to" them. Yet this claim must come with a caveat, since, in his most recent work on identification, Frankfurt argues that both persons and nonpersons can identify with their desires, but the latter will identify with their desires without having volitionally endorsed them after reflecting upon their desirability. (Harry G. Frankfurt, "The Faintest Passion," in Frankfurt, ed., *Necessity, Volition, and Love*, 105–6.) Since this is the case, and since Frankfurt argues that personal autonomy is grounded in the volitional structure of a person's will (Frankfurt, "Autonomy, Necessity, and Love," 132), it would be more accurate to claim that a person is autonomous with respect to those desires that she *reflectively* identifies with (i.e., those that she identifies with after volitionally endorsing them). This point is discussed more fully in James Stacey Taylor, "Book Review of *Necessity, Volition, and Love* by Harry G. Frankfurt," *Philosophical Quarterly* 51, no. 202 (2001): 114–16. Finally, if one remains unconvinced by this line of reasoning and believes that reflectively to identify with one's desires is not identical with being autonomous with respect to them, it should be noted that the arguments in this essay do not depend on one's acceptance of this understanding of the relationship between identification and autonomy. I thank John Martin Fischer for pressing me on this point.

[6] Just as Frankfurt does not use the term "autonomy" in his seminal essay "Freedom of the Will and the Concept of a Person," so, too, does Dworkin eschew this term in his essay "Acting Freely." However, in his later essay "Autonomy and Behavior Control," Dworkin explicitly notes that the authenticity component of the account of autonomy that he is developing herein was "more fully worked out" in "Acting Freely," where it played the same role in his account of acting freely. (Dworkin, "Autonomy and Behavior Control," 280 n. 5.) Given this, it would not be misguided to understand Dworkin's views in "Acting Freely" as being concerned with personal autonomy. However, once again, if one is unconvinced by such an argument, one should note that nothing in this essay rests on these terminological considerations.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

mous with respect to) his effective first-order desire that *x* if he both desires to have the desire that *x* (i.e., he has a *second-order desire* that he have his desire that *x*) *and* he also wants his desire that *x* to move him to act (i.e., he does not merely endorse his desire that *x* with a second-order *desire*, but with a second-order *volition*).[7]

Similarly, on Dworkin's original analysis of autonomy, an "autonomous person is one who does *his own* thing." For Dworkin, it is "the attitude that . . . [the] . . . person takes towards the influences motivating him . . . [that] determines whether or not they are to be considered 'his'."[8] That is to say, on Dworkin's view, for a person to be autonomous with respect to the influences that motivate him to act, it is necessary that he endorse his being moved to act by them. In addition to the requirement that a person's motivations be *authentic* in this way, in his original analysis of autonomy Dworkin also required that a person enjoy both procedural independence and substantive independence with respect to his motivations. Here, a person possesses procedural independence with respect to his motivations if his desire to be moved to act by them has not been produced "by manipulation, deception, the withholding of relevant information, and so on."[9] A person possesses substantive independence with respect to his motivations if he does not "renounce his independence of thought or action" prior to developing them.[10]

On both Frankfurt's and Dworkin's hierarchical analyses, then, a person's autonomy will be impaired if he is moved to act by a desire that he has a second-order desire not to be moved by. In most cases, this is intuitively plausible. For example, if a person is subject to a neurotic compulsion constantly to wash his hands that he desires to be free from, his autonomy will be impaired if he is moved to act by a first-order desire to wash his hands that this neurosis causes him to have, and that he does not wish to be moved by.[11]

B. Thalberg's objections to the hierarchical analyses of autonomy

Thalberg's first objection to these hierarchical analyses is that they have counterintuitive results when they are applied to situations in which a person is forced to perform a certain action by being subjected to duress. To show this, Thalberg focuses on the way in which Frankfurt and Dworkin apply their hierarchical analyses to explain how it is that, when a person is subjected to duress and performs an action that complies with his threatener's desires, it can be said that he acted unwillingly (and thus

[7] Frankfurt, "Freedom of the Will," 14–22.
[8] Dworkin, "Autonomy and Behavior Control," 276.
[9] Ibid.
[10] Ibid.
[11] Although neither Frankfurt nor Dworkin uses this example, it serves well to demonstrate the intuitive plausibility of their hierarchical analyses.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

suffers from impaired autonomy) even though he is performing the action that (given the circumstances that he is in) he most wants to perform.

Dworkin bases his discussion of this problem on an example in which a traveler surrenders his billfold to a highwayman in order to avoid being killed. Dworkin claims that, in this case, "What [the victim] doesn't want to do when faced with the highwayman is to hand money over in these circumstances, for these reasons." More generally, Dworkin asserts that persons "resent acting for certain reasons; they would not choose to be motivated in certain ways. They mind acting simply in order . . . to avoid unpleasant consequences." Dworkin adds that some "part of the human personality . . . takes up an 'attitude' toward the reasons, desires and motives" that determine one's conduct. He further notes that, as a result, "we consider ourselves compelled because we find it painful to act for these reasons."[12]

Thalberg infers from this that, in Dworkin's view, the highwayman's victim suffers from impaired autonomy when he hands over his billfold, because he has adopted a negative attitude toward that desire; he does not desire to be moved by this desire because he does not desire to hand over his billfold solely in order to avoid the unpleasant consequences of resistance. As Thalberg understands it, then, in Dworkin's view a person will suffer from impaired autonomy when he is moved to act by a desire that he does not want to be moved by, where one instance of this would be a case in which the individual concerned does not want to be moved by a desire because he resents acting for the reasons that led him to form the desire. (For example, he performs an action under duress solely in order to avoid the penalty with which he is threatened.) On this understanding of Dworkin's view, a person's autonomy will be impaired with respect to an action he performs, if he is moved to perform that action by a first-order desire whose motivational efficacy he repudiates at the second-order level of his desires.

According to Thalberg, Frankfurt offers a similarly hierarchical analysis of how it is that a person might suffer from impaired autonomy with respect to those actions that he performs as a result of being subject to coercion or duress. Discussing both threats and coercive offers in "Coercion and Moral Responsibility," Frankfurt describes an offer as coercive "when the person is moved into compliance by a desire . . . which he would overcome if he could . . . a desire by which he does not want to be driven." Frankfurt notes that the individual's autonomy may be violated by a threat in the same way: "In submitting to a threat, a person invariably does something that he does not really want to do."[13] As Thalberg

[12] Dworkin, "Acting Freely," 377–78. This quotation of Dworkin's view is from Thalberg, "Hierarchical Analyses," 125.

[13] Harry G. Frankfurt, "Coercion and Moral Responsibility," in Harry G. Frankfurt, ed., *The Importance of What We Care About*, 41–42. This quotation of Frankfurt's view is from Thalberg, "Hierarchical Analyses," 126.

understands this, Frankfurt's view is very similar to Dworkin's. A person will suffer from a lack of autonomy with respect to an action if he is moved to perform that action by a first-order desire that he volitionally repudiates; that is, he is moved to act by a first-order desire that he does not want to be moved by.

Thalberg's first objection to this hierarchical approach to analyzing the concept of autonomy is both elegant and simple. Thalberg argues that, on these hierarchical analyses, a person will only suffer from impaired autonomy with respect to those actions that he is coerced into performing, or that he performs under duress solely in order to avoid the penalty he is threatened with (his "compliant actions"), if he repudiates his first-order desire that actually moved him to act. Thalberg observes, however, that most persons who are subjected to coercion or duress "would, at the time and later, give second-order endorsement to their cautious [and compliant] motives. They are unlikely to yearn, from their elevated tribune, for more defiant ground-floor urges."[14] If this is so, Thalberg notes, then the hierarchical analysts are committed to claiming that most persons who are subject to coercion or duress do *not* thereby suffer from impaired autonomy, for they are not motivated to perform their compliant actions by first-order desires that they repudiate at the second-order level. As Thalberg points out, this conclusion is counterintuitive. However, Thalberg continues, if the hierarchical analysts of autonomy wish to avoid this counterintuitive result, they must claim that, when most persons are subjected to coercion or duress, they *do* wish to repudiate their prudent first-order desires to comply with the demands of their threateners. And this seems to be empirically false. Thus, Thalberg concludes, either the hierarchical analysts of autonomy are committed to the counterintuitive claim that most persons who are subject to coercion or duress are fully autonomous with respect to their compliant actions, or else the hierarchical analysts must attempt to salvage the plausibility of their position by falsely claiming that most persons thus constrained repudiate their effective first-order desires.[15]

In addition to this objection to the hierarchical analyses of autonomy, Thalberg also offers a second: "Dworkin and Frankfurt are mistaken, or

[14] Thalberg, "Hierarchical Analyses," 126.

[15] Thalberg's objection seems to work, not only against Frankfurt's original hierarchical analysis of what it is for a person to be autonomous with respect to these desires, but also against his most recent hierarchical analysis as it is outlined in Harry G. Frankfurt, "The Faintest Passion," 102–6. Here, Frankfurt argues that a person (reflectively) identifies with her first-order desires (i.e., is autonomous with respect to them) if she is satisfied with her volitional endorsement of them, where this satisfaction "is a matter of simply *having no interest* in making changes." ("The Faintest Passion," 104–5. Emphasis in original.) Since it appears that the typical compliant victim of duress would be satisfied, in this sense, with her volitional endorsement of her first-order desire to comply with her threatener's demands, Frankfurt once again appears committed either to the counterintuitive claim that the typical victim of duress is fully autonomous with respect to her compliant actions, or else to the false claim that such a victim will repudiate her effective first-order desires.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

anyway guilty of exaggeration, when they suppose that what a constrained person 'doesn't want' is for some desire or other to move him." [16] Thalberg argues that the primary object of aversion of the victim of coercion or duress is more likely to be the consequences of his compliant actions (for example, that he has lost his billfold to a highwayman) than the fact that he was moved to perform his compliant actions by a certain first-order desire, or that he performed it for certain reasons. Since this is the more likely case, Thalberg argues, Dworkin and Frankfurt misdescribe the situation of the typical victim of coercion or duress to the extent that they hold that he is principally averse to his effective first-order desires, or that he acted for certain reasons, rather than to the actual losses that he has suffered. Thalberg therefore holds that, to be descriptively accurate, any account of why it is that a victim of coercion or duress suffers from an impairment in his autonomy should focus on his aversion to the situation that he finds himself in, and not on any aversion that he might be claimed to have toward those desires that moved him to comply or toward his reasons for such compliance.

III. Responses to Thalberg

Thalberg's objections to Frankfurt's and Dworkin's hierarchical analyses of autonomy are both persuasive and widely accepted. It is thus all the more unfortunate that these objections rest on a fundamental misunderstanding of their views.

Thalberg's first objection to the hierarchical approach to analyzing autonomy is based on the view that proponents of this approach are committed to two related claims. The first of these claims is that a person only suffers from impaired autonomy with respect to one of his actions if he is moved to perform that action by a first-order desire that he repudiates at his second-order level of motivation. The second of these claims is the converse of the first: if a person fully endorses the first-order desire that moves him to perform an action, then he will be fully autonomous with respect to that action. When taken together, these claims comprise what might be termed the *Assumption of the Transitivity of Autonomy*: that autonomy as a property of persons is transitive across their desires and actions.[17] On this assumption, if a person is autonomous with respect to

[16] Thalberg, "Hierarchical Analyses," 127.

[17] Although the Assumption of the Transitivity of Autonomy is never explicitly articulated, it underlies much of the discussion of Frankfurt's and Dworkin's work on coercion and duress. Christman, for example, implicitly accepts the assumption in his "Introduction" to *The Inner Citadel*, as do Slote in "Understanding Free Will"; David Zimmerman in "Making Do: Troubling Stoic Tendencies in an Otherwise Compelling Theory of Autonomy," *Canadian Journal of Philosophy* 30, no. 1 (2000): 25–54; and Susan Dimock in "Personal Autonomy, Freedom of Action, and Coercion," in Samantha Brennan, Tracy Isaacs, and Michael Milde, eds., *A Question of Values: New Canadian Perspectives in Ethics and Political Philosophy* (Atlanta, GA: Rodophi, 1997), 65–86. Moreover, some writers mistakenly understand Frank-

a desire *D* to degree *n*, then he will also be autonomous to degree *n* with respect to an action *A* that he is motivated to perform by desire *D*.

If the proponents of hierarchical analyses of autonomy are committed to this assumption, then Thalberg's first objection to them will hold, and these analyses must either be modified to accommodate it or else be rejected entirely. However, careful examination of these analyses will show that their proponents are *not* committed to this assumption. Indeed, Frankfurt even goes to some pains explicitly to reject it.

A. Autonomy and coercion

Before outlining how Frankfurt rejects the Assumption of the Transitivity of Autonomy, one should first outline his account of how a person who is *coerced* into performing an action suffers from impaired autonomy. This is because Frankfurt's discussion of coercion is often cited to support the view that he (and hierarchical analysts of autonomy in general) is only able to account for a person suffering from impaired autonomy by appealing to discord between different levels of desire in her motivational hierarchy. Since this view appears often (and since the defense of the hierarchical analyses of autonomy against Thalberg's objection will rest on the claim that this view of the hierarchical analyses is false), it would be useful to outline Frankfurt's account of the relationship that holds between coercion and autonomy-impairment in order to contrast it with his more nuanced (and overlooked) discussion of how persons who perform compliant actions under *duress* suffer from impaired autonomy.

For Frankfurt, when a person is coerced into performing an action by being threatened with a penalty if he does not comply, he is moved to act "by a desire which is not only irresistible but which he would overcome if he could."[18] In Frankfurt's view, such a person suffers from impaired autonomy because he is moved to act by a first-order desire that he does not volitionally endorse.

Frankfurt is often criticized for accounting for the autonomy-undermining nature of coercion in this way, for (as is frequently pointed out by his critics) persons who are coerced are not typically moved to comply with their threatener's wishes by "irresistible" desires that they would overcome if they could.[19] Instead, the typical victim of coercion is moved to comply by a desire that he could easily resist if he so desired, but which he acts upon because he fully endorses both his possession of

furt's and Dworkin's accounts of what it is for a person to be autonomous with respect to her desires to be accounts of what it is for her to be autonomous with respect to her actions. For example, see Ian Jennings, "Autonomy and Hierarchical Compatibilism," *South African Journal of Philosophy* 16, no. 2 (1997): 44–50.

[18] Frankfurt, "Coercion and Moral Responsibility," 42.

[19] See, for example, Thalberg, "Hierarchical Analyses," 126; and Zimmerman, "Making Do," 34–38. For an account of coercion that draws from Frankfurt's, see Denis G. Arnold, "Coercion and Moral Responsibility," *American Philosophical Quarterly* 38, no. 1 (2001): 53–67.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

this desire and its moving him to act. Weighing the courses of action open to him, the typical victim of coercion decides to pursue the better part of valor and acts accordingly. This critical focus on Frankfurt's account of the autonomy-undermining nature of coercion is unfortunate, for two reasons. First, in focusing on Frankfurt's explicit discussion of coercion, there is a tendency for his commentators to overlook the fact that he does discuss the more "typical" case of coercion, although he refers to persons who are thus forced to act as victims of "duress," rather than as victims of "coercion." Consequently, Frankfurt is not guilty of taking the untypical for the typical, as David Zimmerman charges—although he is guilty of the lesser charge of using untypical terminology at the expense of the typical.[20] Second (and relatedly), by overlooking Frankfurt's discussion of how a person who acts under duress thereby suffers from impaired autonomy, Frankfurt's commentators are led to misconstrue his subtle and nuanced analysis of autonomous agency. This is because if one focuses on Frankfurt's discussion of coercion and is familiar with his hierarchical analysis of identification, it is easy to conclude that Frankfurt *accepts* the Assumption of the Transitivity of Autonomy. If one focuses solely on the combination of these two aspects of his overall view, it is likewise easy to conclude that Frankfurt holds that a person who is subject to a threat will only suffer from an impairment of his autonomy if he is thereby led to act by an irresistible desire that he wished to resist. However, since Frankfurt actually *rejects* this assumption in his discussion of the autonomy-impairing nature of duress (as I will argue in the following section), any account of his analysis of autonomous agency that focuses solely on his work on coercion and identification will fail to accurately represent his view.

B. Rejecting the Assumption of the Transitivity of Autonomy

Frankfurt's clearest rejection of the Assumption of the Transitivity of Autonomy occurs in his "Coercion and Moral Responsibility"—ironically, the very essay that Thalberg focuses on in developing his criticisms of the hierarchical approach.[21] In discussing a variant of Dworkin's highwayman example, in which a man is forced at gunpoint to hand over his billfold, Frankfurt notes that the person who is forced by the highway-

[20] Zimmerman, "Making Do," 38.

[21] Although the discussion in this section will focus on the work of Harry Frankfurt, since it contains the most explicit rejection of the Assumption of the Transitivity of Autonomy, the same arguments can also be used to defend Dworkin's hierarchical analysis of autonomy against Thalberg's objection. It should be noted, however, that at one point Dworkin does appear to endorse this assumption, when he writes that the victim of coercion (in the non-Frankfurt sense of duress) is motivated similarly to a kleptomaniac who is moved to act by resistible desires that she does not endorse. (Dworkin, "Acting Freely," 378.) However, this implied endorsement of the Assumption of the Transitivity of Autonomy is not central to Dworkin's view and could be repudiated without loss.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

man into handing over his money "would not have defied the threat even if he had been capable of doing so." Instead, "since he really does prefer to give up his money than to die, he would doubtless have fought against any impulse toward defiance which might have arisen in him."[22] Here, then, Frankfurt allows that the motivational structure of the highwayman's victim might have been such that not only did he have a first-order desire to hand over his billfold, but he also possessed a second-order volition that this first-order desire lead him to act. However, despite this harmony in the victim's hierarchy of desires, Frankfurt still holds that, because this person acted to avoid a threatened harm, his autonomy *was* impaired. It is clear that Frankfurt does *not* accept the Assumption of the Transitivity of Autonomy, for if he did accept this assumption, he would be committed to claiming that the autonomy of the highwayman's victim was *not impaired at all*. Since this victim volitionally endorsed his effective first-order desire, he was, according to Frankfurt's hierarchical analysis of autonomy, fully autonomous with respect to it. Thus, if Frankfurt did accept the Assumption of the Transitivity of Autonomy, he would also be committed to claiming that the highwayman's victim was therefore also fully autonomous with respect to his actions—and this is something that Frankfurt explicitly denies.

The fact that Frankfurt allows that a person might volitionally endorse his effective first-order desire to x when he does x under duress, and yet still suffer from an impairment of his autonomy when he does x, shows that Frankfurt does not accept the Assumption of the Transitivity of Autonomy. However, the same observation does *not* show that Thalberg's first objection to the hierarchical approach to analyzing autonomy (as exemplified by Frankfurt's work) is mistaken. Even though Frankfurt might *desire* to reject the Assumption of the Transitivity of Autonomy (and claim to do so), the real question is whether or not he (and hierarchical analysts of autonomy in general) *can* so reject it, given his (and their) previous theoretical commitments.

Fortunately for Frankfurt, he can reject the Assumption of the Transitivity of Autonomy despite his previous theoretical commitments. (This is also true for hierarchical analysts of autonomy in general.) Thalberg was correct to note that, when hierarchical analysts of autonomy attempt to explain why it is that a person who is coerced to act or who acts under duress suffers from impaired autonomy, they focus on the claim that such a person "acts on a motive by which he would prefer not to be moved."[23] However, Thalberg fails to recognize that this central claim is ambiguous. On one reading, it could be understood as the claim that the compliant victim of duress has an *absolute* preference not to be moved by the first-order desire that actually moves him to act (i.e., he repudiates this desire).

[22] Frankfurt, "Coercion and Moral Responsibility," 43.
[23] Ibid., 44. See also Dworkin, "Acting Freely," 377.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

Alternatively, it could be understood as the claim that the victim of duress has a *relative* preference that he not be moved by the desire that actually moves him to act (i.e., he would *prefer* to be moved by some *other* desire, but this preference does not entail that he repudiates the desire that actually moves him to act).

Recognizing this ambiguity is crucial if one is to understand correctly the hierarchical analysts' account of why a person who performs a compliant action under duress suffers from impaired autonomy. As is clear from his objection, Thalberg understands this central claim to attribute an absolute preference to the victim of duress not to be moved by his actual desire to comply; that is, the victim of duress will volitionally *repudiate* his first-order desire to comply. However, as I argued above, this cannot be the correct understanding of what the hierarchical analysts mean by their ambiguous central claim (i.e., that a person who is coerced to act or who acts under duress "acts on a motive by which he would prefer not to be moved"). Frankfurt allows that a person who performs an action under duress does thereby suffer from impaired autonomy *even if* he *endorses* his first-order desire to submit. Thus, the central claim of the hierarchical analysts should instead be understood as noting the fact that a person who acts under duress has a *relative* preference to not be moved by his actual desire to comply. In other words, he does not want to be moved by his desire to submit because (as Thalberg acknowledges) he would prefer to be in a different situation where he would be able to volitionally endorse a first-order desire *different* from that which he *actually* endorses without fear of incurring some harm. (And where, according to both Frankfurt and Dworkin, the typical victim of duress will also prefer to be in a different situation because he resents being in a situation where he has to act "simply in order to avoid unpleasant consequences with no attendant promotion of . . . [his] . . . own interests and welfare.")[24] A person who performs a compliant action under duress, and who thereby "acts on a motive by which he would prefer not to be moved," might volitionally *endorse* his first-order desire to comply, while still preferring that he not be in a situation where such endorsement is prudent. He thus prefers *not* to be moved by this first-order desire to comply, even though he volitionally endorses it, because he prefers to be in a situation such that he could act on a desire that would not move him to perform this compliant action. When their central claim is understood in this way, it is clear that the hierarchical analysts of autonomy are not subject to Thalberg's first objection, for they can accept that a person subject to duress might endorse his effective first-order desire to comply with his threatener's demands, and yet they still recognize that his autonomy was impaired. The hierarchical analysts of autonomy can thus retain their hierarchical

[24] Dworkin, "Acting Freely," 377. Frankfurt explicitly endorses Dworkin's view here in Frankfurt, "Coercion and Moral Responsibility," 44.

analyses of what it is for a person to be autonomous with respect to his desires, while rejecting the Assumption of the Transitivity of Autonomy.

From this discussion it can be seen that, on a generalized version of the hierarchical analysts' view, a person will suffer from impaired autonomy with respect to an action *A* if (i) he is motivated to perform action *A* by a first-order desire that he is not autonomous with respect to (i.e., one that he does not volitionally endorse); or (ii) he is motivated to perform action *A* by a desire that he is autonomous with respect to, but this desire is one that he would prefer not to be moved to act by, because he would prefer to be in a situation other than the one that he is actually in, where this preference for a different situation stems from his belief that, in the situation that he is in, he cannot act to improve his condition, but only to prevent it from becoming worse.

C. Autonomy and actions performed in situations of Type A

Although it is uncontroversial to state that condition (i) can be shared by all hierarchical analyses of autonomy, one might object that condition (ii) cannot similarly be shared by all such analyses, for it cannot be attributed to Frankfurt as part of his considered position. This is because, one might argue, condition (ii) appears to refer to actions that would be performed in a set of situations that comprise a subset of a larger set of situations that Frankfurt later describes (in "Three Concepts of Free Action") as "situations of Type A."[25] A situation of Type A is one in which a person's feeling "that he acted unwillingly derives from the fact that the external circumstances under which he acted were, as he perceived them, discordant with his desires," such that he was "actively discontented" with, or "resistant" to, the state of affairs in which he was forced to act. However, even though a person might be discontented with, or resistant to, the state of affairs in which he acts, Frankfurt holds that a person who acts in a situation of Type A nonetheless remains "autonomous within the limits of an unsatisfactory set of alternatives," and that the actions that he performs in this situation are still ones that he performs freely (i.e., autonomously). Even more worrisome for those who wish to regard Frankfurt as accepting condition (ii), Frankfurt goes on to claim explicitly that since a person can act autonomously in a situation of Type A, "it follows that actions may be performed freely [i.e., a person might be autonomous with respect to them] even when they are performed under duress."[26]

[25] Harry G. Frankfurt, "Three Concepts of Free Action," in Frankfurt, ed., *The Importance of What We Care About*, 47. The set of situations that are referred to in condition (ii) will comprise a subset of situations of Type A, rather then being coextensive with the set of situations of Type A, because not all situations of Type A will be ones in which a person resents having to act in order simply to maintain his condition rather than to improve it.

[26] Ibid., 49, 55.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

Frankfurt's later position (as expressed in "Three Concepts of Free Action") certainly appears to be directly opposed to that which is attributed to him when condition (ii) is ascribed to him by virtue of his being a hierarchical analyst of autonomy. Nevertheless, two possible approaches can be taken to dispel Frankfurt's apparent opposition to condition (ii). The first approach is simply to note that Frankfurt makes it clear that he is not wedded to the claim that a person may act freely even if he acts under duress, for Frankfurt accepts that "it would be reasonable to require that no action be construed as having been performed freely if it was performed under duress, or under duress of a certain degree of harshness."[27] (A requirement that for the sake of clarity can be termed *Requirement 1*.) Consequently, attributing condition (ii) to Frankfurt need not be contrary to his expressed view.

However, to respond in this way to the objection that condition (ii) appears to be directly at odds with Frankfurt's views, as explicitly stated in "Three Concepts of Free Action," has a whiff about it of cheating. Rather than simply adopting this response, then, it would be far more satisfactory to show that the most coherent reading of Frankfurt's view lends support to the claim that he *should* accept Requirement 1, and thus condition (ii). To offer this more satisfactory response, one should first note that Frankfurt's reluctance explicitly to endorse Requirement 1 stems from the fact that he holds two further beliefs concerning the relationship between a person's acting freely (i.e., autonomously) and the question of whether he is morally responsible for his action. First, in addition to acknowledging that Requirement 1 is reasonable, Frankfurt also acknowledges that another requirement (*Requirement 2*) is also reasonable: "that a person can bear no moral responsibility except for what he has done freely."[28] Second, in addition to accepting the reasonableness of both Requirements 1 and 2, Frankfurt believes that a person who performs an action under duress is morally responsible for his performance of that action, even though, owing to the circumstances in which he performed it, he might not be praised or blamed for it.

Frankfurt thus finds himself in a quandary. If he accepts Requirement 1, a person who performs an action *A* under duress (or under duress of a certain degree of harshness) does not act freely; and so, if Frankfurt also accepts Requirement 2, this person is *not* morally responsible for his performance of action *A*. And Frankfurt does not accept that this is so. However, if Frankfurt wishes to maintain that this person *is* morally responsible for his performance of action *A*, by Requirement 2 he has to accept that he does *A* freely. But Frankfurt recognizes that he cannot accept this if he cleaves to Requirement 1.

[27] Ibid., 56.
[28] Ibid.

To show that it is not exegetically inaccurate to hold that Frankfurt accepts condition (ii), one must show that, given his expressed views, Frankfurt should hold Requirement 1. To achieve this, one must outline a way in which Requirements 1 and 2 can be reconciled such that Frankfurt can hold that although a person does not act freely (i.e., autonomously) when he acts under duress of a certain degree of harshness, he is *still* morally responsible for his actions.

Fortunately, such reconciliation can be achieved once it is recognized that Frankfurt does not accept the Assumption of the Transitivity of Autonomy; hence Frankfurt can distinguish between the degree to which a person is autonomous with respect to his desires and the degree to which he is autonomous with respect to his consequent actions. On Frankfurt's hierarchical analysis of what it is for a person to be autonomous with respect to his desires, a person who acts under duress (even under duress of a certain degree of harshness) will be autonomous with respect to the first-order desire that moved him to comply with the threat that confronted him. However, it is clear from his statements in "Coercion and Moral Responsibility," concerning the man who was forced at gunpoint to hand over his money to a highwayman, that Frankfurt does not accept that such a person is thereby fully autonomous with respect to his compliant action. Frankfurt might claim that attributions of moral responsibility should rest on whether the person being so assessed was autonomous with respect to the *desire* that moved him to perform the act that is in question. Thus, rather than adopting a stance of indifference as to which of the reasonable requirements he should accept and which he should reject, Frankfurt could modify Requirement 2 to read: a person must freely desire to perform an action (i.e., must be autonomous with respect to that first-order desire that led him to perform it) for him to be morally responsible for it. In this way, Frankfurt could retain both Requirement 1 and Requirement 2, and also his view that a person who performs an action under duress is morally responsible for it.

Having reconciled Requirements 1 and 2 in this way, one would not be mistaken to attribute condition (ii) to Frankfurt, even though his acceptance of both the (unmodified) Requirement 2, that a person is only morally responsible for what he does freely, and the view that a person may be morally responsible for the actions that he performs under duress, seems to imply that he cannot accept it. Not only is Frankfurt able to avoid Thalberg's first objection, since he is not committed to accepting the Assumption of the Transitivity of Autonomy (despite what is standardly believed), but he is also able to escape from the quandary in which he finds himself at the end of "Three Concepts of Free Action."

D. The response to Thalberg's second objection

In addition to avoiding Thalberg's first objection, the generalized hierarchical approach to analyzing autonomy is also able to avoid Thalberg's

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

second objection, that a person under duress is not principally averse to being moved to act by a certain desire, but to the situation in which he finds himself. On this understanding of the hierarchical approach, its proponents can *agree* with Thalberg that, in cases of duress, the principal object of the victim's aversion is *not* the first-order desire that moves him to act, but, as expressed in condition (ii), the situation in which he finds himself. Of course, this response to Thalberg does not extend to cases in which a person is subject to coercion, for the hierarchical analysts would still claim that the autonomy of a victim of coercion is impaired, since he is moved to act by a desire that he does not volitionally endorse (by condition [i]). Despite appearances, however, such cases of coercion (rather than duress) are not susceptible to Thalberg's second objection. The reason they are not susceptible is simple: They are *defined* in such a way that they only include those cases where the victim's aversion is to his effective first-order desire, and not to his act or to the situation that he is in.

IV. Further Objections to Frankfurt's and Dworkin's Analyses

The generally accepted view that Thalberg's objections to Frankfurt's and Dworkin's hierarchical analyses of personal autonomy are either fatal to such analyses or demonstrate that they require serious modifications is, as I have thus far argued, mistaken. Moreover, once Frankfurt's and Dworkin's views have been properly understood (i.e., it is recognized that they are not committed to the Assumption of the Transitivity of Autonomy), it will be realized that they do not actually offer *hierarchical* analyses of why it is that a victim of duress who performs a compliant action thereby suffers from impaired autonomy. Instead, these apparently hierarchical analyses of what it is for a person to be autonomous with respect to her actions should be recognized as (to use a better term) *Situation-Averse* analyses, since they are based on the claim that the victim's autonomy is impaired owing to the situation that she is in, and not because she suffers from any disharmony in her motivational hierarchy.

A. Impairments of autonomy can differ in degree

Even though Frankfurt's and Dworkin's approaches to analyzing personal autonomy manage to avoid Thalberg's objections, they still suffer from significant theoretical difficulties. The first of these arises from the fact that a person can be more or less autonomous with respect to her actions, where (roughly) the degree to which she is autonomous with respect to them correlates to the degree to which it is she, and not some-

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

one else, who controls her performance of them.[29] It thereby follows that a person will suffer impaired autonomy with respect to her actions in proportion to the degree that she cedes control over them to someone else. This being so, a person who is threatened with a severe penalty unless she performs "an action within the week whose performance she disvalues" will suffer *less* impairment of her autonomy with respect to her consequent compliant action than will another victim of duress (such as the victim of Dworkin's highwayman) whose compliant action is specified in greater detail by his threatener. This is because the former victim of duress is able to exercise a greater degree of control over both *what* action she is to perform to avoid incurring the threatened penalty and *when* she is to perform it, than is the latter victim, who lacks control over both the nature ("Hand over your billfold!") and the timing ("Now!") of his compliant action. However, the Situation-Averse analyses outlined above aim only to explain how a person might suffer from impaired autonomy, *simpliciter*, when she performs compliant actions as a result of being subjected to duress. Such analyses cannot readily account for the fact that victims of duress might suffer from differing degrees of autonomy-impairment with respect to their compliant actions. Consequently, these analyses fail to offer a complete account of the relationship that holds between a person's being subject to duress and the consequent impairment of her autonomy when she performs the compliant actions that are required of her.

It might be tempting to respond to this first theoretical difficulty by modifying the Situation-Averse analyses, such that the autonomy of a victim of duress will be impaired in proportion to the degree that she resents acting solely to prevent her situation from worsening, rather than acting to improve it. This temptation should be resisted, however, for it commits the proponents of Situation-Averse analyses to claiming that, if the person in the above example whose compliant action was left unspecified by her threatener resents being subject to duress to a higher degree than does the victim of Dworkin's highwayman, then it will be she, and not the highway-robbery victim, who will suffer from the greater impairment in autonomy. Yet to claim this is to claim that the person who retains a greater degree of control over her actions suffers from a greater impairment of her autonomy with respect to those actions than the person who retains less control over his actions—and this is obviously mistaken.

B. The Anti-Stoic Objection

The difficulty that is posed by this suggested modification to the Situation-Averse analyses leads to the second major objection that such

[29] For an informative discussion of this point, see James W. Child, "Specific Commands, General Rules, and Degrees of Autonomy," *Canadian Journal of Law and Jurisprudence* 8, no. 2 (1995): 245–58.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

analyses face: they have troubling Stoic implications. On these Situation-Averse analyses, a victim of duress will suffer from impaired autonomy with respect to her compliant actions if and only if she resents that she is performing these actions only to prevent her situation from getting worse, rather than to improve it. Thus, like the Stoic view of freedom, in which a person can become free by altering her expectations in the face of externally imposed constraints, so that these constraints are no longer felt as such, according to these Situation-Averse analyses, a victim of duress will not suffer from impaired autonomy if she *fails* to resent that she is moved to act only to prevent her situation from getting worse, rather than to improve it.

This view is mistaken, though. To see just how mistaken it is, consider the cases of Vinnie and Vera, who have both been held up at gunpoint in New York City and forced to hand over their billfolds. When confronted by the gunman, both Vinnie and Vera quickly decide that it would be prudent to comply with his demands, and they rapidly, volitionally endorse their first-order desires to comply with his threat. Vera, however, is outraged by what she takes to be a personal affront to the sanctity of her person—"Such a thing would never have happened to her in her hometown in Minnesota!"—and so deeply resents having to hand over her billfold in order to prevent her situation from (rapidly and bloodily) getting worse. As one would expect, then, the Situation-Averse analyses hold that Vera's autonomy is impaired with respect to her compliant action. By contrast, Vinnie, a native of the Bronx, simply regards the whole episode as part of the cost of living in New York City. (Indeed, rather than resenting his having to act for the reason that he did, Vinnie admires the gunman's *chutzpah* in holding him up in Macy's on a Saturday morning.) In this case, the Situation-Averse analyses hold that Vinnie does *not* suffer from any impairment of his autonomy. This cannot be correct, however, for to avoid being shot, Vinnie cedes to the gunman the *same* degree of control over his actions as did Vera. Since both Vinnie and Vera lack the same degree of self-direction with regard to their compliant actions, their autonomy with respect to their respective compliant actions must (*ceteris paribus*) be impaired to the same degree.

This objection to the Situation-Averse analyses can be termed the *Anti-Stoic Objection*, since it is based on denying the Stoic view of freedom as it is outlined above. This Anti-Stoic Objection can be generalized. Consider, for example, the slave who believes that it is his God-given lot in life to work solely to avoid being beaten, the political prisoner who fails to resent his incarceration because he believes that he will not have the courage openly to avow his political views outside the Gulag, and the downtrodden wife who has resigned herself to expect nothing better in life than constantly being told what to do by her husband. All three of these people suffer from impaired autonomy, even though each fails to resent being moved just to prevent his or her situation from becoming

worse.[30] This because it is not the slave, the political prisoner, or the wife, but someone else, who controls which actions they perform and when. From considerations such as this, it appears that the Situation-Averse analyses have confused the degree to which a person suffers from impairment in her autonomy with the degree to which she suffers from a diminution in her well-being. The contented slave, the political prisoner, and the downtrodden wife might all be unaffected by the autonomy-undermining situations in which they find themselves—and this is why they fail to resent them. However, the fact that these persons fail to resent being moved to act for the reasons that they are moved to act does not show that their autonomy with respect to their actions is thereby unimpaired. Instead, it merely shows that these situations do not adversely affect the well-being of these individuals as much as one would expect. The Stoic retreat into the "inner citadel" of one's reason, soul, or noumenal self might protect one's well-being, but it still leaves one's autonomy undefended.[31]

V. The Analyses of Christman and Slote

Although Frankfurt's and Dworkin's Situation-Averse analyses of why a person who acts under duress thereby suffers from impaired autonomy are unsatisfactory, both John Christman and Michael Slote have held that the general approach that Dworkin and Frankfurt take to this problem is sufficiently promising to warrant further exploration.

A. Christman's analysis

In response to Thalberg's objections, Christman argues that a proponent of the hierarchical approach to analyzing personal autonomy need not evaluate the compliant action of a victim of duress as being one that he was fully autonomous with respect to, as Thalberg claimed.[32] Unfor-

[30] The example of a (real-life) Stoic slave is discussed in Sigurdur Krisstonson, "The Limits of Neutrality: Toward a Weakly Substantive Account of Autonomy," *Canadian Journal of Philosophy* 30, no. 2 (2000): 257–86. The example of the political prisoners who fail to resent their incarceration is discussed in Flint Schier, "The Kantian Gulag: Autonomy and the Liberal Conception of Freedom," in Dudley Knowles and John Skorupski, eds., *Virtue and Taste: Essays on Politics, Ethics, and Aesthetics* (Oxford: Blackwell, 1993), 1–17.

[31] Concerns about such a retreat were famously raised by Isaiah Berlin in *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), 135–36. The worry that such Stoicism conflates the preservation of well-being with the preservation of one's ability to act freely (i.e., autonomously) has been expressed by Wright Neely in "Freedom and Desire," *Philosophical Review* 83, no. 1 (1974): 38. A similar objection has also been offered by David Zimmerman, "Making Do," 25–30.

[32] Christman, "Introduction," 8–9. Christman's failure to recognize Thalberg's misunderstanding of Frankfurt's and Dworkin's analyses indicates that he similarly failed to understand them properly, for, had he done so, he would have recognized that they were not susceptible to Thalberg's objections. Christman's modification of these analyses is thus

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

tunately, however, Christman did not reach this (correct) conclusion after recognizing that Thalberg had misunderstood the hierarchical analyses, for Christman accepted the conventional view that Thalberg's first objection to these analyses showed that they were in need of modification. Christman's response to Thalberg is based on an example of a bank teller who hands over his bank's money to a robber, solely in order to avoid being shot. Christman argues that the hierarchical analyses of autonomy could be salvaged if one accepts that, in such a case, "there are actually two operative desires that the agent has, both with the same object: giving over the money. The first is the desire to not be heroic in threatening situations. . . . [T]he other relevant desire is to give over the money *in this particular case*."[33] Christman argues that these two desires "are importantly different because they have different origins: they were adopted for differing reasons and under different circumstances."[34] The bank teller's *first* desire, Christman argues, is one that he is autonomous with respect to, for he approves of the way in which he came to possess it.[35] According to Christman, however, the bank teller's *second* desire is not one that he is autonomous with respect to, for he does not approve of the way in which he came to possess this desire. He would not have formed this desire had he not been placed under duress. From this example, Christman concludes that "if one's account of autonomy requires that all the relevant desires causing an action be autonomously formed, then . . . [the bank teller] . . . is not autonomous" when he compliantly hands over the money to the robber.[36]

Although Christman's historical analysis of autonomy leads him to conclude correctly that persons who perform compliant actions under duress thereby suffer from impaired autonomy, his analysis suffers from several flaws that prevent it from being a fully satisfactory account of why this is so. The first two objections to Christman's historical analysis of why the compliant victim of duress suffers from impaired autonomy question the psychological plausibility of the motivational structure that Christman attributes to him. First, and most obviously, it is questionable whether the bank teller who complies with the robber's threat really possesses *two* operative first-order desires that share the object of handing over the money, as Christman claims. Instead, it is more likely that the

redundant, insofar as it is considered a defense of them against Thalberg's objections. This is not also to say, however, that Christman's modification lacks considerable merit in its own right as an attempt to explain how it is that the typical compliant victim of duress suffers from impaired autonomy, owing to his subjection.

[33] Ibid., 8. Emphasis in original.

[34] Ibid.

[35] For Christman's innovative historical account of what it is for a person to be autonomous with respect to his desires, see John Christman, "Autonomy: A Defense of the Split-Level Self," *Southern Journal of Philosophy* 21, no. 3 (1987): 281–93; and John Christman, "Autonomy and Personal History," *Canadian Journal of Philosophy* 21, no. 1 (1991): 1–24.

[36] Christman, "Introduction," 9.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

teller has only *one* operative first-order desire to hand over the money, and it is this desire that moves him to act. Indeed, it seems that the only reason one might have to posit the existence of a *second* first-order desire is if one is a hierarchical analyst of autonomy who wishes to explain how a compliant victim of duress suffers from impaired autonomy, and who accepts Thalberg's (mistaken) view that the only way that one can explain this (while at the same time retaining the hierarchical approach to autonomy) is to posit that such a victim suffers from motivational discord across his different levels of desire. This point leads to the second objection to Christman's account. As Thalberg noted, the most likely object of the teller's aversion is *not* his operative first-order desire (or that operative first-order desire whose conditions of formation he does not approve of) but to the *situation* that he is in.[37] And if this is so, then any account of why a person who performs a compliant action under duress thereby suffers from impaired autonomy that is based on this person's aversion to one or another of his desires will rest on an inaccurate description of such a victim's actual motivational state.

In addition to these two objections, which apply specifically to Christman's historical analysis, Christman's approach is also subject to the objections that beset Frankfurt's and Dworkin's (properly understood) Situation-Averse analyses. Like these analyses, Christman's lacks the theoretical resources to explain why it is that a person who is afforded considerable latitude with respect to both *what* compliant action she will perform and *when* she will perform it, suffers from *less* impairment of her autonomy than a person whose compliant action is specified in detail by her threatener. Furthermore, Christman's analysis also has the same troubling Stoic implications that beset the earlier Situation-Averse analyses of Frankfurt and Dworkin. This is because, on Christman's analysis, a person's autonomy with respect to her compliant action will be unimpaired if two conditions are met: (1) she does *not* object to one of her operative first-order desires being formed by her being placed under duress, and (2) she also has no objection as to how she came to have her other, previously formed, operative first-order desire with the same object. If this is the case, then persons who are Stoically resigned to their lot (such as the slave, the wife, and the political prisoner in the examples that I mentioned above) will be held to suffer from no impairment in their autonomy, even if they are constantly forced to perform actions under duress. And, as was noted above, this is highly counterintuitive.

B. Slote's analysis

Christman's historical analysis, then, is no more persuasive than the Situation-Averse analyses that he intends it to replace—indeed, given its

[37] Thalberg, "Hierarchical Analyses," 127.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

susceptibility to the additional objections that I outlined in the preceding section, it appears to fare worse. Alas, Slote's development of Frankfurt's and Dworkin's analyses fares little better. Analyzing why it is that a person who acts in compliance to a threat thereby suffers from an impairment of his autonomy, Slote begins by focusing on the claim that such a person will *resent* being placed under duress. Slote argues that the resentment that a person is likely to feel when he is placed under duress signifies a desire or a wish to defy the person who is placing him under duress, or to retaliate against that person. This being so, Slote continues, the victim of duress "will have various momentary/fragmentary fantasies of heroic defiance, or foiling, or retaliation, with appropriate accompanying thoughts."[38] Consequently, Slote argues, the victim will be *ambivalent* about his own compliance, and to the extent that he resents having to act in this way and wishes that he had chosen to defy his threatener, he will reproach himself for choosing to save his own skin. In complying with the threat, then, "such a person acts from a desire by which he wishes not to be moved to act, even given the alternatives that he confronts. There is in him a conflict between a first-order desire to comply (and play it safe) and an ultimately frustrated second-order volition that that desire not be effective and that . . . he should overcome that safe-playing desire through appropriate heroics or defiance."[39] Slote concludes that there is no need to reformulate Frankfurt's views to respond to the possibility that Frankfurt is committed to the mistaken claim that victims of duress will be fully autonomous with respect to their compliant actions, since, owing to the motivational ambivalence that they experience, such persons will *not* count as being autonomous with respect to their effective first-order desires to comply.

The first objection to Slote's analysis comes from David Zimmerman, who has noted that it rests on the claim that, if a person resents being moved to perform a compliant action, then he will automatically form a second-order volition that he not be moved by his first-order desire to comply. However, Zimmerman argues, it is not at all clear that this will happen, because the mere fact that a person resents a certain state of affairs need not lead to him judging that he should resist it.[40] Thus, even though the typical compliant victim of duress might resent being forced to comply with the demands of his threatener, he is not necessarily subject to any motivational ambivalence, for his resentment need not automatically lead him to form a second-order volition repudiating his effective first-order desire to comply. If Slote's analysis is correct, then there might be some compliant (and resentful) victims of duress whose autonomy remains unimpaired, despite their subjection.

[38] Slote, "Understanding Free Will," 127.
[39] Ibid.
[40] Zimmerman, "Making Do," 47.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

Yet, even if one accepts, for the sake of argument, that Slote is correct to claim that a person's feelings of resentment will lead to his volitional repudiation of his effective first-order desire to comply, Slote's focus on a person's motivational ambivalence as the source of impaired autonomy has the counterintuitive implication that even a person who exercises significant control over his life does not act autonomously, *unless* he does not suffer from any serious doubts about the advisability of his actions. To see this, consider an example that was developed by Paul Benson, which I am using here to different purpose: Charlotte, a "white, middle-class American woman," comes to adulthood in the 1880s, and "finds the rigid, conventional roles that women of her class are expected to take up—wife, mother and housekeeper—to be nearly intolerable."[41] Charlotte thoroughly detests the conventional role that she is expected to play, and leaves her husband and children in order to pursue an artistic career—an occupation that, owing to her socialization, she agrees is "an unsuitable job for a woman" (if I may borrow that description here, as has another philosopher, from novelist P. D. James).[42] However, Charlotte does not leave her family as a result of having undergone a feminist version of St. Paul's conversion on the road to Damascus. Instead, owing to her internalization of her era's norms of femininity, Charlotte's decision to abandon her family and pursue a career is agonizingly difficult for her, and even after she has made this decision, she continues to experience "various momentary/fragmentary fantasies" of returning to her family and living as a dutiful (if frustrated) Victorian wife and mother. Moreover, these "momentary/fragmentary fantasies" of return that Charlotte experiences are not merely idle daydreams, for they lead her to feel so guilty about leaving her family that she forms a second-order volition that repudiates her effective first-order desire to pursue her career. Nevertheless, despite the guilt that she experiences, Charlotte still, all things considered, believes that she should continue to pursue her artistic yearnings, and so her desire to do so still moves her to act.

Charlotte thus experiences the same sort of motivational ambivalence that Slote's compliant victim of duress experiences, for they both experience a motivational conflict at the level of their second-order volitions. However, rather than claiming that, as a result of this motivational am-

[41] Paul Benson, "Feeling Crazy: Self-Worth and the Social Character of Responsibility," in Catriona Mackenzie and Natalie Stoljar, eds., *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (Oxford: Oxford University Press, 2000), 74. It should be noted that Benson developed the example of Charlotte in order to illustrate his views concerning the relationship between a person's self-worth and her moral responsibility for her actions, rather than to illuminate the conditions that are required for a person to be autonomous with respect to her actions. The following discussion in the text of this example might thus not be pertinent to Benson's own use of it.

[42] A similar example is given by Marina Oshana, who also recognizes the appropriateness of the title of P. D. James's novel, *An Unsuitable Job for a Woman*, in Oshana's discussion of her own example. Marina Oshana, "The Autonomy Bogeyman," *Journal of Value Inquiry* 35 (2001): 220.

bivalence, Charlotte is *not* autonomous with respect to her acts of leaving her family and pursuing her career, it is far more plausible to claim that she is *highly* autonomous with respect to these acts, for they demonstrate that she is directing the course of her own life despite considerable social pressure not to do so. But if this is so, it cannot be the case that a person's experience of motivational ambivalence at the level of her second-order volitions is sufficient for her to be held to suffer from impaired autonomy.

Of course, one might rightly object that this criticism of Slote's analysis trades heavily on the intuition that Charlotte *is* acting autonomously when she leaves her husband to pursue a life of art—and that simply *claiming* that this is so is not the same as *showing* that it is so. In order to make the above objection to Slote's analysis compelling, then, one must fill out the details of Charlotte's motivational structure. For this objection to hold, it must be plausible to attribute to Charlotte a motivational structure that is structurally identical to that which Slote holds is possessed by a compliant victim of duress. Moreover, it must be the case that, when the details of Charlotte's motivational structure are filled out, it is clear that there is a theoretical basis (rather than a merely intuitive one) for considering her to be autonomous with respect to her pursuit of her artistic career.

As I argued above, it is plausible to hold that, like Slote's compliant victim of duress, Charlotte suffers from conflicting second-order volitions. Now, as Thalberg noted, most compliant victims of duress, all things considered, *endorse* their effective first-order desires to comply. To develop Thalberg's insight further, one can attribute to such persons the possession of a *third*-order volition endorsing their pro-compliance, second-order volitions (i.e., all things considered, they prefer to be moved to comply rather than to resist). Since this is so, for Charlotte's motivational structure to be structurally identical to that of the compliant victim of duress, it must be plausible to ascribe to Charlotte a third-order volition endorsing her effective first-order desire to pursue her career. And this is indeed the case, for it is plausible that she might, all things considered, endorse her desire to pursue an artistic career and leave her family. Thus, since one can plausibly ascribe structurally identical motivational patterns to both Charlotte and the compliant victim of duress, and since the details of these motivational patterns (having been filled out) provide a theoretical basis for holding Charlotte to be autonomous with respect to her pursuit of her career, Slote is faced with a dilemma. He can either continue to claim that second-order volitional ambivalence suffices for a person to suffer from impaired autonomy, or he can accept that a person's being subject to such volitional ambivalence is not sufficient for her to suffer from impaired autonomy. If Slote chooses the first horn of this dilemma, he will be committed to the claim that Charlotte's autonomy is impaired when she defies social custom and pursues her own goals. This claim is highly counterintuitive. However, if he chooses the second horn

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

of this dilemma, then his account of why it is that a compliant victim of duress suffers from impaired autonomy is incomplete.[43]

VI. Why the Compliant Victim of Duress Suffers from Impaired Autonomy

From this discussion it is clear that there is something very baffling about the status of actions that a person performs solely in order to avoid incurring a threatened penalty. A person who is subjected to duress *chooses* to perform her compliant actions after deciding that her performance of them offers the least unattractive option from a set of unpalatable alternatives with which she is faced. Since she thus desires to perform these actions, and this desire moves her to perform them, it seems, prima facie, plausible to claim that she is fully self-directed, fully autonomous, with respect to their performance. However, to claim that a person who is *forced* to perform a series of compliant actions by being subjected to duress is a paradigm of someone who is engaged in autonomous self-direction seems clearly mistaken. The most obvious way of resolving this difficulty is, of course, to deny one or the other of the intuitions that these claims express. However, it would be far more satisfying to provide an account of personal autonomy that is able to respond to both of these tugs of intuition, when it is applied to a situation in which a person acts solely in order to comply with a threat.

A. An informative false start

Given that one can reject the Assumption of the Transitivity of Autonomy, the clearest way to provide an account of personal autonomy that is able to respond to both of these tugs of intuition is to follow the lead of Frankfurt and Dworkin, and to develop an account in which it is a necessary but not a sufficient condition for a person to be autonomous with respect to her actions, that she be autonomous with respect to her effective first-order desires. One could thus argue that, although the compliant victim of duress might be autonomous with respect to the effective first-order desire that moves her to perform her compliant action, she is still *not* autonomous with respect to the act itself. Since this is so, it is tempting to argue that a person who performs a compliant action under duress is moved to act by the effective first-order desire that *w*: "to do whatever my threatener tells me I must do to avoid incurring the penalty I am threatened with." Given her situation, this person might be autonomous with respect to this desire. (She might volitionally endorse it and be satisfied with it; it might be authentically hers, and she might be both substan-

[43] Worse yet, Zimmerman argues that Slote's position is incoherent. Zimmerman, "Making Do," 48–50.

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

tively and procedurally independent with respect to it; she might decide to treat it as reason-giving; and so forth.)[44] However, to satisfy her desire that *w*, she will cede control over her actions to her threatener, and thus suffer from impaired autonomy with respect to these actions.

This *Simple Analysis*, as it may be called, certainly satisfies both of the intuitions that are expressed in the aforementioned claims. The intuition that the compliant victim of duress retains full autonomy even under duress because she *chose* to act as she did, is satisfied, since she *is* autonomous with respect to her effective first-order desire to comply. Similarly, the intuition that she suffers from impaired autonomy as a result of being *forced* to act as she did is also satisfied, since to satisfy her desire that *w*, she ceded control over her actions to her threatener. Nevertheless, the temptation to account in this way for the impaired autonomy of a compliant victim of duress should be resisted, for several reasons.

First, it is not clear that the actual first-order desire that moves a compliant victim of duress to perform her compliant action is as general and open-ended as the above desire that *w*. For example, when a highwayman's victim hands over his billfold in order to avoid being shot, it is plausible to assume that at least part of the object of his effective first-order desire concerns the specific action that he is being forced to perform. But this is overlooked by the Simple Analysis, which holds that the object of this victim's effective first-order desire is that of the desire that *w*.

From this first objection to the Simple Analysis, one can also develop a second: the *Objection from Illegitimate Redescription*. The Simple Analysis appears to rest on the presupposition that it is legitimate to redescribe the intentional object of a person's effective first-order desire. (For example, the intentional object of the effective first-order desire of the victim of the highwayman should not be described as "handing over my billfold to avoid being shot," but as "performing an action that will preserve my well-being.") It is doubtful that such redescription is legitimate.[45] However, even if such redescription were legitimate, it would be self-defeating to aim to satisfy *both* of the intuitions expressed earlier by developing an analysis, which rests on this basis, of why it is that the compliant victim of duress suffers from impaired autonomy. This is because, if it is legitimate to redescribe the objects of a person's effective first-order desires in this way, then there seems to be no principled objection to allowing her *actions* to be similarly redescribed as well. If this were the case, then the compliant actions of any such victim of duress might be described as

[44] These accounts of what it is for a person to be autonomous with respect to her effective first-order desires are, respectively, those of Frankfurt in "Freedom of the Will" and "The Faintest Passion," Dworkin in "Autonomy and Behavior Control," and Michael E. Bratman in "Identification, Decision, and Treating as a Reason," *Philosophical Topics* 24, no. 2 (1996): 1–18.

[45] Dworkin expresses doubts about such redescription in "Acting Freely," 372–75.

those "that secure her well-being." With this redescription of the compliant victim's actions in hand, it is a simple matter, on this analysis, to move to the claim that the victim is autonomous with respect to *both* her effective first-order desires and her consequent actions. This is because, under this revised description of her actions, no mention is made of the victim's ceding any control over them to another. Instead, she merely desires to act so as to secure her well-being and she does so. But to make this claim is inconsistent with the Simple Analysis, for in making it one denies that the compliant victim's autonomy is impaired. The Simple Analysis, then, faces two immediate and related difficulties. First, to defend it one will have to establish the legitimacy of redescribing the victim's effective first-order desires. Second, even if the legitimacy of such redescription *can* be established, it seems that there will be no principled objection to holding that the compliant victim of duress is autonomous with respect to both her effective first-order desires and her actions.

In order for one to defend the Simple Analysis further, one might do so by noting that these two objections to it are based on the understanding that the Simple Analysis is committed to the claim that the typical compliant victim of duress is moved to act by the first-order desire that *w*—and this is not so. Instead, it is consistent with the Simple Analysis to hold that the compliant victim of duress forms her desire that *w prior* to her forming her effective (and more specific) first-order desire that *p* ("to hand over her billfold"), such that she forms her desire that *p* to satisfy her desire that *w*. Unfortunately, this revised version of the Simple Analysis will not suffice either, for if this revised description of the motivational structure of the compliant victim of duress were accurate, she would cede control not only of her actions to her threatener, but also of her effective first-order desires. This is because it appears that, after the compliant victim of duress has formed the desire that *w*, it will be up to her threatener what actions she will perform, and thus it will *also* be up to him what specific first-order desires she will need to form, in order to be appropriately moved to act to satisfy her prior desire that *w*. And if this is the case, then the compliant victim of duress will *not* retain her autonomy with respect to her effective first-order desires that *p*. But if this is so, then the Simple Analysis will no longer satisfy both of the intuitions expressed above, for if one accepts this revised version of the Simple Analysis, one will regard the victim of duress as failing to exercise any autonomy at all, after she has acted to satisfy her desire that *w*.

As an attempt to account for both of the intuitions expressed above, the Simple Analysis is thus caught in a dilemma. In its original form, this analysis must establish that it is theoretically legitimate to redescribe the effective first-order desires of the victims of duress who comply with the demands of their threateners. However, if the legitimacy of such redescription is established, it will lead to the development of an alternative analysis of how to characterize the autonomy of a compliant victim of duress—

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

one that does not satisfy the second intuition (i.e., that the victim of duress suffers from impaired autonomy through being forced to act as she did). Yet, if the Simple Analysis is revised to avoid this problem in the way outlined here, it will no longer satisfy the first of the two intuitions (i.e., that the compliant victim of duress retains full autonomy even under duress because she *chose* to act as she did).

B. Learning from the Simple Analysis

Despite these difficulties faced by the Simple Analysis, there is still something intuitively compelling about its core idea that the effective first-order desire that moves the compliant victim of duress to act is one whose satisfaction involves her ceding control over her actions to another. This core idea, then, might be retained as the basis for an alternative analysis that will satisfy both of the intuitions expressed in the preceding subsection. Furthermore, one can also draw from the previously discussed objections to the Simple Analysis an idea of what counterintuitive traps should be avoided by a satisfactory account of why a compliant victim of duress suffers from impaired autonomy.

The first set of objections to the Simple Analysis showed that one should not assume that one can intentionally describe the object of a compliant victim's effective first-order desire without making reference to the particular act that it moves her to perform. In addition to this, the second set of objections to the Simple Analysis showed that, to satisfy both of the intuitions discussed above, one cannot posit that the compliant victim of duress abdicated complete control over her compliant actions to her threatener. This second point is reinforced by the fact that persons who are subject to duress do not simply perform *any* act that their threatener requires of them in order to avoid incurring the threatened penalty. Instead, persons subject to duress will assess whether or not they are willing to perform the particular act required of them in order to avoid the threatened penalty, where the severity of the penalty might be sufficient to move them to perform some actions, but not others. From these two objections to the Simple Analysis, it can be seen that a satisfactory analysis of how it is that a compliant victim of duress suffers from impaired autonomy will have to recognize *both* (1) that the object of the victim's effective first-order desire should include a reference to the specific act that she is required to perform, *and* (2) that the victim will herself decide whether or not to perform the particular act that her threatener requires of her.

It is easy to construe this second condition as requiring that a satisfactory analysis must be one that regards the compliant victim of duress as a person who assesses whether or not she will choose to perform the act her threatener requires of her. However, this is precisely how *not* to understand this second condition, for to view the victim's deliberative pro-

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

cess in this way is to misunderstand her situation. This is because when a person under duress deliberates about whether or not to perform the action that her threatener requires of her, she is not deliberating as to whether or not she should *perform* this action, but whether or not she should *resist* being forced to do it. This distinction between the possible deliberative aims that might be ascribed to a person who is subject to duress is not without a difference, for it points to the different default positions that are held by a person who is not subject to duress and by one who is. For the former, the default conclusion of her deliberative process is (typically) that she *will not* perform the action she is considering (i.e., she will not perform it unless she believes that she has a reason to do so). For the latter, the default position is (typically) that she *will* perform the action that she is considering (i.e., she will perform it unless she believes that she has a reason to resist doing so). Thus, if a person who is subjected to duress decides that she should perform the act that is required of her by her threatener, in order to avoid the penalty that he is threatening her with, then she will not be moved to act by the desire to perform this action *simpliciter*. Instead, she will be moved to act by the desire "not to resist another's attempt to exercise control over me such that I perform act *x* at his behest."

Once it is recognized that it is *this* desire that moves the compliant victim of duress to act, a satisfactory account of how it is that such a victim suffers from impaired autonomy can be developed, along lines similar to that of the Simple Analysis. On this new account (the *Revised Simple Analysis*) the compliant victim of duress *is* autonomous with respect to her effective first-order desire to fail to resist her threatener and thus to perform the act that he requires of her. (This satisfies the first intuition expressed at the beginning of this section.) However, since the satisfaction of this desire involves her relinquishing control to her threatener, the victim of duress suffers from impaired autonomy with respect to her compliant action. (This satisfies the second intuition.) Moreover, in addition to satisfying both of these intuitions, the Revised Simple Analysis also satisfies the two requirements that were outlined earlier in this section: (1) the effective first-order desire that the compliant victim of duress is posited to have includes, as part of its intentional object, a reference to the specific act that she is required to perform; and (2) it is the case that the victim herself decides whether or not to resist performing the act that her threatener requires of her.

VII. Conclusion

It is now time to take stock. I have argued that Frankfurt's and Dworkin's highly influential analyses of identification and personal autonomy have been misunderstood, for it has been assumed by their critics and supporters alike that they are committed to the Assumption of the Tran-

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press

sitivity of Autonomy when, in fact, they are not. It has also been argued in this essay that, once it has been recognized that neither Frankfurt nor Dworkin are committed to this assumption, Thalberg's objections to their hierarchical approach to analyzing autonomy are unfounded. However, that this is so can provide only cold comfort to the proponents of such hierarchical analyses of autonomy, for even when these analyses are properly understood, they still suffer from significant theoretical difficulties.

Nevertheless, this discussion of Frankfurt's and Dworkin's analyses did not arrive only at this rather dispiriting conclusion. Rather, in making clear both what the Assumption of the Transitivity of Autonomy is and the fact that it can be rejected, this discussion paved the way for development of the Revised Simple Analysis of why it is that a compliant victim of duress suffers from impaired autonomy. The Revised Simple Analysis satisfies both of the (apparently conflicting) intuitions that surround the question of whether or not a person who compliantly acts under duress thereby suffers from impaired autonomy.

I noted at the start of this essay that the issues that I planned to discuss are not only of parochial interest to those who are concerned with Thalberg's criticisms of the hierarchical approach to analyzing personal autonomy, but are also of interest to others who are engaged by the perennial question of whether a person who compliantly performs an action under duress thereby suffers from impaired autonomy. However, the issues in this essay are of even broader interest than this, for once the possibility of rejecting the Assumption of the Transitivity of Autonomy has been explicitly recognized, one can also see that it is possible that a theoretically complete analysis of personal autonomy might require that different conditions hold for a person to be autonomous with respect to her desires and with respect to her actions. This recognition is important, for such distinctions may well have significant implications for those areas of moral philosophy in which the nature and value of personal autonomy play a central role.

Philosophy, Louisiana State University

https://doi.org/10.1017/S0265052503202053 Published online by Cambridge University Press