https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

HOW MUCH SHOULD WE VALUE AUTONOMY?*

By Marina Oshana

I. Introduction

Autonomy generally is a valued condition for persons in liberal cultures such as the United States. We uphold autonomous agents as the exemplar of persons who, by their judgment and action, authenticate the social and political principles and policies that advance their interests. I will begin by examining the concept of autonomy in Section II of this essay. In Section III, I will explore the idea that autonomy is valued because autonomous agents are persons whose judgment and actions serve to advance their interests in a democratic society. But the focus of this essay is on the phenomenon, which is not implausible in a culture such as that of the United States, of being "blinded" by the ideal of autonomy. What happens if we value autonomy too much?

I will examine three possible outcomes. One, our commitment might lead us to believe, falsely, that all persons deserve autonomy, such that incursions of autonomy are never justified. This situation can be avoided once we settle upon a class of persons who, by their conduct, deserve protection of their autonomy. This classification rests on a negative test for desert modeled on John Stuart Mill's *harm principle*. Section IV will be devoted to an exploration of this issue.

A second problematic outcome arises when the premium that we place upon autonomy is tested by the very independence that it secures for us. The events of September 11, 2001, have compelled us to question the extent to which we should (and do) continue to regard autonomy as a good. We must confront the possibility that extreme measures that abrogate full autonomy will need to be undertaken in exchange for heightened security. Relaxed civil liberties protections, intensified scrutiny, and modified legal standards illustrate this tension.[1] Because the scope of this

* I thank the editors of *Social Philosophy & Policy* for their thoughtful comments. I am also grateful to colleagues who critiqued an early version of this essay. In particular, I benefitted from the insights of Ellen Frankel Paul and James Stacey Taylor.

[1] The following account offers some insight:

> Consider the case of Rabih Haddad, a Lebanese national active in the Muslim-American community in Ann Arbor, Mich[igan]. Mr. Haddad, 41, is among 326 individuals jailed in the federal investigation being conducted in unprecedented secrecy since the Sept. 11 attacks. . . . Arrested at his Ann Arbor home on Dec. 14 on a minor immigration violation, Mr. Haddad has been shunted through a series of closed-door court hearings and detention facilities in Detroit, Monroe County, and Chicago. His crime? No official will say, although it apparently has something to do with the Global

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

essay does not permit an exploration of this issue, I mention it primarily for consideration, briefly touching on the subject in Section V.

An injudicious attachment to the ideal of autonomy might result in a third state of affairs that is more pernicious and less easily resolved than the previous two. The worry is that, if we value autonomy too much, we might advocate the use of paternalistic measures to compel persons whom we identify as nonautonomous, or insufficiently autonomous, to become (more) self-directed. Is it coherent to attempt to force autonomy in a person by means that deny autonomy? I will turn to this discussion in Sections VI and VII.

II. The Concept of Autonomy

I understand personal autonomy as the condition of being self-directed, of having authority over one's choices and actions whenever these are significant to the direction of one's life. Personal autonomy as I construe it is a "global" phenomenon, a property of a person's life that expresses and unifies the will and choices of the person.[2] By contrast, the "local" or occurrent sense of autonomy is a property of a person's acts or desires considered individually, and pertains to the manner in which a person acts in particular situations.

In the global sense, a self-directed individual is one who sets goals for her life, goals that she has selected from a range of options and that she can hope to achieve as the result of her own action. Such goals are formulated according to values, desires, and convictions that have developed in an uncoerced fashion. They are goals that the individual would affirm as important to her were she to reflect upon their origin and content. One who does not care about her goals, or lacks goals altogether, might be moved about by others to the extent that self-direction becomes illusory. In addition, and most important, persons who are autonomous are parties to ongoing social relations that enable them to direct their lives

> Relief Foundation of Chicago, which he co-founded. Muslims say the organization is a charity but federal officials suspect it was funneling money to the al-Qaeda terrorist organization. All of the immigration court proceedings involving Mr. Haddad have been secret. . . . He is being held indefinitely and, in effect, incommunicado. . . . Mr. Haddad may be guilty, or he may be innocent, although that's not the issue here. It's the manner in which he is judged. If his detention and trial—assuming there is one—were to be handled in public, there could be little room for accusations of unfairness. As it is, secrecy invites skepticism that is corrosive to the trust that underlies our system of democratic government.
>
> Editorial, "Justice in Open Court," *The Toledo Blade*, March 12, 2002, A6.

[2] A "global" or dispositional phenomenon of autonomy is developed by Robert Young in his *Personal Autonomy: Beyond Negative and Positive Liberty* (New York: St. Martin's Press, 1986). Also see Paul Benson, "Autonomy and Oppressive Socialization," *Social Theory and Practice* 17 (1991): 385–408; and Paul Benson, "Free Agency and Self-Worth," *Journal of Philosophy* 91, no. 12 (1994): 650–68.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

with a minimum of interference.[3] An autonomous person is able to meet her goals without depending upon the judgments of others as to the goals' validity and importance. One is autonomous when one is "an independent source of activity in the world."[4]

This definition suggests that an autonomous person is in control of her choices, her actions, and her will. Some philosophers suggest that a weak form of control suffices for autonomy—or, more properly, for responsibility—charging that a person can remain in "guidance control" of his choices, actions, and will even when subject to conditions that could undermine self-governance.[5] For example, a person who, for reasons of drug addiction, coercion, subordinate rank, or weakness of will, could not do otherwise than perform a particular act (ingest a drug, relinquish money to a mugger, execute a military order, or lapse from a diet) might nevertheless be deemed in control of his actions, and responsible for them, if he would have performed the act anyway, independently and of his own free will. Thus, guidance control is possible even in the face of factors that are sufficient to determine one's actions.[6]

While guidance control might suffice for responsibility, I advocate a more stringent interpretation of being in control for autonomy. When we say that a person is self-governing because she is in control of her actions and choices, we are saying more than that the person's actions coincide with preferences or values that are her own. We are also saying that the person has the power to determine how she will live. Being autonomous is not simply a matter of having values and preferences that mirror those a person holds under conditions in which control is absent. Rather, being autonomous is a matter of directing one's life according to such values and preferences.

Autonomy or self-directedness, so described, calls for positive freedom. This is not just independence from the directives of others, not just unobstructed authority over the domain of one's life, and not simply rugged individualism of the sort that negative freedom is said to supply. It is positive liberty of the sort associated with the desire and ability for self-governance, or the psychological resources for self-governance.

Control of the sort that autonomy requires assumes of an agent certain psychological characteristics and a history of experiences conducive to

[3] See Marina Oshana, "Personal Autonomy and Society," *The Journal of Social Philosophy* 29, no. 1 (1998): 81–102. This conception of autonomy incorporates Thomas Hurka's view of autonomy as the condition of intentional causal agency that involves deliberate choice among a variety of actions. See Thomas Hurka, "Why Value Autonomy?" *Social Theory and Practice* 13, no. 3 (1987): 361–82.

[4] Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1998), 206.

[5] For an extensive discussion of guidance control, see John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge University Press, 1998).

[6] Note that the species of control that I describe could be specified in compatibilist or incompatibilist terms. I make no claim about the metaphysics of control in this essay.

self-directed agency. These suggest a capacity for autonomy, consisting of the minimum of qualities that a person must possess in order to lead a self-directed life. This capacity is not a bare, potential condition of the kind that we expect might be realized at some future moment, for where capacity is understood in this looser sense—as potentiality—an infant, a comatose being, and an intelligent computer would all be capable of self-government. A threshold must be satisfied: one must, say, have reasonably astute cognitive skills and a developed set of values in order to be actually autonomous. As the possession of these qualities is a matter of degree, so the capacity for autonomy is a matter of degree and can be cultivated more or less successfully in persons.

So, let us assume the following. An autonomous individual has knowledge of her circumstances and of the effective forces that are operative in these circumstances.[7] Control of the relevant sort requires that an autonomous person have the capacity to make decisions about matters that are pertinent to the nature and the direction of her life, and that she be disposed to do so. Such decisions, for example, concern a person's choice of lifestyle, partners, and career. Impediments to a person's control may consist of inner, psychological obstacles such as neurotic compulsion, excessively low self-esteem, weakness of will, or addiction. But impediments to autonomy are frequently of an external, or social, nature.[8] Manipulation and intimidation carried out by others on the individual, unreasonable conformist attitudes and role expectations, sexism, racism, or poverty might all count as external or social impediments of the relevant sort. An autonomous person not only has the capacity for independent decision but also exercises it; the individual must not succumb to the

[7] To be in control in this sense is to be "sensitive to environmental circumstances so as to allow oneself as much elbow room as possible." See Daniel Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, MA: MIT Press, 1984). This requirement suggests that autonomous agents must be self-aware and able to avoid situations that undermine the pursuit of their life-plans. It also implies that they be rational in the sense described by Robert Young. Young says:

> [B]eing rational can be seen as significant [to autonomy] in the following two positive ways. First, it brings coherence into the relationship between a person's general purposes and his or her particular actions. Some degree of understanding of this relationship will be needed to ensure that actions performed on particular occasions do not seriously thwart or impede more dispositional concerns. Second, and more importantly, perhaps, rationality equips a person to assess critically the advice tendered by others, an increasingly important safeguard given the extent to which we are reliant on the testimony of others about matters of great moment like health, welfare, education, economic and political affairs and so on.

Young, *Personal Autonomy*, 13.

[8] Following Joel Feinberg, constraints upon autonomy may be categorized as encompassing internal positive and negative impediments, such as neuroses or lack of skill, respectively, and external positive and negative impediments, such as physical barriers or coercive threats and inadequate economic resources, respectively. See Joel Feinberg, *Social Philosophy* (Englewood Cliffs, NJ: Prentice-Hall, 1973), 13.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

well-intentioned or malevolent attempts of others to control her decisions, nor must she be disposed to impose impediments upon herself.[9]

Clearly, a number of individuals lack this capacity as I have described it. For example, a small child, an individual afflicted with Alzheimer's disease, and an insane person lack the rudimentary ability to be self-governing. Absent from all three is the characteristic of being a good "local sociologist," of apprehending the complexities of one's external environment, of consistently distinguishing malevolence from benevolence, and of comprehending the normative expectations of other persons and adapting one's behavior accordingly. Absent from all three is the power of self-appraisal and the ability to plan, to fix on preferences, and to function in a farseeing, deliberative, and self-protective manner. All three are creatures for whom certain forms of supervision and protection are appropriate. (Ideally, of course, children will acquire the necessary characteristics for self-determination as they mature.)

Now, if we value autonomy too highly, we might attempt to free a child from the supervision of her parents before she is qualified to care for herself. (As a product of the alternative education trend and a former emancipated minor of the 1970s, I can attest to the questionable benefit of premature liberation.) Valuing autonomy too much, we might act in haste, expecting those who cannot assume direction for themselves to do so.[10] Indiscriminately embracing the ideal of autonomy has social costs as well, because not infrequently, the community, or the state, will have to assume support for those persons who are ill-equipped to care for themselves. Too often, the result is that the full autonomy of these persons will be abridged.

These cases must be distinguished from circumstances that are faced by persons who possess the general capacity-conditions for self-determination, but who are prevented for various reasons from living autonomous lives. An imprisoned individual, for example, lacks liberty to interact with others in a manner that gives her control over the direction of her life. She also lacks a range of options, which we expect of one who can direct her own life. Negative freedom—an absence of interference or constraint—is, therefore, a necessary condition of autonomy. Positive freedom—liberty of the sort associated with the desire and ability for self-governance, or the psychological resources for self-governance—is needed for autonomy as well. But negative and positive freedom do not suffice for autonomy. A paraplegic who requires, but does not receive, adequate physical therapy, or perhaps lacks a vehicle to provide him with a minimum of mo-

[9] To quote Young: "[T]o be autonomous is not merely to have a capacity, nor the opportunity to exercise the capacity. Autonomy is an *exercise-concept*, to use Charles Taylor's phrase." Young, *Personal Autonomy*, 49.

[10] James Stacey Taylor has pointed out, in correspondence, that this may just indicate a mistake on the part of the parent in how to respond to the value of autonomy, or a misunderstanding of how to foster properly the development of autonomy. Taylor is correct, but in some cases—my own, for example—this misunderstanding was accompanied by a disproportionate emphasis on the worth of autonomy.

bility, has both negative and positive freedom, but because he must depend on others he may fail to live in a self-governing fashion. Similarly, a person suffers no lack of freedom simply because abject poverty forces him to depend on the willingness of his government and the good graces of others for the availability of social services that are essential to his support and survival. Nevertheless, he can be described as self-governing only in an attenuated sense. Autonomy, then, calls for the presence of certain social, political, and economic arrangements. An autonomous person's choices must not merely be unobstructed (by others or by internal obstacles) but, where realistic, these choices must be socially, politically, and economically within his or her reach.

Some persons may deliberately forge lives in which autonomy is absent, though they possess the capacity and the freedom to do otherwise. Consider the situation of a woman living under a Taliban regime such as that which controlled Afghanistan until 2001. Suppose that this woman has embraced the role of subservience and the abdication of independence that it demands, out of reverence, a sense of purpose, and an earnest belief in the sanctity of this role as espoused in certain passages of the Qu'ran. Having previously enjoyed a successful career as a physician, this woman has since chosen, under conditions free of whatever factors might disable self-awareness, and with a considered appreciation of the implications of her decision, a life of utter dependence. She can no longer practice medicine (indeed, she is no longer permitted access to information about the science of medicine). She is not permitted to support herself financially. She has no voice in the manner and duration of any schooling that her children, particularly her daughters, may receive. She must remain costumed in cumbersome garb—a burqa—when in public. She is forbidden to enter common places of worship. She knows that any transgression, any show of independence counts as heretical defiance and invites punishment both swift and harsh. But a life of subservience is consistent with the Taliban woman's spiritual and social values, provides her with a sense of worth, and satisfies her notion of well-being.

I think that it is evident that the Taliban woman is not autonomous. In a "local" or occurrent sense of the term, she has chosen autonomously. Nevertheless, she fails to be autonomous in a "global" sense for the obvious reason that the life that she chooses, and toward which she experiences no alienation, is a life in which she is systematically subject to the ultimate will of others. Although the Taliban woman is "master of her will"—her original decision was made autonomously, she willingly renounces her rights, and she continues to express satisfaction with the life that she has selected for herself—she now has no practical authority over her situation.[11] Although she lives in a manner consonant with her pref-

[11] The case of the individual who willfully relinquishes his rights of self-government (and does so in an authentic manner, under suitable psychological and historical conditions),

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

erences, and succeeds in achieving what she believes is in her best interests, the choices that she makes are guided almost entirely by the judgments and recommendations of others. Although the Taliban woman does what she wants, what she wants frustrates the exercise of autonomy.[12]

Since autonomy calls for more than upholding a person's values, it is the conditions under which a person lives that must provide the framework against which personal autonomy is continually assessed. A lack of autonomy on the part of the Taliban woman is not merely due to the fact that she comes to depend on others. Nor can this lack of autonomy be minimized by the fact that she might retain the respect of others. Rather, her lack of autonomy is determined by what this dependency entails for her in her daily life, and the respect of others does not compensate for this loss.

It is a stretch to call a person "autonomous" whose genuine valuing of subservience or unquestioned adherence to religious tradition leads her to live a life of dependency. It diminishes the concept of autonomy to call such a human being autonomous in these conditions, for human beings are distinguished from other creatures precisely because of their deliberative and creative capacities. We may call such a person autonomous if we mistake well-being for autonomy, but the road to autonomy is not always the road to achieving one's aims. What a person might have reason to do in order to secure autonomy can diverge from what she has reason to do in order to secure what she values, or what comports with her conception of well-being. For example, deeply religious persons might believe that their interests are best served by following, without question, the edicts of their leaders. Such persons will not value or seek autonomy.[13]

---

preferring to live under the dictates of a religious order, offers an interesting case of one who might preserve some important measure of autonomy. It also illustrates the extent to which a person's autonomy is a function of his relations with others. Contrast the situation of the Taliban woman with that of a monk. In the case of the monk, some autonomy is preserved in that, every day, it is up to the individual to decide whether to remain in the order and to continue living in a manner that denies him a fuller range of freedom. The monk retains autonomy over a series of ongoing decisions to be subservient. Although an institution has power over him sufficient to compel him to behave in a certain way, the monk can recall this power within whatever frame of time is designated by the terms that he has accepted, in much the same way that individuals have the freedom to renege on their marriage vows or the legal authority to dissolve the terms of certain contracts. (Breaking the vow or dissolving the contract may, of course, carry a penalty that is sufficiently burdensome to make autonomy an impossibility. And of course, the nature of the contract, in terms of what is required of the individual, will be important for assessments of autonomy.)

[12] This means that "identification" and satisfaction are not sufficient for autonomy, even in the absence of certain negative constraints, as analyses sympathetic to the work of Harry Frankfurt maintain. As Young states, "[W]e may identify with a certain occurrent motivation in such a way as to undermine comprehensive or dispositional autonomy." Young, *Personal Autonomy*, 43.

[13] The distinction between the two different ideals of self-realization and autonomy is important. Joseph Raz raises the distinction, stating that although autonomy is a factor that contributes to self-realization, "[t]he autonomous person is the one who makes his own life, and he may choose the path of self-realization or reject it. Nor is autonomy a precondition of self-realization, for one may stumble into a life of self-realization or be manipulated into

Alternatively, we may want to call the Taliban woman autonomous because we think of autonomy as a condition relativized to the satisfaction of a person's desires, or decided entirely by the stance that a person adopts toward her choices, desires, affective states, and personal relations. But this mistake results from confusing autonomy with one of its potential advantages, namely, that it can lead to personal contentment. If I am correct, then it is false that chosen social roles of any variety must be permissible if a person is to be self-directed.

For why *should* autonomy be a condition compatible with any conception of the good, with any social role, or with every life-plan? I suspect that those of us in liberal societies are drawn to the idea that autonomy can "in principle be satisfied by an indefinite number of ways of life," [14] and are wary of the idea that only certain social arrangements befit self-directed agents. Not everyone will include an autonomous life among the goals that he or she regards as integral to well-being, and to suggest that there is *an* ideal end or way of life is more likely than not to impose *our* ideal upon others. This smacks of hegemony.

Autonomy as I have described it can comport with a variety of social arrangements. But, as Robert Kane notes, what has been called "value pluralism," the idea that the legitimate "ends of men are many, and not all of them are compatible with others" [15] need not entail value relativism. Value pluralism of the kind that is required for free will and for autonomy "does not imply that any end or way of life is just as good as any other—for all persons, or for a particular person at a particular time." [16] It is not insensitive to state that the types of lives that a properly autonomous person can live are limited. Such lives need not accord with an archetype of autonomy that reflects the alleged perspective of Western society, where autonomy is valorized ("fetishized" might be more accurate if one accepts the broader allegation) as a condition of atomistic, self-created individuals, insulated from the influence and guidance of others.

To deny the autonomy of the Taliban woman is not to show disrespect or to demonstrate insensitivity to her values, choices, and commitments. One might well admire the woman for the depth of her commitment and the richness and fulfillment that this brings to her life, while rightly recognizing that such a life lacks autonomy. Similarly, one might esteem—even desire to emulate—an individual who forsakes all of her aspirations to lead a life of service. Nonetheless, it is not uncontroversial that we should value such a life and the social roles that it mandates. This is not because *our* culture does not happen to value this way of life, but because

it or reach it in some other way which is inconsistent with autonomy." Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), 376–77.

[14] Isaiah Berlin, "Two Concepts of Liberty," in Isaiah Berlin, *Four Essays on Liberty* (Oxford: Oxford University Press, 1969), 169.

[15] Kane, *The Significance of Free Will*, 203.

[16] Ibid., 200.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

it is a way of life that is inconsonant with autonomy, and autonomy itself is of considerable objective importance for all persons, whether or not it is of subjective importance to a particular individual.

III. Autonomy and Liberal Society

Some philosophers worry that autonomy is a condition that, at worst, necessitates the detachment of an agent from others and, at best, encourages this. Understanding autonomy in this manner, these philosophers question whether it is a desirable trait of persons. For example, these philosophers worry that insofar as the autonomous agent may be unresponsive to shared values and objective standards of good judgment, autonomy is a condition likely to disrupt various cooperative enterprises and relations premised on values such as caring and commitment. The worry is not that autonomy may nonetheless be valued immoderately. Rather, the concern is whether autonomy ought to be valued at all. I take issue elsewhere with these objections to an ideal of autonomy, and I will pursue the subject here only in an attenuated fashion.[17]

That autonomy as I have described it is valued as a means to realizing the ideal of democratic societies cannot be disputed. A democracy is, roughly, a sociopolitical alliance of agents who, through representation or direct participation, engage in the task of political governance. The judicial, the political, and the legislative systems of any liberal democracy rely on the participation of self-reliant, self-directed persons. It is through the activity of such persons that liberal political principles and policies are generated and garner legitimacy. Hence, a person who fails to be self-governing, whose reasons for choice and action are appropriated by others, lacks an essential component of democratic citizenship, namely, the authority to speak for oneself and to be an active participant in an important range of one's experiences.

In some circumstances this person may display a veneer of democratic involvement. She may, for example, vote, seek public office, or have legal action taken against her. Because she does not manage her own life-choices, however, the responsibility for representative governance that is assigned to her by the democratic ideal remains illusory. While such a person has liberty, she lacks autonomy. But the liberty of persons that is so essential to democratic society is of little practical value if the more primary value of autonomy is absent. Lawrence Haworth comments that in a situation of this sort, "although one confronts numerous and fecund options, one totally lacks [the] capacity to exercise them autonomously. One has choices to make, but consistently makes them heteronomously." [18]

[17] I address this concern in Marina Oshana, "The Autonomy Bogeyman," *Journal of Value Inquiry* 35, no. 2 (2001): 209–26.

[18] Lawrence Haworth, *Autonomy: An Essay in Philosophical Psychology and Ethics* (New Haven, CT: Yale University Press, 1986), 143.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

More generally, agent autonomy is essential if any society, democratic or otherwise, is to survive. Consider a society in which a despotic regime is sovereign. Safeguarding the status quo appears to call for an absence of autonomy on the part of the citizenry. Autonomy is not broadly valued; the populace consists of persons who, worthy or not, capable or otherwise, are accorded a status similar to that of children. Those few who are autonomous undertake a parental role, supervising the activities and lives of others, by measures typically harsh. A political regime of this character can be sustained only with difficulty, even assuming that the populace is able to exercise self-restraint and to configure their lives to the arrangements determined by the authoritarian regime. Because the majority of persons would not be autonomous and so would be unaccustomed to decision-making, risk-taking, and innovation, there simply will be a minimum of persons who are qualified to assume the administrative burden of such a society.

Sustaining a despotic regime is a challenge for a more fundamental reason. As persons, we care about our ability to leave a trace, or a lasting impression, upon the world, a legacy made unique by our involvement. And we want this legacy to reflect a life of self-governance. We do not, as a rule, wish simply to reside in the world or to be moved through it. We are not simply objects made happy through the activities of others. Rather, we want to engage in activities that reflect self-competence and to pursue projects that bear our imprint rather than the imprint of other persons. This claim is contentious, but I think that evidence bears its truth. Whether our choices and actions be for good or for evil, we care that these choices and actions are ours, and that in forming them and in executing them we do not simply borrow from others but realize our autonomy.

One may experience contentment as a member of a despotic society. Indeed, absent autonomy, the intensity of pleasure and preference satisfaction may increase—think of a member of Walden Two, the fictional utopian society described by B. F. Skinner, or of a person attached to Robert Nozick's "experience machine."[19] But these people know a different and, I would contend, inferior variety of satisfaction than the contentment that is known by one who is a subject and not just an object, an agent and not a patient. Haworth has a point when he states: "Pleasure and preference satisfaction lose value in proportion as the pleased or satisfied individual lacks autonomy."[20]

IV. Autonomy and Desert

Whether one who is capable of autonomy, or self-determination, is *worthy* of an autonomous life remains an open question. Autonomy is the

[19] See B. F. Skinner, *Walden Two* (New York: Macmillan Press, 1962); and Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), 42–45.

[20] Haworth, *Autonomy*, 183.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

default position; we assume that its *possession* is independent of the question of desert. But the idea that self-determination is a condition to be *promoted* only for deserving persons is quite plausible when self-determination is thought of as a characteristic that persons develop and retain in virtue of their relations to others. Although autonomy as I have been discussing it is a "global" state of persons, the issue of desert arises primarily when autonomy is regarded "locally," as a condition predicated of persons vis-à-vis a certain activity, situation, or class of rights enjoyed in a social context.[21]

The class of undeserving persons may be divided between those who are undeserving on moral grounds and those who are so for pragmatic reasons. Not every undeserving person is morally destitute. Consider the case of Nancy, a narcoleptic. Assume that Nancy is fully capable of assessing her motives for action and bearing responsibility for her actions. Assume, too, that although driving is not usually a necessary condition for autonomy, Nancy is employed as a travelling saleswoman. Driving is something that she must do if she is to keep her job, and jobs are hard to come by.

Nancy's occupation gives her reason to drive, although at grave risk to herself and others. By continuing to drive, Nancy exploits the qualities that befit her for self-determination in such a manner as to pose a danger to herself and to others. Denying Nancy the freedom to drive would diminish her autonomy, but in exercising self-determination, in acting autonomously, persons must be sensitive to the interests of others. Nancy's economic interests are important, and her subsistence needs must be met, but only in a way that poses no unjustified threat to others. (I say "unjustified" because there are, of course, numerous occupations that pose justified threats to others. Military activity, law enforcement, and the practice of medicine are examples.) As a rational agent and a member of society, Nancy has an obligation not to imperil others needlessly. Because she chooses to ignore this responsibility, she does not deserve autonomy with respect to her freedom to drive, and measures that restrict her autonomy are called for.

In a case such as Nancy's, the agent is regarded as an autonomous being, as a being in control of herself. She is also a responsible agent, able to appreciate the normative import of her conduct and accountable for her actions. Yet we find it neither necessary nor desirable to respect Nancy's autonomy in a way that we might think befits an ideal. The criterion

[21] In conversation and correspondence, James Stacey Taylor has pressed the following important point, which I can only mention in abbreviated fashion. This is that the idea that autonomy is conditional upon desert allows incursions upon autonomy whenever it is decided that a person has fallen short of some independently specified ideal of character or behavior. If, instead, autonomy is the default condition, and it is assumed that persons have a right (of sorts) to autonomy, then the burden of proof for justifiable incursions upon autonomy rests on the encroacher.

of desert acts as a constraint upon what we will do and should do in order to promote a person's autonomy. Valuing autonomy too highly may tempt us to overlook the desirability of this constraint.

Consider the following case.[22] Shelley is a schizophrenic. She is also a talented artist whose work is celebrated by the art establishment and commands a substantial price on the market. Unfortunately, Shelley is at her creative best when working under the momentum of her illness. Although her schizophrenia subjects her to torment, delusion, and unsettling antisocial behavior, it also supplies the thematic framework of her work, and enhances the vision and the breadth of her work. Taking her medication gives her life a prosaic quality: she feels lifeless and uninspired. Citing the values of artistic license and individual flourishing (not to mention the lucrative benefits that her artwork secures), Shelley chooses to discontinue use of her medication. Once off her medication, Shelley lacks autonomy; it is the schizophrenia, and not she, that motivates her behavior. But she regards her self-imposed, voluntary abridgment of personal autonomy as a manifestation of her individuality or self-expression. Are incursions to Shelley's autonomy justified? That is, is it permissible to force persons such as Shelley to take their medication when doing so means an irreplaceable loss of individual fulfillment?

What I will call the *negative test for autonomy* raises the issue of when a person's autonomy may justifiably be violated or rescinded. This test focuses on the criteria for desert. Under what conditions does the capable person no longer deserve protection of her autonomy? One way of formulating the test can be found in John Stuart Mill's effort to fuse an alliance between the spheres of liberty and authority. Mill argues that, since people are members of society (and of narrower, more formal associations), tensions will inevitably arise between the provinces of individual freedom and social authority.

Following Mill, I hold that under normal circumstances, individuality, as a manifestation of autonomy that we value, must be protected from the "tyranny" of social custom and the collective authority of society. This tyranny seeks

> to impose, by means other than civil penalties, its own ideas and practices as rules of conduct on those who dissent from them; to fetter the development and, if possible, prevent the formation of any individuality not in harmony with its aims. . . .[23]

[22] The example of the mentally ill artist was raised by Ellen Frankel Paul and others in discussion.

[23] John Stuart Mill, *On Liberty*, ed. Currin V. Shields (1859; reprint, New York: Macmillan, 1956), 7.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

Because some threats to individuality are unavoidable, the variety of control and the extent of the power that can rightfully be exercised by society over an individual must be ascertained.

Mill does this by appealing to a general-interest principle known as the liberty principle. Negatively formulated as the *harm principle*, it states that

> the sole end for which mankind are warranted, individually and collectively, in interfering with the liberty of action of any of their number is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. *His own good, either physical or moral, is not a sufficient warrant*.[24]

We rely on something very much like Mill's harm principle in determining when autonomy ought to be protected and preserved, and when it ought not to receive such protection. Like individuality, autonomy is a highly valued state for capable and deserving persons and requires for its preservation freedom from undue interference, or negative freedom.[25]

Additionally, the harm principle establishes useful parameters for individual autonomy by showing that autonomy can be cultivated "within the limits imposed by the rights and interests of others."[26] It does this by employing a distinction between self-regarding and other-regarding conduct. Self-regarding actions, thoughts, and opinions are, on Mill's account, those in which other members of society have no interest, because the expression of these affects only the individual himself directly and in the first instance. Mill maintains that *any* interference with those actions of an individual that are purely self-regarding constitutes an illegitimate violation of that person's self-determination. Thus,

> [m]en should be free to act upon their opinions . . . to carry these out in their lives without hindrance, either physical or moral, from their

[24] Ibid., 13, my emphasis.

[25] Berlin challenges Mill's claim that negative freedom—freedom from undue interference—is a necessary condition for human flourishing. Berlin argues that the link between the notions of negative liberty and the ideal of character that Mill envisions "is, at best, empirical," since ". . . integrity, love of truth, and fiery individualism" can be highly valued, and indeed can flourish, "in severely disciplined communities" as readily as they can in more tolerant communities. (Berlin, "Two Concepts of Liberty," 128.) I am not certain that this is true. Consider that Mill defends the value of not only truth and integrity, but also social progress. In order for the latter to ensue, the community must be responsive to the creatively nonconformist individual, and so it must loosen its reign upon the individual. Moreover, Mill claims only that *undue* interference must be absent; it is questionable whether he would call every instance of influence and discipline (such as that which certain religious communities embrace) "undue."

[26] Mill, *On Liberty*, 76.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

> fellow men, so long as it is at their own risk and peril. This last proviso is of course indispensable.[27]

When a person behaves in ways that we find injurious to the self-determination of others, either by "doing evil" or by failing to prevent evil, or when a person violates certain accepted codes for social behavior, or abdicates "a distinct and assignable obligation to any other person or persons," even when this action consists of injury that the person does to himself, "the case is taken out of the self-regarding class and becomes amenable to moral disapprobation in the proper sense of the term."[28] No longer in the realm of liberty, such conduct becomes subject to morality and law, where punitive measures that include a diminution of personal autonomy may be taken as justified and legitimate. We may interfere with the affairs of another person, even those that are of profound importance and interest to that person's life, when these pose a threat to others.

V. When Valuing Autonomy Is an Issue

What are the legitimate restrictions that can be placed upon autonomy when circumstances are *not* normal? Are the harm principle and the corollary distinction between self-regarding and other-regarding conduct subtle enough to provide parameters for autonomy at such times?

In less than normal circumstances, threats to the autonomy of deserving individuals can be great. Protections from the tyranny of majority sentiment, from encroachment upon civil liberties and constitutionally mandated rights, and from the infliction of loss, falsehood, duplicity, and the like take on a heightened urgency. In the aftermath of the December 7, 1941 attack on Pearl Harbor, Japanese-American citizens who had harmed no one were dispossessed of their autonomy and their property, and were treated in this way with the overwhelming approval of the United States populace. In the aftermath of the unprecedented terrorist events of September 11, 2001, circumstances have been anything but normal. A rekindled patriotism has tempered the willingness of many American citizens

[27] Ibid., 67. Mill correctly notes a "distinction between the loss of consideration which a person may rightly incur by defect of prudence or personal dignity, and the reprobation which is due to him for an offense against the rights of others." (Ibid., 96.) The difference here is one of a loss of respect or esteem from others in the first instance, and a loss of autonomy, or individuality, in the second. When a person's self-regarding actions are foolish, others may regard him with distaste, pity, or even contempt, but this lack of respect must never translate into a punitive act against the agent's individuality. In response to "that portion of [a person's] character and conduct which concern his own good . . . which do not affect the interests of others in their relations with him," we are entitled to visit upon the agent certain inconveniences or penalties. Such penalties may be quite severe—they may include forms of ostracism and discriminatory treatment. But these must be only the natural and inseparable result of the actor's conduct. (Ibid., 95.) And they must go no further than is permitted by the expression of our own individuality.

[28] Ibid., 99.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

to criticize openly, or even question, the political and military policies of the United States. Restrictions that fetter individual autonomy have become commonplace and enjoy increasing public approval.[29] There is heightened support for surveillance cameras on select street corners of select urban neighborhoods. Passenger profiling and personal searches at public transport facilities have increased. Investigations of personal computer use and electronic communication have intensified. Persons have been detained, incommunicado, for unspecified crimes. Although restrictions on autonomy are not always or principally privacy concerns, privacy protections are generally essential for autonomy. The aforementioned security measures, both overt and covert, yield, *at a minimum*, a decline in the level of civil liberty that most of us cherish and regard as crucial to unimpaired self-governance. (Detainment by federal authorities for weeks or months unambiguously assails the detainee's exercise of self-government.)

If anything positive has come out of the horrific events of September 11, 2001, it is that the United States is a more unified, kinder, and gentler nation. Or so the story goes. But it is also a more vigilant and suspicious nation. So perhaps the following scenario is not too great a departure into fiction:

> The citizens of Bordertown are agitated by the events of September 11, 2001. Since that date, they have garnered increasing support for intensifying anti-immigration policies in the nation. Anti-immigration sentiment in the community runs high, and with some good reason. Bordertown is located in uncomfortable proximity to Migrantville, which is suspected as a breeding ground for terrorists. Although the law enforcement officials of Bordertown and of Migrantville have yet to confront indiscriminate violence against innocent members of the "migrant foreign element," the officials impose a curfew, rescind concealed weapons laws, and, in collusion with the media and local ecumenical councils, embark on a disinformation campaign that is intended to foster social unity.

[29] The Associated Press reported (June 12, 2002) that "four in five Americans would give up some freedoms to gain security": One-third of Americans surveyed favor making it easier for authorities to access private e-mail and phone conversations; more than 70 percent favor requiring U.S. citizens to carry identification cards with fingerprints. The *USA Patriot Act*, 107 P.L. 56, which was passed by the United States Congress in 2001, provides the Central Intelligence Agency with access to U.S. Justice Department records, including secret testimony developed in grand jury investigations. U.S. Attorney General John Ashcroft has issued executive directives allowing the Federal Bureau of Investigation (FBI) to eavesdrop on conversations between lawyers and their imprisoned clients in certain cases, and has given the FBI greater freedom to use wiretaps and to investigate people, even if they are not suspected of committing any crime. Certainly these amount to incursions upon privacy. How they affect the ability of persons to live self-governed lives remains to be seen: Self-governance and an open society go hand in hand.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

These measures plainly erode the autonomy of the citizens. Are such measures justified as being in the best interest of the populace? Even if the end supports the means in this case, it is still an open question whether the cost in diminished autonomy is excessive.

When are such encroachments on autonomy warranted? Post-September 11, 2001, do we confront a situation akin to that described by Thomas E. Hill, Jr.? Hill remarks:

> If . . . the only way to persuade someone to make a decision that will prevent a riot or a series of murders were to make an otherwise impermissible threat or a nonrational appeal to his weaknesses, then surely most would grant that such interference would be justified. Though important, autonomy need not be considered an absolute right.[30]

Such scenarios remind us that although autonomy is a condition that we value in persons, it is not the sole or overarching good. Security may surpass autonomy in value and may justify constraints on the freedom necessary for autonomy. Attributes of character such as compassion, generosity, service to others, or a willingness to compromise may also exceed autonomy in importance.[31] Isaiah Berlin assumes that there are values higher than autonomy that bear protection, and the protection of which calls for restrictions on individual freedom. He states:

> I do not wish to say that individual freedom is, even in the most liberal societies, the sole, or even the dominant, criterion of social action. We compel children to be educated, and we forbid public executions. These are certainly curbs to freedom. We justify them on the ground that ignorance, or a barbarian upbringing, or cruel pleasures and excitements are worse for us than the amount of restraint needed to repress them.[32]

VI. Paternalism for the Sake of Autonomy

By expanding the criteria for desert, we might rightly narrow the class of persons against whom infringements to autonomy are in general unjustified. Candidates must respect the harm principle, and they must

[30] Thomas E. Hill, Jr., "Self-Respect Reconsidered," in Thomas E. Hill, Jr., *Autonomy and Self-Respect*, (Cambridge: Cambridge University Press, 1991), 259.

[31] I am, then, in agreement with Feinberg when he notes that "autonomy is not the whole of virtue, and may be made to look bad if it keeps bad company. (Imagine an inflexibly conscientious Robbespierre)." Joel Feinberg, *Harm to Self*, vol. 3 of *The Moral Limits of Criminal Law* (New York: Oxford University Press, 1986), 40.

[32] Berlin, "Two Concepts of Liberty," 169.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

meet other conditions as well. Strengthening the criteria enables us to test our belief that autonomy is an optimal condition for persons. If autonomy is a condition that we think persons ought to enjoy, then to what extent should pains be taken to protect and promote that state? Is it permissible—and for the best—to take paternalistic steps to restore autonomy in deserving persons? When does interference with affairs that are of profound importance and interest to another person constitute a harm, a violation of the person's autonomy? As the stories of the Taliban woman and the schizophrenic artist suggest, certain attributes that might promote or strengthen a person's autonomy might be quite undesirable from the perspective of the good of the individual.[33]

While the harm principle permits the state to protect individuals against the choices and activities of others, and to restrict certain actions on the grounds of preventing harm to others, paternalistic actions cannot be justified by appealing to other-regarding harm. Let us understand paternalistic action as interference with a person's autonomy that seeks justification "by reasons referring exclusively to the welfare, good, happiness, needs, interests or values of the person. . . ."[34] We offend a person's autonomy by paternalistic means when we endeavor to impose on the person a conception of what is a worthy and proper life. We do this either by preventing the person from doing whatever she has decided to do, or by interfering with the way in which she reaches her decision.

[33] Consider the case of Mark, a megalomaniac. Mark can recover from his megalomania if provided with proper treatment. But in virtue of his neurosis, Mark is motivated to care for his family and to fulfill the role of a valued community leader. Because Mark is dominated by a neurotic disorder, he is not autonomous with regard to his behavior. Yet Mark is a more attentive parent and spouse, and a more estimable neighbor when motivated nonautonomously, by his neurosis, than he would be were he to undergo treatment for his megalomania and thus experience a full state of autonomy. Even if Mark would be better off were he to be freed of his neurosis, those who are close to him, whose interests he, presumably, holds dear, would likely suffer as a result.

[34] Gerald Dworkin, "Paternalism," *The Monist* 56 (1972): 65. I will confine my remarks to what Dworkin calls "pure" paternalism, and what Feinberg labels "direct" or "one-party" paternalism, where "the class of persons whose freedom is restricted . . . is identical with the class of persons whose benefit is intended to be promoted" by such restrictions (Dworkin, "Paternalism," 68). Feinberg contrasts paternalistic activity with the phenomenon of legal paternalism. Legal paternalism is expressed by the principle that "the need to prevent self-inflicted harm [is] a legitimizing reason for coercive legislation. . . ." (Feinberg, *Harm to Self*, 8). Legal paternalism concerns the manner in which criminal legislation can be granted moral legitimacy, and its tools are penal *regulations* that function as "coercive interferences with liberty" rather than paternalistic *actions* that need not involve or endorse coercive legislation. Hence, the target of paternalistic rules is always liberty of action, and paternalistic rules always involve the use of coercion or force. Legal paternalism, unlike the harm principle, is not mediated by the maxim that "a person is not wronged by that to which he consents. . . . B's consent to A's action, even though that action is harmful or dangerous to B's interests, exempts A from criminal liability under the harm principle, but does not exempt him under indirect legal paternalism." (Feinberg, *Harm to Self*, 11.)

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

Four things are required in order for an act to be paternalistic:

1. Since paternalism *just is* that which offends autonomy, those who are subject to paternalistic action must be capable of autonomy (though not necessarily deserving of autonomy). Therefore, children cannot be subjects of paternalistic action so construed.
2. The act uses compulsion (although not necessarily coercion) to promote the well-being of a subject, or to achieve some benefit that may or may not be recognized as a benefit by the one for whom it is intended.[35]
3. In cases where we wish to protect a person from a harm, incurring the harm requires the active cooperation of the victim.
4. To the extent that the agent is aware of the paternalistic measure, the agent does not (or would not) want to be treated in this way.[36]

Paternalism usurps autonomy because it substitutes one person's judgment for another's. Suppose that a friend in the United States military is tormented about his homosexuality, torn between his need to be forthright about his sexual orientation and his allegiance to the military policy of "don't ask, don't tell" on openness about homosexuality. If I reveal the homosexuality of my friend to his supervising officer because I believe that doing so will alleviate the unhappiness that my friend is experiencing, and in the long run will best promote his psychological esteem, I invade my friend's privacy, and I invade his autonomy. And my intervention most certainly invites substantial changes in my friend's life that he might fail to welcome. My motives need not be malicious. But because I have acted to promote what I perceive to be this person's well-being, where doing so is not welcome, my action is paternalistic.

Disquieting scenarios involving paternalism occur when persons who possess the capacity for self-determination *and* who are at least minimally deserving of self-determination (since they do not violate the harm principle) find their self-determination jeopardized by being subjected to even the most benevolent of paternalistic gestures. Interferences of this sort are worrying, if only because they threaten to make systematic violations of liberties more acceptable. Mill, Hill, and Berlin, among others, voice such worries.

[35] To "compel" action is to necessitate an action by force of some sort: it may include gentle persuasion or temptation. To "coerce" is to compel behavior by use of pressure, threats, or intimidation.

[36] Note that this list does not specify anything about the *manner* in which a person is prevented from exercising her autonomy. Whether manipulation, brute force, or rational persuasion is used is insufficient to determine cases of paternalism. Paternalism just consists in the practice of imposing a putative good on a person when that person does not welcome the imposition. See Dworkin, "Paternalism," 68.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

In speaking of the value of individuality and the nature of self-regarding conduct, for example, Mill tells us that the person most interested in the well-being of an individual is the individual himself, and if the individual is a reasonable, mature adult, he should be granted "perfect freedom, legal and social, to do the action and stand the consequences."[37] No other person should tell him "not to do with his life for his own benefit what he chooses to do with it."[38]

Even when persons engage in behavior "which experience has shown not to be useful to any person's individuality," or act in ways that are "injurious to happiness and . . . a hindrance to improvement,"[39] intervention, says Mill, is permitted only to prevent a graver harm to the larger social body. In order to justify compelling a person to do what another believes is best, Mill adds, "the conduct from which it is desired to deter him *must be calculated to produce evil in someone else*. . . . Over his own body and mind, the individual is sovereign."[40]

Hill's concern is with the "many ways in which narrow utilitarian thinking can foster unwarranted interference in others' lives."[41] Although Hill restricts his discussion to the phenomenon of lying, his remarks have implications for the more general phenomenon of paternalistic interference. He offers several conceptions of autonomy, and he argues that each of these incorporates moral principles that oppose the phenomenon of benevolent lying as one that demonstrates a lack of respect for the autonomy of persons.

Hill claims that it is irrelevant whether the paternalistic intervention is intended to protect a person from discomfort or to augment that person's independence. When one knowledgeable, sane person is subject either to deliberate deception or to a voluntary withholding of information by another, "one's opportunity to live in rational control of one's life" is curtailed.[42] Paternalistic intervention displays a disrespect for persons by treating them as incapable of choosing their own courses of life.

Hill argues that, if we agree that individuals have a right to autonomy, then interferences to this right via paternalistic gestures such as benevolent lying will be justified only if the action that would occur in the absence of the paternalistic gesture is so disastrous as to override the individual's claim to autonomy as a right. What would count as suffi-

[37] Mill, *On Liberty*, 92.
[38] Ibid., 93.
[39] Ibid., 98.
[40] Ibid., 13, my emphasis. Note that what Mill appeals to here in defending the right of noninterference is the utility principle. His injunctions against paternalism as well as the harm principle emanate from the principle of utility. Berlin points out that other principles, such as the categorical imperative, prescriptions of natural law, or the inviolability of a social contract, might equally be employed. See Berlin, "Two Concepts of Liberty," 127.
[41] Thomas E. Hill, Jr., "Autonomy and Benevolent Lies," *The Journal of Value Inquiry* 18 (1984): 251.
[42] Ibid., 265. The offensive character of a paternalistic lie is commensurate with the extent to which the activity affected by the lie is significant for the individual's life-plan.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

ciently disastrous remains an open question. Would the self-regarding aim of the individual to destroy his own autonomy qualify?

Berlin expresses similar concerns, though he restricts his comments to cautioning us against paternalistic acts that arise out of a particular interpretation of positive freedom. Specifically, Berlin asserts that the positive idea of freedom, understood as self-mastery, invites unwarranted interference when it is interpreted as the domination of the agent by his "real" or "ideal" self, as opposed to his baser, impulsive, and passionate nature.[43] Although positive freedom understood as "freedom by self-control" lends itself most readily to this partitioning of the self, Berlin notes that an account of negative freedom can lead to a similar result. Berlin states: "[T]he self that should not be interfered with is [not] the individual with his actual wishes and needs . . . but the 'real' man within, identified with the pursuit of some ideal purpose not dreamed of by his empirical self."[44] This account of the person accords reason the role of the dispassionate "true" self, and relegates the more empirical aspect of the agent to the lower, animal self. These persona are taken to be discordant: A struggle ensues between "the transcendent, dominant controller, and the empirical bundle of desires and passions to be disciplined and brought to heel."[45]

This idea of self-mastery assumes that it is only the self *qua* rational will whose wishes must be heeded, even when the individual demands otherwise. The outcome, Berlin argues, is that this true self takes on a fetishized character. It somehow becomes identified with whatever larger social or institutional collective body is given reign over the expressed wishes of the individual. Berlin remarks that just as

> [t]he reason within me . . . must eliminate and suppress my "lower" instincts, my passions and desires, which render me a slave; similarly . . . the higher elements in society—the better educated, the more rational, those who "possess the highest insight of their time and people"—may exercise compulsion to rationalize the irrational section of society. For . . . by obeying the rational man we obey ourselves: not indeed as we are, sunk in our ignorance and our passions . . . but as we could be if we were rational.[46]

The phenomenon that Berlin describes gives rise to what I will call *rational consent* varieties of paternalism. If the rational consent model is correct, then paternalistic action will be justified as long as it is taken in the name of the agent's true self. The action will not be viewed as one that

[43] Berlin, "Two Concepts of Liberty," 132, 134.
[44] Ibid., 134.
[45] Ibid.
[46] Ibid., 150.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

the individual would want *were he otherwise*, but will be said to express what the individual does in fact really (though not expressly) want. Under such circumstances, argue advocates of the rational-consent justificatory scheme, no interference to the individual's freedom in fact occurs. Since the action is carried out by the individual's real self against his "less real" self, the individual is, in fact, the actor. As long as a person is not interfered with by factors alien to the person's rational (even if unrecognized) self, the person is autonomous.

However, this belief—that actions made by proxy for the rational true self are efforts that the individual *in fact* wills autonomously, even when these actions evoke the most intense rejection and resistance from the individual—is counterintuitive. It also carries tremendous potential for exploitation. Berlin, for example, worries that the two-self ideal breeds moral and political despotism. The assumption that the agent who does not act from his rational self fails to act freely entails that no brand of paternalistic action against such a person (no matter how sane that person appears and regardless of how offensive the paternalistic action) can be deemed a violation of the person's autonomy. Once the rational consent position is adopted,

> I am in a position to ignore the actual wishes of men or societies, to bully, oppress, torture them in the name, and on behalf, of their 'real' selves, in the secure knowledge that whatever is the true goal of man . . . [it] must be identical with . . . the free choice of his 'true', albeit often submerged and inarticulate, self.[47]

Requiring that the autonomous individual be ideally rational, disposed to sublimate her empirical nature for her dispassionate side, and making this the locus of individualism and personal freedom countenances a state of affairs potentially hostile to the negative species of freedom espoused in Mill's liberalism. It also severely limits the class of people whom we call autonomous. The idea that there is a particularly correct way of life, and a correct set of creeds that sustain the rationalist vision, invites intolerance and the suppression of opinions. No longer is autonomy an ideal associated with individual responsibility and self-mastery but, ironically, autonomy becomes a species of self-perfection. The case of the Taliban woman commands our attention in this regard. Ought paternalistic measures be taken in the name of her true self to remove her from the grip of the Taliban?

VII. Autonomy and the Value of Paternalism

As the preceding discussion indicates, it is especially difficult to weigh the value and desirability of autonomy against that of paternalism in

[47] Ibid., 133.

situations where a person's self-directed actions threaten the person's autonomy, and where paternalism might avert this threat.

Furthermore, there are varieties of paternalism. Gerald Dworkin distinguishes between "soft" (or "weak") and "hard" paternalistic acts. A necessary condition of soft paternalism is that the person for whom we act paternalistically is incompetent in some sense.[48] Interference with a person's autonomy is permitted when there is thought to be an occurrent or a dispositional defect in the decision-making capacities of the person who is interfered with, and when it is believed that consent would be forthcoming were the person's decision-making capacities revived.

The real challenge is posed by cases that we might want to characterize as appropriate targets for paternalistic intervention, but in which we lack a credible reason for deeming the agents incompetent and their actions involuntary. In such cases, Dworkin notes, "hard paternalism may be the only position which can justify restrictions on such actions."[49] "Hard" paternalism is the view that an acceptable reason for paternalistic legislation is the necessity of protecting competent adults, against their wills, from the harmful consequences of even their fully voluntary undertakings. Dworkin attempts to avoid the hard paternalistic position in such cases by arguing for a "hypothetical consent scheme for justifying paternalism" via the "soft" view. Justifications of paternalism based on hypothetical consent turn on the assumption that, to some degree, competent adult individuals are vulnerable to the same cognitive, emotional, and epistemic failings as are their less competent counterparts. It is also assumed that rational persons wish to protect themselves from cognitive deficiencies and incompetence. Dworkin's idea is that soft paternalism is justified when an agent's action is voluntarily undertaken, but it poses a real risk, of which the agent is not aware or does not sufficiently appreciate. Paternalistic protections or limitations on certain conduct are appropriate if they are ones that rational individuals antecedently would willingly and collectively establish, and would consent to as "social insurance policies."[50]

[48] Gerald Dworkin, "Paternalism: Some Second Thoughts," in Rolf Sartorius, ed., *Paternalism* (Minneapolis: University of Minnesota Press, 1983), 107. Feinberg also distinguishes between "soft" and "hard" paternalism but, instead of making soft paternalism a feature of the actor's competence, views it as determined by the involuntary nature of the harm suffered by the actor. The idea is that we can permissibly protect a person from self-regarding harm only when the harm is substantially nonvoluntary (or when intervention is required in order to determine whether it is voluntary or not). Soft paternalism is intended to affect only "wrongfully" suffered harm, i.e., harm that the subject does not consent to. Given that the person fails to consent to the harm, Feinberg rightly questions how action or legislation taken against the harm can truly be called paternalistic.

[49] Ibid., 109.

[50] Dworkin, "Paternalism," 78. The parameters of justifiable interferences will be determined in either of two ways. We might focus on the rationality and competence of the individuals involved or we might examine the decisions that the individual makes. On the former approach we ask whether the individual suffers from some degree of cognitive

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

For example, Dworkin notes that if we can show that a person who values slavery for its instrumental powers—believing, perhaps, that slavery will maximize some other, more highly desired good—suffers from a misapprehension, or a mistaken calculation about how best to obtain the primary good that he seeks, then we will be justified in imposing certain mandates regarding this person's behavior. Doing so "minimizes the risk of harm . . . at the cost of a trivial interference with . . . freedom."[51]

But when the targets of intervention are adults, the hypothetical consent model that is invoked to justify soft paternalistic restrictions lends itself to the worries that Berlin explores regarding rational consent models. Surely, every interference with freedom is not trivial, especially from the perspective of the one who is interfered with. Forcing the Taliban woman to abandon her way of life counts to her as a substantial harm, the cost of which is her purported flourishing. Yet Berlin concedes that there are times when it is defensible to compel a person to do one's will, or the will of a larger group, under the guise of promoting what the person would, in a more informed, mature, and reasonable state prefer. (Recall that civil liberties were retracted for the alleged benefit of the citizens of Bordertown and Migrantville.) Which interferences with individual freedom are defensible will reflect our understanding of what constitutes a well-developed life. Although action is taken for the sake of the agent, is said to be in the agent's interests, and is against the express wishes of the agent, in this case the difference—which Berlin believes to be a crucial one—is that interference is made in the name of an action that the individual would want *were he otherwise*. Paternalism is warranted in light of presumed future consent, as opposed to the presence of a "true, rational self." Again, the assumption is that the person will come to recognize and welcome these interferences as being in his or her best interest.

I do not think that Berlin's willingness to embrace a "best interests" defense of paternalism is any less subject to abuse than is the "true self" model for justification. Justifications for weak paternalism based on assumptions of rational consent or of future consent are problematic even though, as Dworkin notes, paternalistic interference is a less bitter pill where it (allegedly) "preserves and enhances for the individual his ability to rationally consider and carry out his own decisions."[52] Nonetheless, I

incompetence that causes her to weight abnormally some of her values, or to "discount unreasonably the probability or seriousness of future injury" ("Paternalism: Some Second Thoughts," 108). On the latter approach we wonder whether the person selects courses of action that are "far reaching, dangerous, and irreversible" ("Paternalism," 80), or are made under conditions of duress, or involve dangers not sufficiently understood or appreciated (ibid., 82).

[51] Dworkin, "Paternalism: Some Second Thoughts," 110. Dworkin believes that once the consent of the subject is acquired, "in interfering with such people we are in effect doing what they would do if they were fully rational. Hence we are not really opposing their will, hence we are not really interfering with their freedom" ("Paternalism," 77). But it is obvious that if consent is granted, then the action is no longer paternalistic.

[52] Dworkin, "Paternalism," 83.

believe that there are cases in which *strong* paternalistic measures ought to be employed as a means of discouraging certain autonomous behavior, of compelling certain autonomous behavior, and of enhancing autonomy where the capacity is present, but unfulfilled. I believe that this holds even when the target of the paternalistic gesture has not violated the harm principle, and so has not acted in ways that clearly permit infringements of autonomy.[53] Strong paternalistic intervention is sometimes needed to preserve the autonomy that is threatened by a competent and deserving person's self-regarding conduct. Robert Young argues for this point:

> Suppose *S* knows that heroin addiction causes severe physical harm and likely death before 30 years of age, but still chooses to take the drug because he wants the pleasure of the moment more than anything else. Assume, furthermore, that we independently have good grounds for believing *S* is emotionally stable and of sound reason. A policy of weak [soft] paternalism cannot in such a case justify intervention to prevent *S*'s taking heroin. A strong paternalist . . . would argue for intervention where the consequences of *S*'s action would be to undermine other more dispositional commitments.[54]

Since employing paternalistic measures to compel autonomy seems contradictory, it is the phenomenon on which I will focus. I believe that an argument in favor of strong paternalistic measures, even those taken against self-regarding harms, can be given when such measures are employed to minimize the sort of self-regarding harms that jeopardize autonomy, or individuality. This kind of argument challenges Mill's claim that the individual is always the best judge of his own interests, and that the evils produced by interferences with liberty are worse than any benefit obtained. Although I agree with Mill that there is a general principle that the individual is the best judge of his own interests, my view is that there are important exceptions to the principle that justify paternalistic interference, exceptions that surpass those that Mill himself acknowledged.

Mill argues that because paternalism affects conduct that is entirely self-regarding, it cannot be justified by citing the need to protect the interests of others. And because the individual is, arguably, the best judge of his own welfare, it is unlikely that compulsion would advance the interests of the individual, and it is quite likely that compulsion would

[53] The case of Mark, the megalomaniac, above at note 33, is an illustration of discouraging certain autonomous behavior. In challenging the assumption that autonomy is, for Mill, to be desired on instrumental grounds alone, James Bogen and Daniel M. Farrell note: "Autonomy may be desired for its own sake, even if it ceases to produce the mental state for which it was originally desired." See James Bogen and Daniel M. Farrell, "Freedom and Happiness in Mill's Defence of Liberty," *Philosophical Quarterly* 28, no. 113 (1978): 334.

[54] Young, *Personal Autonomy*, 68. The idea is that while *S* exhibits autonomy of desire, paternalistic measures might be employed to preserve *S*'s autonomy of future action.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

produce greater evil than it would prevent. Mill's reasoning is founded on the utilitarian cost-benefit calculus, but the claim that he wishes to defend is not supported by the cost-benefit analysis.

First, there is little proof that adult individuals always know their own interests. It is in a person's interest to be autonomous, but a person can be mistaken about whether or not he is autonomous, and about what his autonomy consists in. If autonomy is something we want to promote, then the failure of people to decide accurately about their autonomy might offer one reason in favor of paternalistic interferences, even when a person has decided in what he believes is his best interest. (Recall Hill's claim that the right to autonomy must be preserved, even when it is the self-regarding aim of the individual to destroy this right in himself.)

Second, as Mill himself should acknowledge, paternalistic action sometimes constitutes a lesser evil, or a greater good, than would obtain should the agent's action that it restrains go unchecked. Allowing a person to autonomously pursue a life in which more dispositional or global interests are circumscribed does not always make the person better off, even if the person is happy with the result. Mill writes:

> [An] . . . exception to the doctrine that individuals are the best judges of their own interest, is when an individual attempts to decide irrevocably now what will be best for his interest at some future and distant time. The presumption in favor of individual judgment is only legitimate, where the judgment is grounded on actual, and especially on present personal experience; not where it is formed antecedently to experience, and not suffered to be reversed even after experience has condemned it.[55]

For example, the value that Mill accords autonomy can be used to justify an argument against self-imposed slavery. Mill counts freely chosen slavery among those actions over which the individual is not sovereign, and says:

> The ground for thus limiting [the slave's] power of voluntarily disposing of his own lot in life is apparent, and is very clearly seen in this extreme case. The reason for not interfering, unless for the sake of others, with a person's voluntary acts is consideration for his liberty. . . . But by selling himself for a slave, he abdicates his liberty; he forgoes any future use of it beyond that single act. He therefore defeats, in his own case, the very purpose which is the justification of

[55] John Stuart Mill, *Principles of Political Economy*, 2 vols. (1848; reprint, New York: P. F. Collier & Sons, 1900), 459.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

> allowing him to dispose of himself. He is no longer free; but is thenceforth in a position which has no longer the presumption in its favour that would be afforded by his voluntarily remaining in it. *The principle of freedom cannot require that he should be free not to be free*. It is not freedom to be allowed to alienate his freedom.[56]

Mill's use of the term "freedom" equivocates between "liberty of action" and "individuality" or "autonomy." Thus, a better statement might be that people should not be at liberty to relinquish their autonomy. But his point is clear. Individuality or autonomy is too important a characteristic for persons to be without. Since it is that element of persons upon which all other forms of freedom are grounded—including the freedom to act—autonomy cannot be something that we are at liberty to dismantle. Consensual slavery, regardless of the gains that it might provide and aside from any benefit to the enslaved, transforms the human subject into a possession or object of another and accordingly defiles the enslaved individual's autonomy. Once autonomy has been cast off, the person is in no position to expect others to treat him with the respect for his individuality that typically disallows paternalism.

The following passage from Chapter 3 of *On Liberty* lends credence to Mill's support of paternalism in such cases. Mill writes:

> He who lets the world, or his own portion of it, choose his plan of life for him has no need of any other faculty than the ape-like one of imitation. . . . It is possible he may be guided in some good path, and kept out of harm's way. . . . But what will be his comparative worth as a human being? It really is of importance, not only what men do, but also what manner of men they are that do it.[57]

The idea here is that autonomy acquires its value from other goods that it makes viable, and from what it enables persons to do with their lives. Strong paternalism is warranted solely to preserve more extensive autonomy or self-direction, greater freedom, and moral agency. We care about being moral agents, even if, on occasion, we autonomously act in ways that are not expressive of that concern. Thus, cases of voluntary slavery, and perhaps cases in which a person willingly embraces a life of subservience, as does the Taliban woman, offer an exception to prohibitions on paternalistic conduct. Such voluntary contractual obligations are prohibited because maintaining the freedom to make future choices is key. If we accept this account, as I am inclined to do, then restrictions on autono-

[56] Mill, *On Liberty*, 125, my emphasis.
[57] Ibid., 71–72.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press

mously executed acts that eradicate one's dispositional or global autonomy can be upheld under a policy of judicious strong paternalism.[58]

Actions taken against conduct such as consensual slavery might even be explained by nonpaternalistic reasons, as when the explanation appeals to the interests of others. A nonpaternalistic argument in support of justified encroachments upon a person's individuality can be derived from Mill's contention that when a person "disables himself, by conduct purely self-regarding, from the performance of some definite duty"[59] to the public, intervening steps may be taken. When the sovereign individual lays claim to conduct that, though directly and primarily affecting himself also indirectly or derivatively affects the larger sphere within which he interacts, the utility principle is implemented.

Does the person who happily relinquishes his capacity for self-determination, as does the compliant slave and as does the Taliban woman, negatively affect the larger environment within which the person acts? No definite duty to the public has been abridged, but it is arguable in these two examples that a social offense has been committed in collusion with the society in which the action is performed. If the offense is of a magnitude that imperils the social basis for productive interchange within the social, political, and economic realms, then the utility principle could be invoked to support injunctions against the action. Classifying self-imposed slavery as a social offense of this proportion involves analyzing the implications of self-imposed slavery for the social structure in which it occurs. The same is true of a society in which 50 percent of the adult population is relegated, willingly or not, to the status of the Taliban woman. The same is also true of a society that, citing the value of artistic license and individuality, refuses to compel socially unstable, schizophrenic artists to take their medication.

Of course, a utilitarian calculus could conclude that autonomy is of less value to a society than is the sovereignty of the individual who chooses nonautonomy; society might greatly benefit from voluntary slavery, Taliban women, and schizophrenic artists. But if an absence of autonomy does significantly damage the very fabric of a person's cultural and political environment, then the calculus will decide that the benefits to all from the interchange among free persons outweigh those that accrue from interchange among free persons and those who lack autonomy. The decision to utilize paternalistic measures must, I think, depend on the extent to which a society counts autonomy as important for the integrity

[58] Young would include arguably less egregious behaviors among the acts for which strong paternalistic interference is justified. He mentions "restrictions placed on professional boxing between grossly ill-matched boxers, . . . [and] voluntary participation in unnecessary, risky experiments," behaviors that are unquestionably imprudent but not clearly risks to, or sacrifices of, future autonomy. Young's claim lends itself to Berlin's worry discussed above in the text that one person's autonomy could become the subject of another's autocratic control. See Young, *Personal Autonomy*, 68.

[59] Mill, *On Liberty*, 70–71.

of its structure. In a society that bestows less social worth on autonomy, actions that diminish a person's autonomy, such as self-regarding slavery, might be targets for reproof or regret, but would not be regarded as social evils. Still, they might constitute a social evil if the longevity of the society could not be secured without autonomous persons.

VIII. Conclusion

The arguments that I have considered surrounding paternalism and interferences with autonomous behavior generally lend themselves to the following conclusions. First, autonomy *qua* individuality or individual freedom is not sacrosanct. There are times when we shall be permitted to override a person's right to noninterference, as when there are other, more exacting prerogatives that we want to protect.

I am not in a position to declare resolutely when the value of autonomy should be overridden, despite the cost to the individual. As Young notes, "[A]n individual's other values may come to assume at some time more importance for him than his autonomy, so the commitment we have to the value of autonomy is at best a defeasible commitment."[60]

There are many penumbral cases in which an individual's voluntary abdications of autonomy may fail to provide us with good or sufficient reason to intervene on a person's behalf. I am thinking of the lives embraced by persons in certain insular religious communities—the Amish, for example. Our respect for the good of value pluralism cautions restraint in such cases.

If we distinguish between autonomy and freedom, or liberty, then it may be that autonomy is one of those values for which freedom can be sacrificed. But autonomy, unlike freedom, is compatible with (and occasionally calls for) restrictions upon the parameters of individual behavior, or upon the sources of authority that govern an agent. The fact that it might be morally and legally incumbent upon us to caution others against their own behavior, to warn them of the punitive consequences that might follow their behavior, and to actually take steps to curtail their autonomy, does not mean that autonomy is not a valued state. This ideal remains intact, although uninstantiated in certain cases.

Even when a person's autonomy bears upon only those affairs in her life that are self-regarding, and is of fundamental importance to the pursuit and outcome of these affairs, autonomy is not inviolable. There are good arguments against respecting the autonomy of persons whose behavior undermines their self-governance, as in cases of consensual slavery. When these cases present themselves, interference may be justified and legitimate.

Philosophy, Bowling Green State University

[60] Young, *Personal Autonomy*, 73.

https://doi.org/10.1017/S0265052503202041 Published online by Cambridge University Press