https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

IDENTIFICATION, THE SELF, AND AUTONOMY

By Bernard Berofsky

I. Introduction

Autonomy, we suppose, is self-regulation or self-direction. There is a distinct idea that is easily confused with self-direction, namely, self-expression, self-fulfillment, or self-realization. (I do not mean to suggest that the latter three terms are all synonymous. But in this essay, whatever differences there are among them play no role, so I will use them interchangeably.) Although it will turn out paradoxically that autonomy is *neither* self-regulation *nor* self-realization, it is reasonable to suppose that the former is a superior candidate. My teacher of Indian religion, Dr. Subodh Roy, blind from birth, chose not to undergo an operation that would have made him sighted because he believed, perhaps rightly, that the ability to see would interfere with his religious quest. He thereby chose not to realize one of his fundamental human capacities, one whose cultivation has produced some of the finest fruits of civilization. Joseph Raz describes a case in which a man places his life in jeopardy by undertaking a trip to deliver medical aid to a group of people in a distant place. Since he will be unable to secure food for several days, he, in effect, subordinates one of his own basic needs or interests to a goal that he deems more important.[1] There is no reason to believe that, in refusing to express or realize a dimension of self, either Dr. Roy or Raz's philanthropist have failed to act autonomously.

Of course, there is also no reason to believe that these two men's actions do not "express the self" even if they suppress other facets of the self. What then renders self-regulation, suitably qualified, superior as the analysis of autonomy to self-realization?

II. Autonomy and Choice

To begin to address this question, consider the following example. Malcolm has recently become a father. He never chose to sire a child and never dreamed of assuming the responsibilities of fatherhood. To his initial surprise, he discovers himself taking on the role of a father with all of its trappings and entanglements. He finds that a powerful bond be-

[1] Joseph Raz, *The Morality of Freedom* (Oxford: Clarendon Press, 1986), 296.

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

tween him and his son is developing. Malcolm also finds that much of his time and energy are directed toward nurturing their relationship. His free time, his finances, and the range of his options are becoming drastically diminished. He is unable to resist certain sorts of entreaties and to do or even to contemplate actions that would place his child in harm's way. Of course, Malcolm cannot imagine doing or even choosing to do serious harm to his son; the prospect is sufficiently repellent as to paralyze his will. Indeed, were he given the opportunity to acquire power to harm his son, he would unhesitatingly refuse. Yet, in spite of the extensive diminution of his time, finances, and options, Malcolm also finds that he is thoroughly pleased about the sort of person whom he has become and the life that he is leading. Indeed, the very encumbrances that Malcolm experiences are essential to the expansiveness that his self undergoes through the relationship. He thoroughly endorses the limits on his will: a loving father is precisely one who cannot choose to harm his son.

Eventually, Malcolm is unable even to imagine living any other life. He comes now to be defined by the connections that have arisen by chance, not by choice. He is subject to what Harry G. Frankfurt calls "volitional necessity."[2]

Malcolm is not only thoroughly content with his life and its limitations, he is also pleased that he has reached this point in ways other than through acts of choice. Malcolm has fallen in love and to fall is precisely not to choose. In "falling" in love, one finds oneself bound to a person (or perhaps an ideology, a movement, or a job), and one is happy to be swept along by the rhythm and demands of the relationship. One prefers not to choose to engage in many of the activities, nor to undergo the emotions and patterns of thinking, that are intrinsic to an *act* of love. For if Malcolm had been presented with an opportunity to choose, he would have had the power to exercise a degree of control that would have undermined the nature and special quality of his involvement with his son.

Many of the most important activities in life are the very ones in which we happily lose a great deal of control to the other. When mathematician Andrew Wiles discovered a glitch in his proof of Fermat's Last Theorem, he did not want to will his way out of his deep frustrations. He could not decide to fix the proof. He was bound by the nature of the structures that he was studying and would have it no other way. He was glad that he was not the voluntarist God of Descartes, who has full control over mathematical relationships. Wiles had to and wanted to *discover* what was wrong with his proof.

Given that Malcolm identifies himself so completely with his life as a father, we may justly regard him as self-realized—his actions reflect the fundamental project that defines him. But he took no active regulatory

[2] Harry G. Frankfurt, "On the Necessity of Ideals," in Frankfurt, *Necessity, Volition, and Love* (Cambridge: Cambridge University Press, 1999), 108–16.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

role in the process of becoming a father. So why might one be inclined to describe Malcolm as an autonomous agent? Yes, Malcolm is a thoroughly fulfilled individual, preferring not to be different in any fundamental way, especially in those spheres in which his life is now largely governed by outside forces. But this is a description that could be provided for paradigmatic instances of *heteronomy*, for example, Sister Rossetta of the Heaven's Gate cult, who is a completely brainwashed individual under the domination of Do, the cult leader. She, too, identifies herself with her life and regards her actions as an expression of her underlying nature.

Malcolm *might* have reached this point autonomously. But to control the direction of his life, he would have had to *engage* in *independent, reflective decision-making*. For consider scenarios without reflective decision. Had Malcolm, having chosen to eschew fatherhood, been the victim of a deception that brought him grudgingly into this state, the fact that he eventually embraced his condition would appear to be a basis for regarding him as perhaps fortunate, but hardly autonomous. Or, had Malcolm actually reflected on the merits of fatherhood once it was upon him, the process might have taken place under the powerful influence of Malcolm's feelings for his son and their relationship, an influence that might have been sufficiently strong to overwhelm his capacity for independent judgment. So even if the heteronomous attachment of Sister Rossetta to Do is different from the attachment of Malcolm to his son, the latter relationship may incorporate a genuine loss of autonomy. An autonomous Malcolm, in contrast, would have been able to reflect on the wisdom of nourishing powerful impulses and emotions and respond accordingly. He would not have had to be constrained to endorse the ideal of fatherhood and might well have embraced it freely. Under *these* conditions, the reduction of control implicit in parenthood would no longer be a ground for reduced autonomy. But without having chosen the ideal of parental love at some key stages along the way, Malcolm will simply be contented and fulfilled, perhaps self-realized, but not autonomous. Thus, reflection on Malcolm's status reveals the superiority of an account in terms of self-regulation rather than self-realization.

But is it not unrealistic to suppose that we choose our ideals? Do they not draw us in by their compelling attractiveness? Yes. But the point is that beings who are under the sway of ideals in this way *may* fail to meet the conditions of autonomous decision-making—conditions that we assume can be spelled out, although we have yet to do so—and appeals to self-realization are not going to change this fact.

Two distinct conditions bearing on Malcolm's autonomy have been identified. In order to be autonomous Malcolm must (1) determine his relationship to his child, and (2) determine it on the basis of reflection that is sufficiently independent to be counted autonomous. If Malcolm fails to

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

choose at all (that is, if he fails the first condition), then he is content or fulfilled rather than autonomous. If he chooses under the undue influence of external forces (if he fails the second condition), then he is too dependent on others, and possibly lacking in sufficient self-control, to be counted autonomous.

Without choice, even in an ideal environment, all we would have are creatures revealing their natures. These might be fully contented animals, or human members of a tribe for whom individual expression has not yet arisen as an issue. Autonomy, as the Western ideal that it is, demands tension between an individual and the group, tension that is resolved in decision-making. One does not control one's affairs in the relevant sense by living a life that is totally fulfilling and conflict-free.

Nonhuman systems, for example, robots, are "autonomous" when they incorporate self-correcting mechanisms such as an internal thermostat, which could be compared to the mechanism within us that regulates internal bodily temperature. The robots' mechanisms operate automatically, that is, an environmental stimulus initiates action by the mechanism in such a way as to produce the "desired" outcome. We do not conceive our *selves* as operating in this automatic way and, therefore, we import into our conception of autonomy the demand that control be exercised at the conscious level. Self-regulation in the case of personal autonomy presupposes a conception of the self that links it to our status as conscious beings. I am not autonomous if I am acting on "autopilot," no matter how independent of the environment the relevant mechanism is and no matter how desirable or sophisticated the outcome happens to be—hence the demand that control in the human sphere be exercised through (conscious) decision-making.

Of course, we can imagine a creature, Jill, who is so fortunate that, at any stage of her life, she discovers, even after ideal reflection, that the values that she would like to guide the formation of specific motives happen, in all cases, to be the ones that do guide her. Jill need make no adjustments in her will. Hence, if she also discovers that she can never decide otherwise, this would not faze her. She may endorse, but she does not determine, her life. Is she autonomous?

We can raise the hypothetical question: Would Jill have had the power to bring her decisions into line with her deliberations were there a discrepancy? If we think of Jill as autonomous, then I believe that we are presupposing an affirmative answer to this question. Jill approves of her life, that is, her values and desires, as well as the specific decisions that she makes on the basis of them after a thorough, rational, and independent evaluation. She cannot be faulted for failing to exercise powers to change when there are no good reasons to exercise such powers. Although no actual practical decisions are called for, the judgment of autonomy introduces at least counterfactual choice for Jill. If Jill's decisions were ultimately those of a distinct coercive power, her acquiescence would

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

not confer autonomy on her.[3] Thus, to return to Malcolm, in saying that he must choose parental love in order to be autonomous, we do not mean that the state of love must be *created* by his choice. His endorsement suffices so long as it is rendered autonomously and is accompanied by the capacity and disposition to make adjustments in accordance with the results of reflection.

A harder line on the importance of choice—an insistence on actual determination by one's decision—would fail because this demand would lead to incoherence. It is a truism that we do not create ourselves ab initio. Once we reach the stage at which our first gropings for independence and self-mastery take place, a good deal of our nature is already in place. The wildest-eyed believer in free will asks only for free selection from options that, in reality, have already been drastically reduced by the forces of heredity and environment. Similarly, an account of autonomy for human beings, creatures who begin life as completely heteronomous, must permit a transition to autonomy that is achieved through critical appraisal of inculcated values and desires. An adult who is willing and able to make adjustments in accordance with an independent critical appraisal is not deficient in autonomy just because he approves of some elements that are already in place.[4]

Another consideration undermines the radical demand that all elements that enter into decision-making be the result of an act of choice by a mature agent. Decision-making cannot be undertaken unless the context permits certain elements to be fixed—not themselves the objects of deliberation—in order that practical reasoning about other elements may proceed. If everything is "up for grabs," then any particular choice is bound to be arbitrary and nonrational. One cannot hope to canvas all possible objections to, or be in certain possession of all relevant knowledge concerning, a course of action; the process of reason-giving must end, necessarily leaving us with a surd, at least in this context. We risk immobilization if we seek perfect rationality or perfect control.

Yet, we can say more to the lovers of "ideal autonomy." For example, even if some elements must be assumed so that practical reasoning may sensibly commence, one need not suppose that assumed matters may not themselves be the subject of evaluation in another context.[5] Moreover, the fact that much important decision-making depends on perspectives that have been provided to an agent in a process of molding that takes place

[3] We do not even have to suppose that a content Jill has the ability to change the direction of her life during her content phase. All we need to suppose for counterfactual choice is that she *would be able and disposed to change* if, through (autonomous) reflection, she came to disapprove of her life. Since the ability (and/or disposition) to change may arise only when Jill detects problems, we need not, therefore, even take a stand on the relation of autonomy to alternate possibilities.

[4] I am grateful to James Stacey Taylor for helpful suggestions on this point.

[5] See Bernard Berofsky, *Liberation from Self: A Theory of Personal Autonomy* (Cambridge: Cambridge University Press, 1995), 124–25.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

prior to the age of reason need not undermine the mature agent's autonomy as much as is ordinarily assumed. Although I cannot present arguments here, I believe that the arbitrariness of some perspectives may be removed through objective grounding (for example, inculcated beliefs may become known to be true), and the arbitrariness of those that cannot be so removed may be shown not to be intrinsically inferior to other possibilities.

Given that Malcolm exercises reflective and independent endorsement and has not reached the state of fatherhood through manipulation or coercion, we can then concede that his current inability to change does not undo his autonomy. For counterfactual control is the power to change *if* one detects a reason to do so. Discontent may be a necessary condition of the power to change, in which case Malcolm is currently powerless. He cannot alter his condition, but he could alter it if he wished to. There is, then, a real distinction between Malcolm and an agent whose autonomy is genuinely impaired, that is, one who cannot now or under hypothetical conditions, restore his life to a satisfactory state.

Even if we understand self-regulation as the proposed analysans of autonomy along the more moderate lines that I have just laid out, decision-making is a core notion in the account. Since decisions can be made under duress, or on the basis of values inculcated from without, or by one who is emotionally disturbed, it would be natural now to undertake the project of more carefully delineating the conditions of autonomous decision-making, for example, critical competence, independence, relevant knowledge, emotional stability, and the like. The picture of self-regulation that we would then be projecting is that of the regulation *of* decisions *by* various states and conditions that meet the requirements of autonomy. But then it would appear that the self, as this system of states and conditions, would be essentially passive. Although, like any set of states or conditions, the self can have causal efficacy, decisions would be the *result of* rather than *constitutive of* self-activity. In recognizing the importance of the difference between a decision that emanates from the self and one that really arises from external sources through manipulation, coercion, selective socialization, and the like, we are led to construe the self as a system that is distinct from the decisions that it occasionally explains.

This position appears to abandon the very starting point of the discussion. For I claimed above that autonomy or self-direction for human persons presupposes a self that is essentially a conscious decision-maker. I rejected the idea that mere explanation of action in terms of features of the self suffices as an account of autonomy; this idea would assimilate our autonomy to that of a self-adjusting mechanism. In an effort to reaffirm the original position, to insist on the fundamentally active nature of the self, some philosophers posit a special sort of self-constituting decision that disregards or bypasses the psychic system. In this paper, I wish to evaluate this point of view, which I shall call *Self-Constituting Decision Theory* or *SCDT*.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

III. Decision-Making and Self-Control

Before we turn directly to SCDT, we may still wonder why decision-making is so important to self-control. What does decision-making provide that makes us identify our *selves* with it? First of all, if the element of consciousness is crucial, then we should consider whether it is possible for decisions to take place without our realizing it. We do indeed say things like, "I see by the coordinated series of actions that you have taken that you have decided to bring me to ruin." May not the individual to whom this is addressed honestly claim that he is unaware that this is so? Yet may it not also be the case that the person has indeed been motivated by a desire to ruin the speaker? Of course, we do allow motivation by unconscious desire, but we have a problem with unconscious decision. Here, though, the allegation is that there has been a coordinated series of actions that reveals not only intelligence and purposiveness, but also something like a commitment to an outcome. Yet without conscious concurrence from the one who acts, we are loath to regard the agent as fully engaged or responsible.

Consciousness may be deemed important in connection with the ancient tradition in philosophy of identifying agency with rational agency. Thus, direction by an agent can only be understood in terms of guidance by reason. Raz says, "Our life is our own when it is under our control and that means when our various emotions, hopes, desires, intentions, and action are guided by reason."[6] The argument for the essential role of consciousness in agency would then rest on the thought that desires that operate unconsciously do so, not as reasons, but rather, as Donald Davidson says, as mental causes.[7]

Many philosophers have challenged the identification of autonomy with rational action on the grounds that true autonomy would permit the very rejection of rationality itself.[8] But even if this challenge is mistaken or incoherent, a direct rebuttal of the argument for the essential role of consciousness in agency is possible. The assumption that unconscious desires can operate only as causes, never as reasons, may be challenged. In many deliberations, there are considerations that ought to bear on a decision, which an ideally rational agent would take into account. Nomy Arpaly has argued that sometimes such considerations do play a role in

[6] Joseph Raz, *Engaging Reason: On the Theory of Value and Action* (Oxford: Oxford University Press, 1999), 1.

[7] See Donald Davidson, "Two Paradoxes of Irrationality," in Richard A. Wollheim and James Hopkins, eds., *Philosophical Essays on Freud* (Cambridge: Cambridge University Press, 1982), 289–305.

[8] See, for example, Robert Nozick, *Philosophical Explanations* (Cambridge, MA: Harvard University Press, 1981), 354; and Lawrence Crocker, *Positive Liberty: An Essay in Normative Political Philosophy* (The Hague: Martinus Nijhoff, 1980), 36–43.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

decision-making even if they do not enter into reflection.[9] Oftimes, circumstances prevent an agent from knowing at the time or even later what all of her reasons for action are. Sometimes these circumstances are benign. For example, we have to decide quickly and so we rely on our "instincts," that is, we do the sort of thing that we have done before without thinking.[10] In the past we formed a conscious intention on the basis of a consideration. We now confront a very similar situation and are unaware at this time that we are again invoking the same sort of reason to guide us. It just never enters our reflections, although later we may recognize the force of this reason. Sometimes, however, we actively resist the emergence in consciousness of earlier-established reasons, and they function undetected with a directness and a power that isolate them from the normal effects of a rational monologue or dialogue. Then, our reasons are not behaving as reasons, and we relegate them to the status of causes. Failing to respond as rational agents would respond to novel factors bearing on the reasonableness of acting on such reasons, we find that we are responding instead to a brute causal mechanism.

Since we can act from reasons in a rational manner even when they do not enter the conscious arena, why is self-determination for a rational self a matter of *conscious* decision-making? Rational decision-making on the unconscious level may at times occur, but it cannot become the norm. Many important decisions demand a delicate weighting of many factors and a large amount of information from disparate sources. We may have to take into account multiple moral demands and commitments, as well as a variety of nonmoral preferences with varying strengths and degrees of personal significance. Consciousness is then essential for the occasional complex deliberations that lead to the hard and important decisions that we face, especially those that take place when our lives are not going well on "autopilot." If serious reconsideration is called for, we cannot control our lives satisfactorily without conscious deliberation.

IV. Identification as Self-Constituting: SCDT

Since autonomous agents must take an active role in the direction of their lives, much attention has been paid in recent decades to the precise character that this direction should take. Since our own desires can enslave us just as much as an outside force can, the decisions that we regard as crucial are often ones regarding the desirability or acceptability of leading a life dominated by the desires that, in fact, happen to be guiding our actions.

[9] Nomy Arpaly, "On Acting Rationally against One's Best Judgment," *Ethics* 110, no. 3 (2000): 488–513.
[10] Ibid., 506–7.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

It is by now generally recognized that the simple fact that one might want to be moved by a certain desire is insufficient to ensure autonomy. As early discussions of Frankfurt's writings revealed, any worries about the insufficiency of an account of autonomy as the ability to do what one wants to do would extend to an account of autonomy as the ability to act upon those wants that one wants to act upon. Subsequent discussions, by Frankfurt and others, have introduced a variety of other sorts of considerations that are deemed relevant to the issue. Frankfurt himself has introduced the concepts of endorsement and identification as key, and he has striven to interpret these notions in such a way as to preserve the core meaning of autonomy as *self*-rule. Thus, the agent's reflective endorsement of, or identification with, certain of his (effective) desires and passions as the ones that he cares most about is constitutive of the very essence of the agent and thereby establishes his autonomy.[11] (To be sure, in facing certain conflicts, we can identify with desires in a weaker sense. When we need to establish priorities, we may accord certain desires a low level of priority without expelling them from the self.)

If identification, as autonomy-conferring, cannot be a merely passive state such as a (second-order) desire, then we had better not interpret it as a belief—equally passive—that a desire, process, or goal is valuable in one way or another. But the goal of reflection is often the formulation of such a belief. In moments of reflection Malcolm might want to know whether fatherhood is a good thing or, perhaps, whether it is a good thing for him at this stage in his life.

Gary Watson, in a well-known critique of Frankfurt's earlier works, has advanced the view that autonomy is indeed constituted by the coherence of one's motivational system with one's evaluational system, thereby claiming that the autonomy-conferring state might well be just a belief.[12] I am autonomous (or free) if my actions are guided by desires that I happen to believe are ones that it is good (worthy, desirable) to act upon.

But Frankfurt's earlier works insisted upon a more active, higher-order role for an agent. One must *do* something before a desire becomes one's own in the relevant sense. An endorsement is not just a belief; it is an act of appropriation, typically rendered on those desires that one thereby certifies as important, as expressive of what one cares about.

Whatever it is that one is doing to those desires that get endorsed, it is clear that this doctrine of Frankfurt's is very different from Watson's. For one thing, I can believe that some state of affairs is valuable without actively valuing it, that is, without being disposed to take steps to care for or nurture it. Watson came to acknowledge this later, observing that we may not identify with our own evaluations, and I have developed this

[11] Frankfurt, "On the Necessity of Ideals," 110–14.
[12] Gary Watson, "Free Agency," *Journal of Philosophy* 72, no. 8 (1975): 205–20.

theme elsewhere.[13] We can love the life that we are leading even if we must acknowledge that it lacks any "reaI" merit and that we would not recommend it to others. So there can be a breach between a person's evaluation on the one hand and his caring or identification on the other.

Watson regarded people like this as perverse. But I believe that this phenomenon is more common than that characterization would suggest. We agree that it is all too common for one's evaluations to clash with powerful desires. But when this happens, some thinkers hesitate to allow that a person might actually value or care about the object of these powerful desires. Some powerful desires are, to be sure, totally repudiated by a person victimized by them. But the view that one cannot value the object of those desires that one decides are unworthy of endorsement rests upon an unreasonable standard for the conversion of a desire into a (personal) value, and I shall argue this point below. I know that a gambling life has little to recommend it, but I cannot help loving it, and I would choose it all over again.

Independently of this complexity, heteronomy is not constituted just by the discord between one's motivational and evaluational systems. Suppose that I acknowledge that I ought to care for my ailing mother, but I really do not much care for her or for the project of caring for her. Harmony can be restored by providing extrinsic incentives for taking care of my mother, by having discovered, say, that she is thinking of rewriting her will. Here, I now want to care for my mother and believe that a good son is obliged to do so, but it is not this belief that is moving me to act in the right way. If evaluations are important to autonomy, then they must cohere with a mode of motivation that is determined by these evaluations. It will not do just to be moved to act in accordance with the demand that is expressed in the evaluation. I can even care for my mother and believe that I should care for my mother because children ought to assume these responsibilities, yet not be motivated in the right way. I might, for example, harbor powerful incestuous desires that are doing the causal work. The grounds on which one normally judges that one should take care of one's mother do not include the desire to nurture a romantic relationship with her. If evaluations are important to autonomy, then they should mesh with or even determine the reasons that one is motivated.

Suppose, then, that I am moved to assume the burden of caring for my mother for the right reasons (children have an obligation to their parents) without really caring for her. Consequently, I regard my responsibility as a burden, as a chore that I must perform. It would be a mistake to conclude that the assumption of responsibility is not one of my values, for I am committed to acting on it because I believe that I should be so committed. But if my allegiance to morality is accompanied by a sense of

[13] Gary Watson, "Free Action and Free Will," *Mind* 96, no. 382 (1987): 145–72. See Berofsky, *Liberation from Self*, chap. 5.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

alienation rather than a sense of fulfillment, then I may fail to identify with this powerful moral streak in my personality. Although I regard myself as a good person, I view my conscience as a fetter.[14] In this case, insofar as my moral belief induces moral motivation leading to morally praiseworthy behavior, we have even more than just harmony between evaluational and motivational systems. My motivations have been brought into line because of my evaluations. Although it strikes me that in this case I am neither free nor autonomous, I emerge as autonomous on a harmony-style account.

Advocates of a harmony-style account cannot evade this result by appealing to the presence of desires that conflict with the effective ones, for conflict is ubiquitous in those who are judged to be free or autonomous. (After an unwilling smoker finally conquers his addiction, he may harbor urges to smoke that he unambivalently keeps under control.) Thus, we see the merit of an alternative account of autonomy such as Frankfurt's. The smoker identifies with his desire to quit, but I do not identify with my relationship with my mother nor the moral principle that moves me to act. The smoker cares about or values quitting, but I do not care about or value (in the right way) my mother or morality. We need to go beyond evaluations.

If identification or endorsement is not a cognitive state, then what is it? As I stated above, Frankfurt thinks of it as appropriation, as an instance of metaphysical hubris. "The person, in making a decision by which he identifies with a desire, *constitutes himself*."[15] This is mine; that is not; I have spoken. At other times, the language of authority rather than appropriation is used. Through this act, I am authorizing, that is, establishing authorhood or ownership. But what exactly does this mean?

Frankfurt emphasizes the volitional or conative dimension of this act to distinguish it both from the cognitive dimension that we have been discussing, as well as from the affective dimension. Endorsement is not the having of any special sort of feeling, even though it may typically be accompanied by certain feelings. Caring, for Frankfurt, is also a conative state. "That a person cares about or that he loves something has less to do with how things make him feel, or with his opinions about them, than with the more or less stable motivational structures that shape his preferences and that guide and limit his conduct."[16] A variant of this idea is Michael Bratman's version, according to which identification with a desire or the act that establishes "ownership" of the desire is a policy in

[14] One of the problems with Susan Wolf's theory of freedom (*Freedom within Reason* [New York: Oxford University Press, 1990]) is that I count as a free agent on her view because I am acting for the right reasons. See Berofsky, *Liberation from Self*, 72–75.

[15] Harry G. Frankfurt, "The Importance of What We Care About," in Frankfurt, *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press, 1988), 170.

[16] Frankfurt, "Autonomy, Necessity, and Love," in Frankfurt, *Necessity, Volition, and Love*, 129.

favor of treating that desire as reason-providing in motivationally effective practical reasoning.[17] It is a forward-looking commitment to treat certain facts about oneself as justifying reasons for action.

If, as I have argued, identification is not simply wanting, and not simply approving, and not simply believing, and not simply feeling a certain way, yet it is a mental act directed to the will, then it would indeed have to be a decision, or a commitment, or the adoption of a plan or policy. Although, until now, we have associated such acts with caring or love, the fact of caring must be distinguished from the act of identification itself. Just as there can be a breach between evaluation and either caring or identification, so can there be cases of identification without caring *so long as identification is defined in terms of mental action alone*. The case of the reluctant son indicates that one can be committed to act in certain ways and to regard one's (weak) desires to do certain acts, such as helping one's mom, as good reasons to do so, without caring for the object of one's attention. The problem goes deep. One may not in general care about being moral in spite of a commitment to a moral path, as the above example demonstrated. And the commitment itself may involve constraints on future choices and decisions without implicating the phenomenon of caring: for example, one is convinced through moral reflection that desires to be good ought to be cultivated.

Although a conceptual distinction can be made, it may be argued that identification without caring is psychologically impossible. As human beings, we do not form commitments, adopt policies, or make identifications unless we care about the goal that is necessarily driving us. Frankfurt characterizes love (a species of caring) as conceptually linked to the desires that explain the mental actions. "What a person loves helps to determine the choices that he makes and the actions that he is eager or unwilling to perform."[18]

The claim of psychological impossibility can be challenged. There are people who become resigned to a way of life that they believe they are trapped in, and they are weary of continued struggle. Although they care for little, they do not want to rock the boat, and so they commit themselves to staying the course. Yes, a person of this sort cares a little about stability, or loyalty, or getting through the day without disaster, but he is basically acting out of world-weariness. It is a mistake to suppose that a person to whom life has been sufficiently cruel so as to render him incapable of love or passion is thereby unable to order his will.[19] Even if his

[17] Michael E. Bratman, "Identification, Decision, and Treating as a Reason," in Bratman, *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999), 197–98.

[18] Frankfurt, "Autonomy, Necessity, and Love," 129. Does the reference to eagerness import affect into the account of love?

[19] If identification can be severed from care, it can a fortiori be severed from love. Love is a stronger emotion because one can care about something that one does not love. See Eleonore Stump, "Persons: Identification and Freedom," *Philosophical Topics* 24, no. 2 (1996): 183–214.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

will is moved by a desire for stability, it is not at all moved *because* he cares about the object, the future of which he is committed to take into account in his deliberations. If identification can be severed from care and can be severed from judgment and affect, as it was above, then can identification alone ground autonomous agency?

In Bratman's version of this idea, an agent must be satisfied with his decision in the sense that it does not conflict with competing decisions and does not, therefore, create a divided will.[20] Bratman concludes that, should an addict be resigned to his condition, he may be free of a divided will, in which case he *has* identified with his addiction.

Bratman objects to Frankfurt's more recent interpretations of identification just because Frankfurt has abandoned the position that decisions are important. Frankfurt replaces the act of decision with the (passive) state of wholeheartedness, characterized negatively in terms of the absence of any discontent or inclination to change.[21] Bratman and Frankfurt both demand satisfaction, but Bratman insists on the interpretation of identification as a bona fide act.

In spite of the centrality of the concept of identification in recent philosophical accounts of autonomy, the term is pretty much a philosopher's term, and I, therefore, have no pretheoretic intuitions to guide me in resolving these disputes among proponents of SCDT. My subject is autonomy, however, and my pretheoretic intuitions regarding it incline me strongly to believe that Sister Rossetta lacks autonomy even if she "identifies" in either Bratman's or Frankfurt's sense of this term. I turn now to a defense of this position.

V. The Limits of Identification

Identification can be reflective, impulsive, or somewhere in between. No one would suppose that reflection undermines autonomy.[22] Indeed, Frankfurt insists that reflection prevents wantonhood[23] and that wholehearted identification arises only as a consequence of an appraisal of one's psychic condition.[24] So, let us suppose that identification (here, with O) is reflective. Even so, an agent can lack information of different sorts that would have led her not to identify with O had she possessed it. An agent may be misinformed about (1) whether O is really satisfying in the long run, (2) the motives leading her to identify with O, and (3) what it is she is really identifying with.

[20] Bratman, "Identification, Decision, and Treating as a Reason," 200–201.
[21] Harry G. Frankfurt, "The Faintest Passion," in Frankfurt, *Necessity, Volition, and Love*, 102–5.
[22] Although one can be too reflective at times, obviously reflection is often important and not in general antithetical to autonomy.
[23] Wantonhood is the state of not caring about which desire moves us to action.
[24] Frankfurt, "The Faintest Passion," 105.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

The third possibility is controversial. But even Frankfurt concedes that people may be mistaken "about what is moving them in their choices and in their actions."[25] They may also be "mistaken *concerning what they love*."[26] And he makes similar remarks about care. If we cannot help identifying with what we love, then we must sometimes identify with something that we are not aware we are identifying with. A woman can rightly say, as the song says, that her lover is really in love, not with her, but with love (or loving).[27] A scholar can believe herself to be devoted to the pursuit of truth, unaware that she is devoted really to the pursuit of acclaim by her peers.

These cases highlight unclarities in the concept of the object of identification. In each such case, we can try to avoid having to think of the agent as ignorant of the object by relocating the ignorance at the level of reasons. Harry knowingly loves Sally, but does not realize that the reason is that Sally is merely an instance of a love relationship with Harry. Wilma loves her sociological research, but does not realize that the reason is that the research is a means to acclaim. In effect, the relocation strategy would expand the second type of ignorance (ignorance of motive) to protect against the third (ignorance of object). But the strategy fails because what a person loves determines the policies and commitments that one undertakes, including policies concerning the structure of one's motivations. Yet in each of these cases, different commitments are associated with what we have called different reasons. When the excitement of loving dissipates, Harry will seek love elsewhere, thereby revealing that Sally was not really the object of his love at all. And when Wilma fails to receive praise from her colleagues, she will abandon her scholarship, thereby revealing—at least to others, if not herself—that the scholarship itself is not what interested her in the first place. What we have called the reasons must enter into the characterization of the object of identification if we want to preserve the idea that the object determines one's policies vis-à-vis the direction of one's motivations and the specific future choices that one makes.

These results imply that the conscious product of reflection whose goal is the formation of a fundamental commitment is not a case of self-formation ex nihilo. The process is, at least in part, one of discovery, and it can go awry in ordinary creatures in several ways.

Suppose that we now simply append relevant knowledge in all its forms to wholehearted commitment or the like in order to define autonomous agency.[28] This step is not just a simple emendation of SCDT. For

[25] Frankfurt, "Autonomy, Necessity, and Love," 130.

[26] Ibid.

[27] Richard Rodgers (music) and Lorenz Hart (lyrics), "Falling in Love with Love," from the 1938 Rodgers and Hart musical, *The Boys from Syracuse*.

[28] Perhaps this is Frankfurt's intention when he talks of understanding and appreciating one's condition as requirements of wholeheartedness. See Frankfurt, "The Faintest Passion," 105.

we are significantly closer now to the more conventional approach to autonomy suggested in my discussion of the Malcolm example. We insisted there that mental action is indeed crucial to autonomy; Malcolm must decide to continue the path that he is on. But this decision is autonomous only if it meets various constraints such as knowledge, independence, mental health, and so forth.

Before we conclude that, by themselves, acts of identification just cannot live up to the demands that are placed upon them—that SCDT fails—we will ponder the possibility of a last-ditch effort to pull the mental act of identification up by its own bootstraps so that it can do all of the work of self-formation on its own.

Here is the argument. Knowledge be damned; if Selma identifies wholeheartedly with acclaim, structures her motivations accordingly, and, unlike Wilma, is not hampered in the pursuit of her goals, then she is an autonomous agent.

After all, why should knowledge matter in this context? If Fred spends his whole life driven by a desire to reach the top of the corporate ladder, manages to climb many rungs, and faces no serious obstacles on his ascent, but dies before realizing that success would have felt empty and would have led to a reappraisal, the outcome of which would have been Fred's repudiation of his core value, why should we deny that Fred led an autonomous life? Fred wholeheartedly identified with success in terms of climbing the corporate ladder and, as a consequence, led a life in accordance with this value. We all perform value experiments. The failure of this experiment is a testimony to the shallowness of Fred's choice, but not necessarily to his heteronomy.

Although we can disagree about the delineation of conditions that we wish to impose on identification, we cannot be sanguine about the wholesale abandonment of reasonable conditions of autonomous decision-making, including knowledge. Let us expand the picture of Selma, the woman who identifies wholeheartedly with acclaim, structures her motivations accordingly, and proceeds unimpeded to fulfill her goals. Like Wilma, Selma is unaware that she would repudiate acclaim were she to achieve it, for she would find it totally unfulfilling. She is even unaware that she has identified with and seeks acclaim. (This possibility has been conceded by the friends of identification.) She believes that she loves scholarship for its own sake. She lacks self-knowledge, for she would be deeply ashamed of herself were she to discover the truth, and her unconscious fear of feeling this way leads her to bury the knowledge of her motivations. Acclaim is very important to her because she is deeply jealous of her sister, upon whom Selma believes all parental love was heaped. (Actually, Selma's parents loved her equally, but, needing to wallow in self-pity, she does not want to recognize this fact.) This jealousy reveals itself in many ways—in feelings, verbal and nonverbal behavior, and dreams—but Selma continues to misinterpret these manifestations in

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

order to preserve a self-image that she is more comfortable with. If Selma were to discover these facts about herself, she would change her life dramatically. Not only would she abandon the pursuit of her scholarly reputation, but also she would abandon the mistaken object of identification, scholarship. After all, she never really cared about it. And she would change careers altogether. But she never makes these discoveries.

Acts of identification, commitments, and policies do not arise in a vacuum. My intuitive response to Selma—that she completely lacks autonomy—reveals to me that the contours of the self cannot be confined to the immediate environment of acts of identification.[29]

If my conclusion requires more support, then we may compound the story along lines that have become familiar in the critical literature. Not only may the agent lack transparency in the above ways, but also failure to know "the what" and/or "the why" of identification may be explicable in terms of extrinsic origins that are obviously autonomy-nullifying. We may, for example, trot out the philosopher's "demon neurologist," that all-powerful, evil physician who manipulates a subject's brain without regard for the subject's mental life, causing her to have the mental states (for example, acts of identification) that the demon neurologist desires her to have. A similar scenario is that of Do, the cult leader who manipulates Sister Rossetta, causing her to identify with his goals. It would be ludicrous to suppose that this act of identification is an instance of self-creation, hence autonomy. Sister Rossetta identifies with the goals of Do as the result of an act in which she is basically enslaved by him.

It is difficult to imagine a more obvious truism than that we act for reasons and on the basis of beliefs, but the thought that we must link the self to conscious decision-making in order to form a conception of our own autonomy must be modified by this recognition, or we will be defining autonomy not for us, but for God. We must observe again that many of the desires and beliefs that we possess, feel thoroughly comfortable with, and act from, evolved from our childhood, a time during which we lacked control over our destiny. No plausible account of autonomy can relegate *all* of these desires and beliefs to limbo until an agent performs an act of identification, commitment, or endorsement. The project of distinguishing desires that are genuinely mine from the rest requires us to define the right sort of independence from external sources. Desires can arise through interpersonal relationships in an autonomy-preserving way (for example, benign advice) and in an autonomy-nullifying way (for example, brainwashing). There is a variety of intermediate forms—manipulation, temptation, and coercion, for example—each of which comes in degrees, and this variety makes it impossible to draw a sharp boundary

[29] I do not wish to deny the originative role of decision-making. We need not be pushed and pulled by our motives. Also, the fact that identification does not take place in a vacuum does not mean that the act itself cannot generate desires that have great motivating power and steer us along important lines.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

between desires that arise from within and those that are imposed from without.

Also, it has become increasingly clear that much of our behavior, our motives, and our capacities is under the control of our self-conceptions. Much of our conceptualization and interpretation of the environment and our own actions takes place in accord with our need to maintain a self-conception or self-image, or with our need to enhance this self-image, or at least with a need to render our behavior intelligible and coherent, given this self-conception. Cognitive dissonance research and self-perception studies reveal that we distort the subjects of interpretation in order to fulfill these goals, whether or not the self-image or self-conception that we import to interpretation is accurate. In other words, we frequently distort our picture of ourselves and the world in order to maintain a distorted self-image!

In the face of all of these possibilities, proponents of SCDT are burying their heads in the sand when they propose an account of autonomy that would be limited to a pristine act of identification. It makes sense to distinguish desires that belong to the self from those that do not, *antecedent to* acts of identification, in which case we must incorporate into an account of autonomous identification or decision-making a requirement of motivational autonomy, namely, the demand that the motives that are driving me already be genuinely mine.

VI. Alienation

If identification lacks the metaphysical clout that some have ascribed to it, so does alienation. One cannot, in other words, expel a desire from the self just through a pristine act of repudiation, no matter how emphatic. Indeed, as we saw above, if repudiation is constituted by a value judgment, then we cannot even be sure that a sincere repudiation will entail that an object of a desire is of no value to us any longer.

Suppose, for example, that Harry A's whole life has been driven by an overwhelming desire to exterminate Armenians. Having met a nice Armenian, Harry A begins to reflect on his values and eventually repudiates his Armenian project, although he continues to find himself driven by it. He suffers, therefore, from the familiar conflict between motivation and evaluation. But conflicts of this kind are quite complex, as we have seen. For example, Harry A may not harbor just a negative moral evaluation of his conduct. He may also feel utterly alienated from what he now thinks of as a dangerous compulsion. (In calling the impulse a "compulsion," I do not mean to imply that the agent cannot refrain from acting on it, although, obviously, it is at least very difficult for him not to succumb to it. In this way, we preserve the possibility of choices to refrain and do not, therefore, automatically rule out the possibility of autonomy in this sphere.)

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

Although Harry A continues his nefarious activities in spite of sincere efforts at reform, he no longer identifies himself with them, and he no longer sees them as flowing from him. Moreover, suppose that this facet of Harry A's life does not permeate other facets all that much. He does not dwell on it, and his feelings, behavior, and motivational system operate independently of the compulsion to kill Armenians. Every so often these murderous impulses come upon him, and it is difficult for him to refrain from acting upon them. Once the episodes end, he picks up his life as if the compulsion had never existed. Even though the action emanating from the impulse is intentional, the intention is not a component of a pattern of intentional behavior. The action is not part of a project whose history is one component of the agent's sense of his continuing identity as a person. Rather, the impulse and resultant intention rupture his sense of an ongoing life that is, to some extent, under his control. If we wish to characterize this example as a case of expulsion from the self of particular desires, the ground for so doing is not just repudiation, but, rather, the fact that the impulses are cut off from the rest of Harry A's life and are perceived by him as lacking in personal value. These impulses no longer make sense in relation to his other interests and values, and he may be undertaking steps to eliminate these urges. I would concede that Harry A is heteronomous (in this sphere of his life) in a pretty straightforward way.

Other scenarios that incorporate projects that we continue in spite of repudiation are more difficult to characterize. For example, in the case of Harry B, the anti-Armenian compulsion plays a significant role in other aspects of his life. He dwells on it, he has nightmares, he spends much of his time in therapy, and his behavior toward others is adversely affected. Here, Harry B is not just repudiating the behavior; he would prefer to be rid of the ancillary business as well. He thinks of himself as victimized by his compulsion. If he were to be rid of it, there would be no adverse consequences to his psyche. The greater significance of the compulsion is only causal in character and is important to Harry B only in the way that a deadly disease is.

Harry C is in the same situation as Harry B, except that the compulsion does play a key role in Harry C's psychic organization. The behavior serves a basic protective role for Harry C, such that success at its abolition would release fears and unresolved conflicts that would be personally devastating. Harry C is not aware of any of this, but he would choose not to change if he became aware—the pain would be too great. Although he repudiates the compulsion because it lacks any intrinsic personal value, his conscious desire to be rid of it is based on ignorance. The compulsion serves an indispensable personal function for Harry C.

Harry D is like Harry C except that Harry D experiences enormous personal satisfaction when he acts on his impulses and he does not, therefore, wish to be different. He dwells on his behavior, but only be-

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

cause he likes to savor the experience. The very danger of the enterprise is invigorating. The image of himself as an exterminator is deeply embedded in his psyche, and he perceives his own motivations and the world in terms of this image. He finds it depressing that his conscience has issued a negative report. He realizes that there is little to be said in favor of his project beyond the personal fulfillment that he derives from it. He knows that he should stop and is, therefore, taking preliminary steps to do so. But, as in the example above of my failure to identify with my obligation to care for my mother, Harry D does not identify with his evaluations. Although Harry D is unaware of the importance of his compulsion to his very sanity, he "fortunately" is not sufficiently interested in ending his project, anyway. Harry D seems to me to be a clear case in which repudiation does not add up to expulsion.

VII. The Role of Reflection

The inclination to permit a move from repudiation to expulsion is the consequence of conferring undue metaphysical weight on reflection. I argued above for the necessity of conscious deliberation to autonomy. But deliberation as a prelude to decision-making can take place without a self-conscious appraisal of the forces at play in the deliberative process. As we are using the term here, *reflection* refers to the deeper, second-order process in which an agent steps back to undertake such an appraisal. One must concede the significance to autonomy of reflective identification *as a type*. Mature beings think before they act and must occasionally think deeply so as to raise the possibility of adjusting their wills as well as their behavior, depending on the results of their reflective activity. In characterizing individual selves, however, we must guard against the tendency to give excessive due to the specific products of reflection. These outcomes cannot take metaphysical priority over the sum of psychological facts that establish both the centrality of the genocidal impulse to Harry D's psychological economy and the personal, positive significance that it possesses. Yes, I experience occasional, weird promptings that play no significant role in my life, that I am taking steps to eliminate, and that I utterly repudiate. That these sorts of isolated impulses may be treated as "external" to myself does not entitle me to eject fundamental desires that are deeply satisfying, that play a central role in the explanation of my verbal and nonverbal behavior and experience, and the dislodging of which would be deeply disruptive.

If Harry A is heteronomous because he is moved by "external" forces, is Harry D autonomous because he is moved by "internal" forces? Although I pointed out above that conflict is a fact of life even for autonomous agents, Harry D's conflict and ambivalence may be quite deep. He at times exercises choice on behalf of his highly satisfying genocidal project and at other times, under the sway of his moral reflections, he seriously

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

ponders steps that he might take to suppress and to eliminate his genocidal urges.[30]

In order to produce a plausible answer to the above and similar hard questions, we must revisit the conceptual starting point of this discussion. In seeking an account of autonomy, we look to the phenomenon of conscious decision-making as crucial. We then confront the need to expand the sphere of relevance in order to incorporate the causal neighborhood in which we find the bases for decision-making. And now we are forced to acknowledge that this neighborhood also houses some pretty seedy characters: (1) beliefs, the falsity of which is self-induced through the need to protect the self and one's self-conception; and (2) stated reasons for action that disguise genuine ones. That these genuine reasons fail to make their way to consciousness is also explained in terms of needs that often remain unconscious, such as the need to produce consistency in, or to maintain or enhance, one's self-image. These elements are not external to us in the way in which Do's interests are external to those of Sister Rossetta. And, at least for Harry D, they are not external in the sense in which completely repudiated impulses and addictions may be so counted.

At a weekly meeting of Exterminators Anonymous, Harry D speaks truly when he says, "My name is Harry and I am an exterminator." This is true, not in a trivial sense, but rather, in the sense that his behavior, feelings, fantasies, speech, and deepest concerns are driven by the desire to eliminate Armenians. We should not forget that in spite of the negative judgment that he casts upon his feelings, he actually does find himself hating Armenians and relishing his actions against them. Nor should we forget that the abandonment of this project would produce in Harry D profound personal upheaval and great distress. Moreover, his feelings of shame make no sense if he is not really an exterminator. Although Harry repudiates himself, how can one deny that his project expresses one of his core values?

VIII. Reflection, Identification, and Autonomy

In examining the role of reflection for the characters whom we have called Harry, we reach the paradoxical result that self-*direction* is compatible with heteronomy! To see this, let us introduce Harry E, who is similar to Harry D, except that Harry E is equipped with many qualities that are important to autonomy. Unlike Harry D, Harry E is not guilty of self-delusion. He knows exactly why he does what he does and why it is important to him; he has all relevant knowledge, including knowledge of the psychologically harmful results of extinguishing his deadly desires.

[30] Harry D is not a clear case of the sort of ambivalence that Frankfurt describes (see Frankfurt, "The Faintest Passion"), because Harry D does, in the end, side with his genocidal impulses in spite of serious reservations.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press

Although none of us creates our desires from scratch, we hope that some sort of line can be drawn between those that arise "naturally" and those that are the consequence of manipulation, coercion, mental illness, etc. Suppose, then, that Harry E's desire to kill Armenians does not arise in one of these "external" ways. The individual decisions that he has made along the way have been made autonomously: they have been informed, dispassionate, and driven by desires that were not the consequence of manipulation, coercion, etc. Perhaps Harry E had misgivings early on. But when he reaches the stage of mature reflection, he concludes that, taking all considerations into account, moral and nonmoral, he would prefer not to be making the decisions and commitments that he has been making on behalf of the project of extermination. In other words, he would choose to be a different sort of person if he could. His ambivalence is not so deep that he cannot ultimately side with his moral nature.

"Self-regulation" cannot be a synonym of "autonomy," etymology notwithstanding, if it requires that the system that is doing the regulation incorporate the deepest or most important components of the self. For Harry E's actions *are* under the control of forces that represent what is most important to him; he *is* choosing his life. Moreover, given that he is independent, informed, uncoerced, free of mental illness, etc., Harry E fulfills the requirements of autonomy as well as anyone else does. There is no serious case to be made that his life is directed from without. But his reflective condemnation of himself establishes his heteronomy *even if the reflection itself meets the requirements of autonomy*!

What has emerged is that considerations bearing on the autonomous character of a decision arise at two points. We first look at the etiology of decision. If deliberations do not bear on decision, if, for example, one is a slave, then approval or reflective endorsement by the slave of her state does not undo her debilitating condition. Reflection is a sham if one lacks the power to make one's own decisions. Second, we look at reflection, which introduces another set of autonomy requirements. For we must suppose that Harry E's reflective repudiation of his project is itself autonomous, that is, independent, uncoerced, etc. Although conflict is a fact of life even for autonomous individuals, such agents cannot abide a deep schism between their reflective and their active natures.

IX. Conclusion

We have tried, on the one hand, to avoid an irrational reverence toward the rational and, on the other, to pay the right sort of homage to our reflective natures. Even though reflection can be faulty in many ways, its ideal dictates must govern the will of an autonomous agent. A sincere, informed, stable, mentally healthy, independent, and rational evaluation must guide an autonomous agent such that its dictates fundamentally

mesh with his *nonreflective nature*.[31] If one's intellect has not achieved a sufficient foothold in the self, what one professes to be important upon ideal reflection may not yet in fact be so. But one cannot be counted autonomous unless one is actually guided by values and principles endorsed by autonomous reflection. The demands of autonomy will go unrealized if professed ideals do not find sufficient mooring in the self. So Harry E's autonomous rejection of himself renders him heteronomous in spite of the fact that he is expressing himself and is regulating his actions by his choices.

The fact that Harry E sides with morality upon reflection is incidental. One can just as easily concoct a case of self-regulation without autonomy in which a moral agent repudiates her morality for the sake of self-interest, although she finds her morality to be a fetter that she is unable to cast off. Imagine a woman who has been given a moral education that meets the most rigorous standards of wise persons in her community. She grows up to be a model of propriety until, as a young adult, she finds herself developing powerful immoral proclivities. Although she never veers from the path of morality, she eventually comes to view her moral nature as an impediment to being the sort of person whom she would ideally prefer to be. I cannot argue here that she may ultimately form an *autonomous* (and rational) judgment that, all things considered, she would choose to scuttle the moral life in favor of an immoral form of self-expression. But if this is possible, then we have a case like Harry E, a person who is self-directed, yet heteronomous.

I will not proceed to produce even in outline a statement of *necessary and* sufficient conditions of autonomy, that is, to consider examples in which reflective endorsement is short of ideal and/or the conflict between it and self-governance is not as deep. Rather, my goals have been (1) to return the phenomena of identification and endorsement to the context in which they belong, namely, elements of a self whose broader contours must be introduced into an account of autonomy, thereby refuting SCDT; and (2) to concede the importance of reflection as key to our understanding of autonomy, without assigning to it the metaphysical significance implicit in the (mistaken) view that its ideal form must coincide with self-determination.

Philosophy, Columbia University

[31] Meshing, rather than governing, may suffice. When we reach the age of reflection, a great deal is already in place and much of it is untouchable. As I said above in the text, no one would be autonomous if we demanded that ideal reflection actually be the source of one's values, desires, and fundamental choices.

https://doi.org/10.1017/S0265052503202089 Published online by Cambridge University Press