

7

“SHE BUT NOT HERSELF”—SELF-ALIENATION AS INTERNAL DIVISION

Our life as we lead it is just our life, except that some elements in it seem like intruders, interpolators. Some thoughts we have, emotions we feel, some of our beliefs, desires, and actions are experienced as not really ours. It is as if we lost control, as if we were taken over, possessed by a force which is not us.

—JOSEPH RAZ, “WHEN WE ARE OURSELVES”

IN THIS CHAPTER I DEAL with cases in which one experiences one’s own desires and impulses as alien, cases in which one sees oneself as dominated by desires that one has, but as if from an alien power, or cases in which one’s own behavior leads one to feel like a stranger to oneself. These are situations in which one wants to say “that can’t be me,” but in which, at the same time, one is oddly incapable of rejecting the behavior one experiences as alien or of dismissing the desires one feels so distant from. In this sense being alienated from oneself means not being able to identify with oneself or with what one wants and does, which seems to be not really “part of our story” and not really to belong to our own life. How are we to understand that? And, conversely, when are desires really *our own* and when are we really *ourselves*? My claim is that in these forms of self-alienation there is a certain way in which one is not accessible to oneself in one’s own desires and that this phenomenon can be explained without appealing to an authentic “core self.”

I will again proceed by first (1) elaborating the phenomenon with the help of an example in order next to (2) bring out the characteristics that make *self-alienation* a plausible interpretation of the described situation. From this arise two sets of questions. The first has to do with the internal structure that characterizes a division of this type within a person’s own will, the second

with the standard in relation to which certain of our desires are able to claim authority and others not. I address both sets of questions in sections (3) and (4) in conjunction with an extended discussion of Harry Frankfurt's account of the person. I conclude (5) that his model is incapable of resolving the problem of how our desires can acquire authority if what we are interested in is overcoming processes of internal division in the name of emancipation. Finally, (6) the view of self-alienation as practical inaccessibility to oneself will yield some clues as to how the dilemma elaborated here can be resolved.

(1) THE GIGGLING FEMINIST

H., a self-professed, reflective feminist of strong convictions, catches herself over and over again communicating with her lover like a silly, giggling adolescent girl. She rejects such forms of feminine coquetry as unemancipated, as the mannerisms of a "little girl." She has long understood that the idea that women must present themselves as cute, petite, and harmless in order to be attractive is the projection of a world dominated by men. Yet, as she discovers to her irritation, she constantly falls back into these patterns of behavior against her will. She experiences her own behavior, so starkly in contrast to her convictions, her self-conception, and her life plan, as contradictory and as not really part of herself. It triggers in her a feeling of disconcertedness when she sees herself behave in such a manner: "That can't be me." Formulated somewhat dramatically, it is as if in her giggling something were speaking through her that is not herself.

(2) DEMARCATING THE PHENOMENON AND DEFINING ITS CHARACTERISTICS

One can describe the discrepancy that makes the situation depicted here an experience of self-alienation as follows: H. cannot *identify* with her impulse to giggle and with the desires she suspects lie behind it. The talk of feeling alien in relation to herself indicates that more is going on than (or something different from) a mere rejection of certain behaviors. She desires and does things that do not "fit" or belong to her, things that at the same time she has no influence over. She is *internally divided* insofar as she seems to be split into two parts that do not stand in a coherent or meaningful relation to each

other. Similarly to the academic in chapter 1, she feels herself to be a "helpless . . . bystander to the forces that move [her]."¹ And yet what is at issue here is nevertheless (to return to the Heidegger quotation from part 1) "a power that is she herself." How precisely, though, can we explain that a person can do and desire things that at the same time do not belong to her? Who is alien to whom here? When, and having which desires, would she be *herself*?

In what follows I elaborate the features that make the conflict I describe a problem of self-alienation.

1. The Significance of Desires. The behaviors of her own that she rejects are not merely insubstantial inconsistencies that occur at the periphery of her personality and have no vital importance for her. The part of herself that she experiences as alien stands at the center of her personality and is of great significance to it.² Her behavior, then, is no mere quirk, not merely a vestige of previously learned behavior that is inconsequential for her. (She is not someone whose feminism resides merely at the surface of her personality and who would almost be relieved to be able to discard its strong demands; she is, rather, a woman whose identity is deeply informed by her feminist convictions, who owes much to them, and who in many other respects successfully leads an emancipated life.) One can imagine that in reflecting on these issues H. discovers that her pattern of giggling is intertwined with deep-seated desires and thoughts, for example, with the fact that her idea of romantic relationships corresponds far less to the picture of a symmetric relationship between equals than she could admit. Her giggling, she discovers, is an expression of the need she feels to be protected, as little as that fits with her otherwise self-confident manner. (It would be excessive to regard behaviors that are really involuntary and trivial—without explicable meaning³—as symptoms of an internal division.)

2. The Incompatibility of Desires. To speak of an internal division further presupposes that the opposed sets of desires are mutually exclusive or at least that pursuing one of them stands in significant tension with pursuing the other. H.'s desires are incompatible insofar as they—at least for her—suggest relationships and forms of life that are mutually incompatible: whereas, on the one hand, she wants to be an independent woman, she also, on the other hand, longs for a love relationship in which she is dependent but protected. This is a case, then, in which the inconsistency between those desires becomes an explicit problem, at least in the protagonist's own self-conception. (Otherwise there would be no reason for a stronger reaction to her own behavior than

mild bemusement.) The two attitudes sketched earlier are understood by our protagonist as an opposition between an emancipated and an unemancipated way of life. As such they are mutually exclusive. To borrow a distinction of Charles Taylor's: they contradict each other *qualitatively* and are therefore incompatible in a different and stronger sense than that they merely cannot be realized simultaneously.⁴ In contrast to this, for example, the desire she occasionally has to play competitive sports conflicts with the rest of her life only because of time constraints. This desire, even if it were to remain unfulfillable, is not an alien element in her economy of desires, and—as long as it is not reinterpreted and integrated into the framework of more fundamental life decisions—she will not have the impression of being alienated from herself regardless of whether the desire is fulfilled or not.

3. The Inauthenticity of These Desires. That the existence of certain desires is interpreted as alienating implies, further, a very specific attitude to those desires and a corresponding understanding of their nature. When we regard our own desires and behavior as alien to us, we understand them as desires that we do not *truly* (authentically) have. When H. experiences her desire for protection or subordination to a man as an alien part of herself, she distances herself from it not merely in the sense that she rejects it; she understands the desires she interprets as alien as being *not really* her own; they are not her authentic desires. The assertion “they don't belong to me” is more than just a confused way of saying “I don't want that.” What is implied, rather, is that they are *not genuine*. These desires, one could say, masquerade as her desires. This calls into question—places under suspicion or expresses a reservation about—the authority of the desires she in fact has. Talk of self-alienation—this is the important point—presupposes the possibility of criticizing desires, which takes the form of doubting their authenticity.

This feature also distinguishes the conflict in which our protagonist finds herself from inner ambivalence.⁵ Someone is ambivalent when she stands between two of her own desires; in this case both sides represent desires—even if qualitatively incompatible in the sense explained earlier—that are equally her own and can therefore each claim an *equal right* to authority. So understood, a conflict of ambivalence is a *tragic* conflict.⁶ (In this sense, for example, one can stand ambivalently between two lovers or be ambivalent about deciding for a life with or without a child.) In contrast to this, the potential for conflict in *inauthentic* or *alien* desires, or in desires one understands as such, resides in the fact that one rejects a desire and therefore cannot identify with it but still cannot be rid of it. H. would not say of herself that she stands between the

desire for emancipation and the desire not to be emancipated or that she has both desires simultaneously. She rejects her desires not to be emancipated. Whereas a conflict of ambivalence is due merely to the fact that one cannot realize the two desires simultaneously, an inauthentic desire is one that one *truly* does not even have.⁷ In the one case we must simply decide; in the other we must find out what we *really* want. This presupposes not merely that one of the two desires is more important but that one of them corresponds more to oneself. The question then is how we can make sense of these presuppositions—how one can distinguish real from unreal, authentic from inauthentic desires if both are desires one in fact has. Is it so clear whether her giggling or her normally self-confident manner better represents H.'s authentic desires? When is she really herself: when she no longer giggles or when she no longer distances herself from her giggling?

4. Self-conceptions. Understood in this way, conflicts of this type concern what is often called a person's self-conception or identity. What is at issue for the young woman who is in internal conflict over her girlish behavior is clearly who she *is* and how she *conceives of* herself. It is not for her merely a question of whether she should decide in favor of one form of life or another (and about the consequences this would have in each case) but rather of what her actions would make of her and of how she could understand herself in them. The role of interpretation and self-interpretation is crucial here: H.'s behavior is not contradictory as such. It is contradictory insofar as it contradicts her feminist self-conception. Without entering into such reflexive relations oneself—not merely *doing* this or that but *conceiving of* oneself in this or that way while doing it—the entire phenomenon of self-alienation would be inexplicable. One can understand one's behavior or one's desires as an alien element of oneself only because one has an implicit or explicit conception of what belongs or should belong to one, because one can integrate certain things into one's self-conception and not others.

5. Freedom and Emancipation. If self-alienation in the present case means being driven by desires that one in some sense does not really have and thus becoming someone one really is not, then one is not really free when controlled by such desires. As Raymond Geuss writes: "Someone is 'free' in the full sense only if he does what he really wants to do, that is, only if he acts out of a genuine, authentic, or real desire. The authenticity of the desires that motivate action is an essential component of freedom."⁸ Unmasking inauthentic desires has in this respect an emancipatory significance, if processes of emancipation involve more than casting off foreign domination and oppression and

also involve emancipation in relation to the alien power “that we ourselves are.”⁹ Such processes of emancipation typically include a complex process in which someone wants to become different from what she is and at the same time interprets this as corresponding better to who she is—as being “freed into being herself,” as the promising formula would have it.

Now in order to examine the relation between real desires, the self, and self-conceptions, as well as to articulate what justification there can be for speaking of self-alienation here, two sets of questions must be analyzed in relation to the example sketched above:

First, how can we understand the claim that alienating, inauthentic desires are those that are alien to the person and that nevertheless mysteriously compel her in certain ways? How exactly are we to understand the fact that one has such desires, but does not really (authentically) have them, that one supposedly—and paradoxically—at once has and does not have them?

Second, what kind of criteria can there be for establishing which of two conflicting desires is one's own in this strong sense? What authorizes desires? What makes them our own or alien? Bound up with this is the question of under what conditions a self-image or self-conception is appropriate or fitting.

I will discuss these questions by critically examining the views of Harry Frankfurt, whose theory of the person, as we will see, is relevant in various respects to reconstructing a theory of alienation.

(3) THE ALIEN CHARACTER OF ONE'S OWN DESIRES

The first problem—how one's own desires can be alien to oneself and how we should understand the distancing from desires that is bound up with this—can be investigated with the help of the model of the will developed by Harry Frankfurt in his paper “Freedom of the Will and the Concept of a Person.”¹⁰

According to Frankfurt, the defining characteristic of persons resides in a specific structural feature of their wills, namely, that they can relate evaluatively to their own desires. Frankfurt elaborates this claim with the help of a hierarchical model: persons relate to their “first order desires” by means of “second order volitions.” A second order volition is a desire to have or not to have a first order desire. So, for example, one can have a second order volition not to give in to one's first order desire for a cigarette. What makes someone

a person is not merely having desires but being able to take a position with respect to them and to distance oneself from them.

Besides wanting and choosing and being moved *to do* this or that, men may also want to have (or not to have) certain desires and motives. They are capable of wanting to be different, in their preferences and purposes, from what they are.¹¹

This conception of the person has several important implications for our problem:

First, not every desire that one has is *one's own* merely because one has it. It is a positive relation to one's desires—which Frankfurt later calls identification—that makes them one's own in a meaningful sense. Just as one can identify with one's desires in order to make them one's own, one can also be alien with respect to them (be alienated from them) insofar as one does not identify with them.

For Frankfurt the paradigm case of this kind of alienation is the unwilling addict. Someone who is an unwilling addict is dominated by a first order desire (to take drugs) that contradicts her own second order volition (not to take drugs or, more precisely, not to give in to her first order desire for drugs). This is a case, then, in which first and second order desires diverge. In a way that is structurally similar to our feminist, the unwilling addict is internally divided—at odds with herself—because she is driven by desires she does not *really* have, which is to say, by desires she does not have at the level of her second order volitions. She is, as it were, unable to turn her second order volition into effective action. Exactly like our feminist, she experiences her powerlessness in the face of the continued presence of an unwanted desire as alienating. In this sense, her continuing desire to take drugs is one she “does not really have” because she does not affirm it on the second order level. This enables us to explain nonparadoxically how desires can be at once alien and one's own: an alien desire is one that I in fact have—on the level of first order desires—but with which I cannot identify—on the level of second order volitions. Calling a person's desires *alien* does not mean that she does not *have* them; it means, rather, that she has not *made them her own*. Here, too, the situation is not one of merely conflicting desires (as discussed above in conflicts of ambivalence) but rather a rejection of a desire that one experiences as an alien element. Allying oneself with the second order volition, one has, so to speak, taken sides against the lower level desire. And the specifically alien character of these desires is not that one is not *aware of* them—the unwilling addict is

aware of her desire for drugs all too well—but rather that one does not make them one's own. Understood in this way, a condition of self-alienation is one in which a person has failed in some way to bring her first and second order desires into agreement.¹²

Second, the account of the person that emerges from Frankfurt's discussion is of great interest for the problem of alienation. For Frankfurt the decisive criterion for the ascription of personhood is the capacity to develop second order volitions. What distinguishes us from other living beings is not the capacity for rationality (or language or other human characteristics) but the structure of our will. If being able to distance oneself from one's own desires—being able to take a critical or affirmative stance to them—is the distinctive characteristic of persons, it follows that the possibility of a divergence between what one factually *is* and one's *project* (for oneself) is constitutive of personhood. Put differently, a person is not determined by the “raw material” of her desires but rather by how she gives form to them (and along with them herself). Authentic desires, then, are not natural or given but rather higher-level, shaped desires; being oneself or being in agreement with oneself is not a natural or immediate condition but a higher-level process, the result of which Hegel refers to as the “purification of the drives.”¹³

This becomes clear in the contrast between an unwilling addict and a “wanton.” A wanton is an addict who will-lessly gives in to her addiction and allows herself to be determined exclusively by her first order desires without taking a position to them in the form of a second order volition. What distinguishes the wanton from the unwilling addict is not the result—both in the end succumb to their desire for drugs—but rather the fact that the former does not reflexively relate to her desires. For a wanton, who lacks the capacity to take a reflexive position to her desires—and, so, to evaluate her own desires—every desire that moves her is immediately *hers*. She does not distance herself from her desires and hence knows no internal division. For Frankfurt it is for precisely this reason that a wanton does not really have her own desires and is not really a person. She lacks the capacity to distance herself (for example, from her desires)—and therefore lacks the feature that for Frankfurt constitutes the core structure of personhood.¹⁴ If the wanton, therefore, is *one with herself*, it is at the price of the essential feature of personhood. Conversely, it follows that real “being one” with one's desires is achieved only on the higher level of reflective will formation—precisely when first order and second order desires agree. The wanton, who does not give form to the raw material of her desires, is not, as one might think, herself to a particularly strong degree; she is

not authentic in the sense of being immediate or spontaneous. Rather, she is not capable of real authenticity because she lacks personhood, a higher-level structure that first makes it possible for something to be authentic or inauthentic. One might say that because she does not have a relation to herself she also cannot be in agreement with herself. For that reason she is not alienated from herself, but *has lost* her self.

What follows from Frankfurt's account of the will for our example and our two problems? The model of second order volitions allows us to give a nonparadoxical answer to the first question as to how something can simultaneously be an alien desire and one's own—as to how H. can simultaneously have and not have her desires, can both desire and not desire something. Her second order volition to be emancipated is directed against her first order desire to behave coquettishly, like "a little girl."

How, though, does this help us to solve the second problem regarding the authorization and disowning of desires? The position that emerges from Frankfurt's account of the person with respect to the authority of desires at first appears to be simple: *real willing* is not found in lower-level, immediate desires; what one really wills, according to Frankfurt, is what higher-level, reflexive volitions aim at. In the case of H. this means that what is decisive for her as a person is not the giggling that she cannot hold back but rather the desires that lead to emancipated behavior. This answers the question of what makes one's desires one's own or of what *authorizes* them: the higher-level volition is the authority that makes a desire one's own; it has the power to define what is one's own and what is alien and what belongs to a person or does not. And not being alienated means bringing one's desires into agreement with this higher-level volition. This follows simply from the formal structure of the will as Frankfurt defines it.

(4) THE AUTHORIZATION OF DESIRES

It is not so simple, however, to solve the second problem raised by our example. For it is possible also to question—and it is even probable that H. sometimes asks herself this, too—whether the second order volition for emancipation that we have taken to be authoritative in fact corresponds to H., whether it is "in keeping" with who she is. Could it not be that the uncontrollable impulse against which she defends herself expresses something that she does not have at her command but that is nevertheless undeniably a part of her personality?

Copyright © 2014, Columbia University Press. All rights reserved.

We see here that the question of the authenticity of desires—the question of what really belongs and corresponds to us—cannot be answered with the formal, hierarchical structure that Frankfurt's model offers us. An unsympathetic misogynist, for example, might claim that when H. distances herself from her giggling and her fantasy of protection this shows that her feminism is a masquerade. In truth, according to the claim, it is precisely when she wants to rely on a man's instinct to protect that she is “fully a woman” and when she giggles she is really herself. How, then, are we to decide in which position our young feminist is really herself and what corresponds to, or is more appropriate for, her? How are we to decide whether the tension between self-image and reality that reveals itself in H's conflict is due to an illusory self-image or simply to the difficulties that have to be overcome in the course of emancipation? What makes desires or impulses authentic? What authorizes them as really *mine*?

The difficulty before us is as follows: if one were to ask H. why she does not consider her desire for protection to be a real desire of hers, the answer “because I don't want to have it” would be just as inadequate as the assumption that this desire is her own simply because it is there. A “barren assurance” (Hegel) is not enough.¹⁵ Everyone knows that one is not simply and immediately what one would like to be and that one does not merely decide to follow whatever desires one has. The fact that H. prefers to be emancipated is—without further elaboration—not decisive. It would be just as mistaken to think that she is herself precisely in the uncontrollable giggling that spontaneously breaks out in her despite the constraints her feminist self-image imposes on her. For if, to follow Frankfurt, it is constitutive of being a person that we can want to be different from what we are (precisely because we can desire that our desires be different from what they are), then authentic desires are always evaluated and formed. This means that the question of their authenticity is a question about the *appropriateness* of this evaluation. It is a question of which of our second order volitions are appropriate or on what basis we identify with some of our desires and distance ourselves from others. But this means it is a question of their *justification*. When H. distances herself from her disposition to behave coquettishly, she makes a certain claim of authority—and according to the considerations just discussed, she must do this. But what legitimizes this claim? How can she be sure that her second order volition is her real, authentic desire, the one that is “her own”? The question here is the following: what truly authorizes a second order volition? What makes it one's own?

Frankfurt's model does not allow us to answer this question. In order to come closer to an answer, it is necessary to examine more precisely the nature of the process of identification that underlies will formation, as Frankfurt has

done in (among other places) his essay "Identification and Externality."¹⁶ I will argue that even his reflections here (and in other essays) cannot solve the problem; yet examining the reasons for this failure will bring further dimensions of our problem into view.

IDENTIFYING WITH ONE'S OWN DESIRES

What exactly does it mean that we can *identify* with one desire but not with another? On what basis are desires alien or our own? What makes them belong to me or not? And to what do we appeal when we distance ourselves from expressions of emotion? Identifying with something means regarding it as belonging to oneself, as a part of oneself. Conversely, desires, feelings, impulses, and passions with which one cannot identify can be seen as *external*. Frankfurt addresses these issues in discussing a case of a violent outburst of temper. When one apologizes for such an outburst—"I don't know how that could have happened; somehow in this moment I wasn't myself"—one means to show that what was expressed in the outburst does not correspond to what one really feels, but one does this without denying that in that moment one was in fact under the sway of a feeling of rage.

One could understand this distancing such that—as in the case of the feminist—we do not identify with these impulses inasmuch as they cannot be integrated into our self-image: "we regard them as being in some manner incoherent with our preferred conception of ourselves, which we suppose captures what we are more truly than mere undistilled description."¹⁷ It is obvious that this does not solve our problem. As we saw before, the fact that a certain behavior "does not fit" with us, that it does not agree with our preferred self-image, is not sufficient to explain its being external or alien. For, in the end, one's self-ideal can be just as alien or inappropriate as the impulse that does not fit with it.¹⁸ Cannot persistent impulses that run counter to one's self-image even serve to uncover illusory aspects of it?

Even if one is not to be identified with all that spontaneously bursts out of oneself, one also cannot reasonably claim simply to be identical with what one *would like* to be. In any case, simply appealing to what one would like to be, which is nothing more than a declaration of intention, cannot do any justificatory work. The question whether an impulse or a desire is alien or one's own (is internal or external) cannot be determined solely by the person's attitude toward it. As Frankfurt himself says: "It is fundamentally misguided to suggest that a passion's externality is entailed by the person's disapproval of it, or that its internality is entailed by his approval."¹⁹ If this were so, then

instead of saying “That’s not me” it would be more appropriate and less confusing to say “I don’t want to be like that.” This, however, would leave the phenomenon we are attempting to explain untouched. If calling something alien has any explanatory value, and the process of identification is to have a higher authority than that of mere wishing, we cannot simply make a desire, disposition, or feeling into something alien by declaring it to be such, just as a certain ideal image of ourselves cannot become *our authentic* identity by declaring it to be so.²⁰

This is a question of interpretive sovereignty, of the conditions under which it can be said of a certain interpretation that it really captures who we *are*. The problem, which Frankfurt himself recognizes, is that the authority and status of even second order volitions can be called into question; they can themselves be alien or external: “Attitudes towards passions are as susceptible to externality as are passions themselves. This precludes explication of the concepts of internality and externality by appealing merely to the notion of orders of attitudes.”²¹ What we need, then, is a criterion for the internality or externality of desires that goes beyond merely subjective attitudes, a description of what it means to identify with something, a criterion according to which identification means more and is grounded in something other than merely a positive view or attitude. If identification is to *authorize*, it must involve something beyond a mere subjective wishing, something more compelling or decisive. And it is not easy to see where this is supposed to come from if what is at issue is the status of one’s own desires, not an objective account of what one *ought* to desire or an appeal to what content “real desires” can have. (What is at issue in the case of H. is not whether it is right to be a feminist or not but whether she really is or wants to be one, whether she *corresponds to* or *fails to be herself* in what she does; it is not a question, then, of being in agreement with what is objectively good or right but of being in agreement with herself.) The problem of finding a criterion presents itself with such urgency because we also lack any essentialist criteria for being in agreement with oneself; that is, we cannot claim knowledge of a human essence that would make such a judgment possible.

BETWEEN DECISIONISM AND THE CORE MODEL

It is instructive to see why Frankfurt cannot solve this problem within the framework of his model. Frankfurt remains undecided between two opposing models of explanation, which, translated into Heideggerian terminology, one could characterize as *resoluteness* and *thrownness*. The first emphasizes the

active elements of the identification process (making decisions), whereas the second emphasizes the passive (fateful) elements. Both models, however, run into difficulties: the first cannot justify the authority of desires; the second falls back into an essentialism that makes it impossible to capture aspirations for self-transformation and emancipation as they appear in our protagonist's initial feelings of alienation.

(a) Resoluteness. Sometimes (in the essay we have discussed) Frankfurt characterizes the process of identification in which one relates to one's own desires and passions as a kind of decision with a fundamentally active character: "it appears to be by making a particular kind of decision that the relation of the person to his passions is established."²² However, Frankfurt has a difficult time characterizing the specific nature of this kind of decision such that it has the binding force and necessity it is supposed to have: "In any event, the nature of decision is very obscure."²³ There are good reasons for this difficulty. As explained above, the authorization of desires cannot be a merely voluntaristic process. The decision in question must be determined by something that—in a way that is indeed difficult to grasp—comes from a "deeper," "weightier," or better founded stance. A further point is also crucial for my way of posing the question: the resoluteness model cannot really explain the possibility of self-alienation as it appears in our case. If identification is conceived of decisionistically—if we make our desires our own by means of a simple decision—it is possible to fall into a condition of irresoluteness that threatens our identity. This dissolution of identity, however, is not equivalent to self-alienation. According to this model, every decision (as long as it is sufficiently firm) results in an "agreement with self" that cannot be further questioned or evaluated. The question "What do I really want?" is then no longer meaningful. It can refer only to the intensity or resoluteness with which one wills. The question "Am I really resolute?" cannot be meaningfully posed or, at best, only rhetorically. The problem of a desire's authority in the sense of its legitimacy has no place here. Applied to H., who questions her identification with her second order volition for emancipation: in the decisionist version of Frankfurt's position the problem cannot be posed such that there could be a correct answer to the question of what she should identify with. From this perspective the only problem is that she asks this question at all, that she is not sufficiently resolute in leaning toward one of her desires. She cannot fail to be herself in deciding for one side or the other; her identity is threatened merely by the fact that she is undecided.

What, then, authorizes the authenticator? According to this model, the authenticator authorizes itself. A desire is authorized by the fact that one has decided for it. It is this decisiveness that stops the threat of a regress of higher-level desires. This is where we hit bedrock and “the spade is turned.” For the purpose of answering our question, however, the spade hits bedrock too soon. The authority of desires, as it arises for the problem of self-alienation, can be neither questioned nor justified on this model.

(b) Thrownness. The second version of Frankfurt’s account, which I associate with the term *thrownness*, explains the problem of identification differently. Even if identification with one’s own desires still has an essentially active character here, Frankfurt emphasizes the passive dimension—the intractability (*Unverfügbarkeit*) of one’s deepest commitments and identifications—in speaking of “ideals” and “volitional necessities.” The account of the person that Frankfurt develops over time in his writings attempts to do justice to the intuition that persons are characterized by a dimension of intractability. Persons are beings who relate to their desires by shaping them; their will “carves a path” for itself through the desires and needs that confront them on the level of their first order desires. At the same time, however—and Frankfurt emphasizes this more and more as his work develops—this dimension of the will should not be misunderstood voluntaristically. The possibility of relating evaluatively to one’s own desires does not mean that a person’s will is completely unbound or uncommitted: a person cannot will just anything; she is not free to redesign her will from scratch.

Frankfurt is concerned here with the will’s limits, with the limits of what one is free to will. He even goes so far as to claim that it is precisely these limits that make up the character of a person: “The boundaries of his will define his shape as a person.”²⁴ Conversely, someone who could will everything would have no identity as a person: “Since nothing is necessary to him, there is nothing that he can be said essentially to be.”²⁵ Frankfurt gets to the heart of this topic with his concept of volitional necessities: there are things we cannot help but will and, on the other hand, things we cannot will. This in turn depends on what we are really committed to, what we really “care about,” what is unalterably important to us: “Our essential natures, as individuals, are constituted, accordingly, by what we cannot help caring about. The necessities of love, and their relative order or intensity, define our volitional boundaries. They mark our volitional limits, and thus they delineate our shapes as persons.”²⁶ Someone who was capable of everything, who had no volitional

limits, would, according to Frankfurt, have no identity. By the same token, following one's volitional limits means being in agreement with oneself.

Once again Frankfurt elaborates his point with a striking example: the case of a woman who has decided to give her child up for adoption but who realizes in the decisive moment that she simply *cannot*.²⁷ It is significant that Frankfurt does not interpret this decision in what would seem to be the most natural way, as a victory of first order desires over second order volitions—as a triumph, for instance, of spontaneous emotion over reason. For if the woman cannot make the second order volition to give the child up “her own” even though she has it, then the force that prevents her from doing so in the decisive moment operates on a level that, according to Frankfurt's hierarchical model of desires, is higher than that of her second order desires.²⁸ Volitional necessities, then—sometimes Frankfurt speaks of ideals, but this can lead to misleading associations—are the authority that decides which second order volitions a person can embrace. Thus, one cannot follow just any second order volition, not because pure, unevaluated desires conflict with it, but because one *cannot* follow some second order volitions when considered from the higher authority of one's volitional necessities.

The crucial point here is the following: according to Frankfurt's account, the limits set by volitional necessities, although they place constraints on what we can do, do not constitute *compulsion* in the conventional sense. If, as has already been said earlier, these limits make up our identity, then they represent something like our deepest fundamental commitments, and these are ineluctable because they are what constitute us as a person. For Frankfurt, someone who questions or denies her volitional necessities betrays her identity. According to this view, the mother who wants to give up her child is threatened with the loss of her identity—and from this threat comes the necessity she cannot escape, the force she yields to when she finally decides to keep her child instead. This necessity, however—and this is Frankfurt's main point—is not compulsion since conforming to it means remaining in or coming into *agreement with oneself*.

In order to understand this account more precisely and to be able to evaluate its implications for the problem of the authorization of desires, I would like to summarize briefly the implications of Frankfurt's claims for the problem of alienation we have just examined:

1. On the one hand, self-alienation can be understood, with Frankfurt, as being “delivered over to” our own desires and longings. (We could call this

Copyright © 2014, Columbia University Press. All rights reserved.

“first order” alienation.) These desires can take on an overwhelming power that presents itself as a “force alien to ourselves.” This is not due to their irresistible character alone: “It is because we do not *identify* ourselves with them and do not want them to move us.”²⁹

2. These feelings and passions are the raw material that we relate to evaluatively or with respect to which we form our will.

Whether a person identifies himself with these passions, or whether they occur as alien forces that remain outside the boundaries of his volitional identity, depends upon what he himself wants his will to be.³⁰

Hence the volitional attitudes on this level, in contrast to unformed *first-order desires*, can be shaped and structured and are wholly at our command: they are “entirely up to” us. A crucial implication of this account is the distinction between *power* and *authority*. Passions, according to this account, have *volitional power* but no *volitional authority*. Frankfurt elaborates: “In fact, the passions do not really make any *claims* on us at all. . . . Their effectiveness in moving us is entirely a matter of sheer brute force.”³¹

3. What we do not freely have at our command, in contrast, is our *volitional nature*, the deep structure of our will itself. On the level of volitional necessities we are determined; here it is not “entirely up to us” how we determine our will; our volitional nature determines us. Yet our volitional necessities determine us in a different sense from that in which passions or first-order desires do: they compel us, one could say, not as *alien* powers but rather *to be ourselves*. They are not a brute force because they are not an external power but rather the power of what we really want or really *are*. “It is an element of his established volitional nature and hence of his identity as a person.”³² For this reason Frankfurt can claim in his adoption example that the mother experiences the limitation of her will—her “not being able to”—as a kind of liberation. Self-alienation, then, means acting against one’s volitional nature. Hence the mother who wants to give up her child has formed a second order volition that conflicts with her volitional nature. If she acted in accordance with this second order volition, she would alienate herself—a “second order” alienation. This means that it would run counter to what constitutes her as a person; it would undermine the *conditions of her identity*. Self-alienation on this level consists, then, in not being in agreement with one’s own person, with what constitutes oneself as a person.

The assumption of a volitional nature appears, then, to solve the problem of finding a criterion for authentic desires and their authorization that I have

raised in conjunction with the theme of self-alienation. The standard for the appropriateness or inappropriateness of identifying with a desire is our volitional nature; our desires—our *real* desires—are authorized in relation to it. In what follows, however, I will explain why this, too, fails to solve the problem raised in our initial example.

CRITIQUE OF THE MODEL OF THROWNNESS

The talk of attempting to find out *who one is* already raises a suspicion, namely, that this concept takes us back, though in an interesting and methodologically sophisticated way, to an essentialist core model of the self of the type I criticized in the introduction (even if it is not vulnerable to some of the criticisms I raised there). In part 3, chapter 9 I will take a more detailed look at the account of the person or the conception of the self at issue here. For the moment I am interested only in the practical implications that follow from this solution and its model of the person.

Volitional necessities, as Frankfurt describes them, are not only what ineluctably makes up the identity of a person; they are also not subject to questioning or critique—they cannot and need not be justified. They are factual and contingent. According to Frankfurt's conception, asking the mother who cannot give up her child "Why can't you do it?" is no longer a meaningful question. She could answer only by saying "Because that's how it is." This is not due only to the emotional strength of her commitment; it is due to the structure of personhood as Frankfurt conceives of it. The intractable commitments in question are *conditions* of the possibility of her own identity; they are what first make her into a person, into someone who can develop further desires for this or that. If giving up the child undermines the mother's identity, it is no longer possible to ask meaningfully whether keeping the child would destroy her plans for the future. There would then be no basis for such plans. This has an important implication for our inquiry: there is no place from which it is possible to question or criticize the influences and formative processes that have constituted this identity. A volitional nature, though volitional, is in the end *nature* and therefore not something one has at one's command.

One could indeed suspect—contrary to Frankfurt's interpretation—that the mother in question (one assumes a situation in which her life with the child would be very difficult) is not in a position to let her second order volition determine her action because she is too deeply stuck within traditional

ideas of maternal love. These ideas, one could say, prevent her from making a self-determined decision that accords with her plans for the future and her life plan more generally. There would, then, be two possibilities: a) a calling into question of oneself (of the person one has become); and b) a tension between identity and self-determination, which is precisely what Frankfurt excludes with his idea of being “liberated into oneself.”

If Frankfurt's account excludes such questioning and critique, however, then *self-transformation* is no longer a possibility: every radical self-transformation manifested in an abrogation of one's volitional necessities would represent a loss of self. If we ask, from Frankfurt's perspective, how desires can become authentic, the answer can only be through a process of comparing and adjusting one's desires to one's volitional necessities. Asking oneself what one really wants means, then, becoming clear about one's volitional nature, which is accepted as “untractable”—as given and not subject to questioning or alteration. This has implications for the possibility of emancipation as well as for the emancipatory nature of the question concerning the authenticity of desires that interests me here.

It should now be clearer why Frankfurt's model is unsatisfactory as a solution to the problem of authorization: Frankfurt underestimates the role of reflection, justification, and evaluation that accompanies the process of identifying with one's own desires (or that at least must potentially be able to accompany it) if these desires are to be able to become one's own in a robust sense. In this respect his two apparently opposed models meet in a common point: what the idea of volitional necessities shares with the decisionistic model is that neither has room for a process of reflection and evaluation that could guide our taking a position in relation to our own desires. Either we make decisions about what we identify with “just because”—as an ultimately unjustified and unquestionable choice—or there is nothing at all for us to decide, and all we have to do is carry out what our identity as defined by our volitional necessities requires of us. In both cases—when Frankfurt understands the process of evaluating our own desires decisionistically and when he completely brackets out the element of decision—we do not really decide *ourselves*. The unintended implication of such a conception is that our desires remain in a certain respect “raw facts” (Charles Taylor), even in the case of higher level desires. Insofar as both ultimately take place without reflection, there is no question of “forming” one's own desires.³³ This, however, undermines Frankfurt's own intentions: if the process of interpreting and evaluating desires is bracketed out, it is difficult to distinguish the situation of

the mother who *cannot* give up her child out of volitional necessity from that of the addict who *cannot* resist taking drugs, even though, as we have seen, it is precisely this difference that Frankfurt's account aims to explain. Although we find in Frankfurt the suggestion that (unreflected or first order) passions have volitional *power* but no volitional *authority*—that they exercise power over us but possess no authority—he has failed to make clear how precisely the authority of such claims can be justified (if the talk of authority is to have normative significance). To summarize my objection: although for Frankfurt desires become one's own only when one *appropriates* them as such, this process of appropriation or identification can be properly understood only if we are able to distinguish between appropriate and inappropriate identifications, that is, between successful and failed processes of appropriation. In the following section I will again take up the problem of identifying with one's own desires with a view toward its relevance for understanding how processes of emancipation are possible.

(5) BEING ONESELF AND EMANCIPATION

The emancipatory character of questioning the authenticity of desires (or of suspecting them to be inauthentic) is due to the fact that, in inquiring into the appropriateness of given desires and attitudes, one presupposes the possibility of criticizing them and thereby of calling oneself, as one is and has become, into question: one can want to become other than one is. Critically examining our desires and dispositions—the doubt “Does this impulse, this desire *really belong to me?*”—can make it possible to make our life more decisively our own and to “move more freely” within it. Applied to Frankfurt's adoption example, the question of whether the mother who cannot give up her child has allowed herself to be trapped in patterns of socialization that make her unfree is part of such critical examination. H.'s case, too, has this general structure. Whereas Frankfurt would be concerned only with determining which of her two sets of desires corresponded to her volitional nature—is she “fully a woman” or a feminist?—for H. herself the problem poses itself differently: she aspires, in cases of doubt, to question critically even her volitional nature. For Frankfurt she would be alienated from herself if, acting contrary to her volitional nature, she attempted to be something other than what she “is”—independent, perhaps—while still longing to be protected, whereas H. would consider herself to be alienated precisely if she “blindly” followed this nature.

And whereas for Frankfurt she would be *authentically* herself only if she were free of tension and without ambivalence, the alternative view would hold that the struggle against her deepest attachments and formative influences cannot take place without such tensions. According to the one position, such an attempt at emancipation is a condition of not being alienated; according to the other, it is a threat to the unity of the self.

THE DILEMMA OF EMANCIPATION

Now it may be that in certain situations one would do well to recognize that a particular way of life simply does not correspond to who one is; and, of course, it could in principle be the case that the standards H. sets for herself are too demanding. On the other hand, would we not be suspicious if H. were suddenly to reveal to us that the whole story of emancipation was nonsense and was never right for her—that, having recently fallen in love, she was happy to have finally embraced a feminine role? If someone like H., who, because she “could not do otherwise,” ended up in a form of life in which she were protected by her husband while assuming a subordinate position in relation to him, would she not be just as problematic (and a candidate for alienation) as someone who was in constant tension with her aspirations?

We can see the problem that the demand for emancipation is faced with: where is the criterion to come from that enables us to decide which of these two sides is more appropriate to who she is, or which part of herself is really her?

An unevaluated, merely factual “agreement with oneself” can obviously not be the criterion we are looking for. Are not the most preposterous conversions routinely accompanied by a claim to have finally found oneself? And does not resignation also lead to a kind of agreement with oneself? If H. were to adopt a traditionally feminine way of life, one would have to wonder whether doing so was an act of resignation in which she *gave up* rather than *found* herself. By the same token, however, if she were to succeed in overcoming her opposing impulses, one could ask whether she was committing herself to too rigid a self-ideal, one that required her to deny too many parts of herself—possibly resulting in a personality that was completely rigid in its self-control. Formulated somewhat paradoxically, the suspicion is that both might be cases in which being in agreement with oneself is achieved at the price of a loss of self and where her self-conception as a whole would be false, manipulated, inappropriate, or illusory.

It is common enough that we question ourselves and others in this way, and not doing so would be costly. But what can such a doubt about one's own identity—an "objection" to oneself—be based on if we lack a secure standpoint that could reveal one's *true self*, even on the model of a volitional nature? How, then, is emancipation, as an "emancipation from alien powers" that we are ourselves, to be conceived? Here the dilemma of emancipation becomes clear: every attempt to pose such questions and to justify this kind of doubt is like trying to pull the carpet out from under oneself.

The proposal I want to make to solve this dilemma can be formulated as follows: emancipation and the self-critique bound up with it must be understood as a *free-floating enterprise*—an undertaking that cannot be grounded in advance but only in the course of the process itself, a process in which one cannot appeal to something that one already is and in which one can at the same time "come to oneself." This means that emancipation must be conceived of as "rebuilding on the high sea"³⁴—where a critique of alienation is the driving force and means of such a rebuilding. (And here, too, as suggested in part 1, what is important is the *how* and not the *what* of this process.)

In what follows I attempt to describe the processes of self-alienation and emancipation when understood in this sense. In so doing, I will approach the problem gradually, "from the outside," that is, from the negative conditions of authentic will formation (as proposed by John Christman and Raymond Geuss). (Although in the end we will see that even these conditions rely on a positive vision of what it is to be oneself.) In the next step I will then elaborate my proposal (in line with the idea of emancipation sketched previously) that authentic "being oneself" is to be understood as a mode of being freely accessible to oneself.

MANIPULATION AND CONSTRAINT

Under what conditions must the process of will formation have taken place if it is to count as my own will? Two such conditions are:

First, the formation of my desires must not have been due to *manipulation*. We do not recognize desires as our own if we have reason to believe that they came about through manipulation. Desires can be authentically my own only if I was free in their formation and if they came about without the manipulation of others. H., for example, might conjecture—or we could conjecture about her—that her behavior is a product of manipulative conditioning, a consequence of her gender-specific socialization. Thus, to follow

Copyright © 2014, Columbia University Press. All rights reserved.

a suggestion of John Christman,³⁵ we must take into account the history of the development of a trait or the genesis of a desire—and thereby rule out manipulation—in order to judge the authenticity of desires.

Second, as Raymond Geuss emphasizes in his reflections on the problem of true interests, will formation must take place under conditions in which the *alternatives* at my disposal are alternatives I can really want to choose among. Following Geuss's account of the "optimal conditions" of choice,³⁶ we can say that what I really want cannot be the result of a choice between alternatives that are inappropriately constrained. Something will count, then, as an authentic desire only if it emerges from a choice among acceptable alternatives. Applied to H.'s case, the alternative between protection and emancipation that H. sees herself faced with would, if this description is correct, be unacceptable. Neither of the two possibilities—emancipation without protection and protection without emancipation—can rationally be what she really wants. If it should turn out, then, that in a patriarchally organized society women were systematically confronted with the choice between such alternatives, then in important domains—at least for women—the conditions for forming authentic desires would be lacking.

The exclusion of manipulation and the idea of optimal conditions or acceptable alternatives are clearly two basic conditions that must be satisfied. But they help to clarify only the preconditions of authentic will formation. And the difficulties with these suggestions are obvious. Given that one is always influenced by one's environment—is always in some way socialized—when is the boundary of manipulation crossed? And given that alternatives are always constrained, what constitutes sufficient and acceptable alternatives?

What exactly constitutes unacceptable manipulation if one assumes that the formation of desires and dispositions is always shaped by outside influences and that socialization necessarily involves being influenced by others? Apart from the drastic but hardly plausible cases of brainwashing that are frequently discussed, the manipulation condition does not go far enough as long as we cannot distinguish between manipulative and nonmanipulative influences. It pushes back the question of what is *one's* own and what is *alien* without answering it. In our debate, for example, both sides could argue that H. is dominated by alien influences and manipulation: the antifeminist would point to indoctrination by feminist "ideology"; the feminist, in contrast, would trace back the conceptions of love and family that she cannot shake loose to a deep-seated conditioning by a patriarchal environment and its gender-specific socialization, elements she must liberate herself from in order to become

herself. While it is correct that the genesis of desires should be taken into account in judging their authenticity, the absence of alien influences is by itself inadequate as the sole criterion of what counts as a successful genesis. Here we run into a problem we already encountered in discussing the appropriation of roles: to what extent are formative influences *alien* if *one is constituted* by precisely such influences? Just as in the case of roles it made no sense to presuppose a pure self lying underneath or behind its roles, it makes no sense, in the case of will formation, to assume that there are pure desires that exist prior to all social influence. As for the case of appropriating roles, criteria must be developed that enable us to distinguish between what is one's own and what is alien other than by appealing to something "unspoiled" or uninfluenced by others.

If, on the other hand, as Raymond Geuss formulates it, "agents' 'real' interests are the interests they would have formed in 'optimal' (i.e., beneficent) conditions,"³⁷ and if these optimally favorable circumstances are spelled out (as Geuss does in his discussion of true interests) in terms of "perfect knowledge and freedom,"³⁸ then this idea, if it is to refer to more than the absence of external obstacles, also calls for an elaboration of the conditions under which perfect knowledge is available and can be acted on in perfect freedom. Here, too, substantial assumptions come into play that are more difficult to articulate the more subtle the constraints that must be taken into account. (This is easy in the case of the considerations that Geuss first introduces—hunger and extreme deprivation—but it is more difficult when what is at issue are possibilities for development, as is the case in my emancipation example.)

(6) BEING ONESELF AS SELF-ACCESSIBILITY

What we need is a positive description of what it means not to be determined in what one wills by alien powers. I will sketch out a proposal for providing such a description, which, on the one hand, does not rely on an independent criterion—an Archimedean point that defines the true self—and, on the other hand, goes beyond a merely factual "agreement with oneself." This proposal aims to broaden—or, better, to refine—the idea of a coherent self-conception into one of a self-conception that is both *coherent* and *appropriate or fitting*. Appropriateness in turn is to be determined by the criterion of *self-accessibility* (as revealed in one's practical engagement) and of *having oneself at one's command*: one is oneself when one is accessible to oneself and can "move freely"

in what one does: so understood, being oneself is not a state but a process; it is not something one *is* but a way of taking part in what one *does*. In accordance with this, the attempt to identify alien desires and to replace these alien “intruders” with desires that are one’s own (as defined by the idea of emancipation I have discussed) is a free-floating, self-balancing developmental process.

EVALUATED COHERENCE AND APPROPRIATE SELF-CONCEPTIONS

This again leads us back to our example: the source of the dilemma was the difficulty of finding a criterion for determining which of H.’s desires and behaviors correspond to who she is, which of her desires are her *own* and which “intrude” into her personality as *alien* elements.

Given the significance of self-conceptions for the problems we are considering, an obvious possibility for defining what is alien would be to appeal to a self-conception’s *internal coherence*. This relocates, as it were, the question at issue: it is no longer a question of whether what I want and do really *fits me* but whether the various things I want and do—the things I identify with and that matter to me—*fit together with one another*. Instead of looking for a criterion within ourselves that would enable us to determine what really belongs to us, we are now asking whether a person is consistent or coherent in her expressions and activities, whether she is able to bring the diverse parts of her personality into relation with one another and to integrate them. On this criterion, whether a desire, a passion, or an impulse “fits” us cannot be decided by examining an individual desire but only by looking at how our diverse desires “hang together.”

Self-conceptions become important here for various reasons. I understand a self-conception, provisionally, as something that establishes connections among our attitudes and desires and gives them a certain order. These connections are essentially *interpretive*. Neither individual desires nor the connections among them are independent of interpretation or objectively given: whether Gisela Elsner’s predilection for luxury consumer goods fits with her socialist views, or H.’s giggling with her feminism, is a question of interpretation and self-interpretation. It is not traits and desires themselves that must fit together but a person’s interpretations of them. What is important is whether I can integrate what I want into the conception I have of myself as a person. Appropriating or identifying with one’s desires essentially means weaving them into a coherent interpretation. Being internally coherent, then, consists in

having, developing, and pursuing desires and plans that can be integrated into a *coherent self-conception*.

Now it is easy to see that the idea of coherence alone is not sufficient to answer our question: delusions can also be internally coherent! We must therefore be concerned not only with the internal coherence of our self-interpretations but also with their appropriateness. What, though, makes a self-conception *appropriate or fitting*? In order to pursue this question, it is first necessary to further elaborate the idea of a self-conception.

Having a self-conception or *conceiving of oneself as something* involves taking a stance toward oneself and one's life—understanding or interpreting oneself in a certain way—that is distinctive of the type of being we are and is constitutively bound up with the way we lead our lives as persons. (It is for this reason that Charles Taylor speaks of the human being as a self-interpreting animal.)³⁹

Self-conceptions involve two elements. First, we not only do and want particular things; we also relate to the fact that we do and want them and in doing so we *understand* ourselves as someone who does and wants those things. Second, having a self-conception means establishing *connections* among these individual elements. Taken together, this means that a self-conception is based on more than merely an inventory of objectively given traits and actions; it is an internal principle of organization, an attempt to make our desires and actions “hang together” and thereby give them *meaning*. Loosely formulated, developing a self-conception means “making sense of” oneself. In chapter 9 I will return to the difficult questions bound up with this, such as how we are to picture this process and how much coherence a successful relation to oneself requires. Here I only want briefly to anticipate a misunderstanding: self-conceptions should not be taken to refer exclusively to sophisticated, higher-order interpretations of the sort a person has of herself when, for example, she understands herself as a “feminist” or “leftist.” Self-conceptions need not always be fully explicit; they can sometimes inform what we do without being expressly articulated. And they are not always fully coherent from the start; for our purposes it is enough to say that they are oriented toward coherence.

It is also important that a self-conception has a dual character: it is at once an *interpretation* and a *project*—a self-interpretation as well as a projecting of oneself. In my self-conception I understand myself as the person I am and at the same time I project—or fashion—myself as the person I want to be. Neither involves merely an objective inventory of facts.

As an *interpretation* of what constitutes us, the self-conception we develop starts with a given material, the facts of our life history. On the basis of this we attempt to understand who we are and what constitutes us. At the same time, this material is chosen, interpreted, and structured. There is no “naked truth” here; the various aspects of our lives are meaningful only if we make them so. It might, for example, be true of someone, objectively considered, that he comes from a “good family,” has enjoyed the best of educations, is a connoisseur of art and wine, and has had asthma since early childhood. But this does not mean that all these traits automatically belong to his self-conception. He could take his asthma as the basis for an intense identification with Marcel Proust and understand his chronic suffering and the distance from the world it involves as an essential aspect of his personality or he could, surprised each time he has a new attack of shortness of breath, barely perceive it as an annoyance and repress its implications. Even his solid bourgeois background need not become the object of a positive or negative identification (pride or shame); coming from a particular kind of family can also be more or less significant for one’s self-conception.

As a *project*, on the other hand, a self-conception defines who one would like to be or what one thinks one ought to be. If I understand myself as a feminist or as someone who looks after her friends, I not only interpret who I *am*; I also ask myself who I *want to be* and I orient my future actions and desires toward conduct that fits this conception. There also belongs to this an implicit value orientation—that I find it correct to orient myself in this direction. In this respect a self-conception is closely related to a self- or ego-ideal that expresses “what is important for me in life, what kind of human being I would like to be, what I would like to strive to be.”⁴⁰ It seems clear that both components, interpretation and project, are interrelated. Thus my project for myself will more or less shape my self-interpretation; conversely, my project for myself can be the result of a particular self-interpretation: I become a feminist because I interpret certain of my life experiences in a particular way; I interpret them in that way because I am a feminist. And, of course, both interpretation and project shape what one is. As Jonathan Glover notes: “the way we think of ourselves helps to shape what we are like.”⁴¹ This tension, between interpreting and projecting what one is, is essential to the idea of a self-conception. One might call this the reconstructive-constructive or the hermeneutic-creative character of self-conceptions.⁴²

Now the following implications of this account are of interest for our question concerning the distinction between appropriate and inappropriate

self-conceptions: what constitutes someone as a person does not lie before or behind her self-conception but is interwoven with it. Identity is not an objective fact beyond interpretation. This accentuates our problem. If everything is interpretation, and if interpretation always creates its object, how can we distinguish correct from false interpretations? How is it possible to distinguish self-conceptions that correspond to who we are from those that do not? When does someone understand herself correctly or incorrectly? How can a self-conception, as I have just described it, be false, illusory, distorted, or inappropriate?

My proposal for answering these questions rests on the following assumption: self-conceptions neither merely reflect objective facts, nor are they mere inventions. Self-conceptions have foundations to which they can do justice in varying degrees. It could be, for example, that someone understands himself incorrectly if he denies the role that asthma has had since early childhood for his relation to himself and to the world, or that he understands himself incorrectly if, under the influence of a particular subculture, he denies his connection to bourgeois values and forms of life. Our suspicion that such a person understands himself incorrectly and that his self-image is illusory does not rest only on the fact that we know he actually had a particular sickness or family background; it is typically also based on what we take to be signs that his background is of greater significance than he thinks it is. We notice, for example, that some of his behavior contradicts his self-conception or, more generally, that it is difficult for him to act in accordance with the image he has made of himself.

Here there are parallels to the general problem of interpretation: an interpretation—including that of a text or a work of art—is powerful when, among other things, it can bring together a large number of significant details and establish connections among them. It is the more compelling the less it is forced to exclude facts that do not sit well with it. Naturally, these are not hard and fast criteria; in the end, the question whether certain aspects of an interpretation run counter to it is also a question of interpretation. And this is an unending process: when there is doubt, an interpretation is valid so long as it is not replaced by a more compelling one.

My thesis, then, is as follows: in the theoretical realm (or when we observe things) an interpretation that does not fit means that we do not interpret things appropriately, we do not *understand* them. If this is carried over to the domain of practical relations to self and self-conceptions, an interpretation that does not fit means that something “stands in their way,” that the practical relations

to self and world that correspond to them or by which they are guided are marked by *functional disturbances* that are expressed in various forms of distortions and inhibitions of action. Whether one's self-conception and the way of leading one's life that follows from it fail to fit with who one is depends on whether there are practical inhibitions and contradictions in what one does and in how one understands oneself in doing it.

Whereas an inappropriate interpretation of a picture or a text means that we are unable to *understand* it sufficiently,⁴³ having an inadequate self-conception means that it does not “work” or *function*: we cannot live with it or act within it. In both cases there are, as mentioned above, only “soft” criteria. Yet these criteria are so frequently applied and appear so self-evident that it is difficult for us to imagine our relations to world and self without these (mostly implicit) practices of judging and understanding.

What follows from this for my initial question concerning the authenticity of desires is that authentic desires are those that can be fit into an appropriate conception of oneself, where appropriateness is determined by whether that conception “works” or functions.⁴⁴

SELF-ACCESSIBILITY AND THE INHIBITION OF ACTIONS

What does it mean, though, that a self-conception does not “function”? I would like to elaborate this idea as follows: the crucial point is whether my self-image, my self-conception, and the desires and projects bound up with it result in my being or remaining *accessible* to myself in them and in being able to act freely on the basis of them. At issue, then, is a kind of inner mobility and self-accessibility.

If H., for example, adapts herself to the traditional role of a woman, is she not forced to do so at the price of closing off essential parts of her personality, not only parts of her history and her social surroundings but also some of her fundamental desires and longings? But a closing off of this kind means that a part “of herself” (as we can now say without being too sensational) is not accessible, that she is forced to avoid certain things and is unable to integrate them, that there will be taboos and “no-go areas” in leading her life that she cannot integrate into her self-conception. There will then typically be strategies of avoidance and rigidification—a familiar phenomenon.

That functional disturbances are to be conceived of as disturbances of accessibility to oneself and to the world is still a very vague and unexplained claim. I propose to elaborate it (tentatively and incompletely) by listing some symptoms of such disturbances:

■ *Rigidity* in relation to oneself can, for example, be identified as one such functional disturbance. Here one is not accessible to oneself insofar as one rigidly holds to previously made decisions and is thereby unable to integrate opposing impulses. As Martin Löw-Beer has shown,⁴⁵ a certain lack of contact with oneself can also be seen in the moralistic traits that characterize rigid personalities. The way in which rigid personalities hold on to things that have outlived themselves or to things they are unable to live out gives rise to an objectifying attitude toward themselves that can be described as a lack of vitality and an inability to take part in their own lives. Löw-Beer articulates this point in his striking example of a man who believes that he must love a woman because she meets certain criteria he considers important. It is not only that these criteria are too impersonal; the stance he takes to himself when demanding this of himself is an objectifying one, not one appropriate to leading a life. Similarly, it could also turn out (contrary to the interpretation I set out previously) that it is rigidity that hinders H. from giving in to her "feminine" impulses.

■ Very generally one can consider *rigidity*—rigidly holding on to previously established norms and self-images without being able to adapt them to new situations—to be a disturbance of self-accessibility, insofar as it means closing oneself off from new experiences and conflicting emotions. In line with this suggestion, Richard Sennett characterizes a "purified identity" as a pathology. The search for too much coherence that does not allow itself to be troubled by anything that conflicts with it is, then, just as problematic as too much discontinuity: "the enterprise involved is an attempt to build an image or identity that coheres, is unified, and filters out threats in social experience."⁴⁶ An adequate self-conception must be open to different outcomes and experiences; an inadequate self-conception is not.

■ While inaccessibility to oneself means being closed off to experiences, it is also characterized by an inaccessibility to reasons. A person who is not accessible to herself cannot translate rational insights about herself or her life into action. The asthmatic, for example, who denies his sickness because he cannot integrate it into his self-image does not succeed in acting on the entirely available intellectual insight that he needs medical care. A self-conception in which one was accessible to oneself, in contrast, would not block insights of this type.

■ Similar to this is the phenomenon of not having access to one's emotions or of having inappropriate emotional reactions. Part of self-accessibility is reacting in emotionally appropriate ways—for example, mourning in the case of illness and loss⁴⁷—and being able to relate to these reactions. An appropriate

self-conception is able to integrate such reactions; an inappropriate one suppresses them. Illusory self-images, for example, are often those that suppress experiences of illness or failure. Phenomena of self-deception also belong to this set of problems. Presumably someone who is constantly self-deceived will be inaccessible to herself, since she must protect herself from evidence that might help correct her self-deception.

Hence self-accessibility in general can be characterized as a complex cognitive and emotional state that includes being sufficiently familiar with oneself to be able to perceive one's own needs, to interpret them, and to draw practical consequences from them. Our true self, then, is not merely one that is acquired in the absence of compulsion; it is, formulated positively, a self-relation in which we can move freely and in which we are accessible to ourselves.

In relation to the topics discussed in this chapter, self-alienation means not being able to move freely in one's life, being inaccessible to oneself in what one wants and does. It includes a form of internal division and estrangement from one's own desires that consists in a limitation of one's power to have oneself at one's command in all that power's complex manifestations. Conversely, overcoming self-alienation occurs through a gradual recovery of self-accessibility, without this requiring the Archimedean point of the true self that defines one's real needs. Thus the form of the question H. must pose to herself is not "What do I really want?" but rather "What am I actually doing in what I already do, and *how* does that happen?" It would then be an unforced, transparent relation to her desires and behavior and an openness in dealing with them that allowed us to determine her unalienated or "true" desires.

THE NORMATIVE STATUS OF ALIENATION CRITIQUE

I conclude with a question that also emerged in each of our earlier chapters: why is self-alienation problematic? What normative standard justifies regarding a condition of internal division with respect to one's own desires, as I have described it here, as problematic? In the case of internal division the immanent character of such a standard is obvious: can I will that my will not be my own or the desires I pursue not be mine? Here, with the help of Tugendhat's terminology, introduced previously, one can speak of a hindrance in the well-functioning of volition itself. From what standpoint, though, can we identify a well-functioning will or hindrances to the same?

If individuals are not immediately given to themselves, they are also, to a certain extent, capable of being deceived about themselves; they can understand themselves falsely. As external observers, for example, we can draw their attention to contradictions between the desires they express and particular behaviors that run counter to those desires. As a first step, one can claim that in uncovering such contradictions the subject has no privileged access to itself and that for this reason one can criticize individuals from the outside in a way that remains immanent.⁴⁸ As with a psychoanalyst's interpretations, interpretation from the outside and self-interpretation must agree if an interpretation is to claim validity. Even such agreement, however, is no final guarantee against shared deceptions. Here, too, there is no Archimedean point, although this does not imply that interpretation and reflection are simply arbitrary and subject to no constraints.

COMPLEXITY AND COHERENCE

In concluding I would like to discuss one more objection: is my model's idea of having oneself at one's command, as developed in relation to the problem of internal division, also too robust, and does the view I have sketched invoke too harmonious a conception of coherence? Do not alien desires and parts of ourselves—the existence of different, not always compatible, parts of ourselves that we do not have at our command—belong just as much to our “own life” as those constitutive, intractable aspects of ourselves outside our command (cf. chapter 5)? Are not inconsistencies and contradictions part of the complexity of persons, without which we would not be ourselves and which we therefore cannot (and ought not to want to) banish from our lives? I will return to these questions in connection with the postmodern critique of the subject in part 3. For now I will note only briefly: first, whether certain experiences and parts of ourselves are part of our lives depends on how we can integrate these initially alien elements at a deeper level. In order to count as our experiences (however disconcerting), there must be someone who can have these experiences. Perhaps this capacity for integration must even be all the stronger the more one comes into contact with experiences of “otherness.” What is required, then, is not coherence in the sense of harmony or a seamlessly unified meaning of life but a capacity, underlying one's discontinuities, for relating to what one feels and does.

Second, self-accessibility and the functional capacity of one's own will are open criteria the nonfulfillment of which manifests itself in practical conflicts

or, more precisely, functional deficits. Self-accessibility, then, is a question of degrees, and it can manifest itself and be attained only in dealing with such conflicts. The problem arises because, and to the extent that, H.'s internal division hinders her in doing what she really wants, in being able to move freely in her life. What an unalienated life requires, then, is not that all ambivalences, disunities, or disharmonies in a person be completely sorted out; it requires instead the capacity to be able to react to such problems—or inhibitions of functioning—when they appear. In cases of doubt that can also mean—contrary to the accusation that the unalienated self is too harmonious—first making them into conflicts.

Third, self-accessibility, or having oneself at one's command, as our distinguishing it from rigidity has shown, does not mean having “everything under control” or keeping to a strict model of who one is at any cost. Perhaps one should characterize the capacity in question in a more favorable way: it is about being familiar and able to deal with oneself.