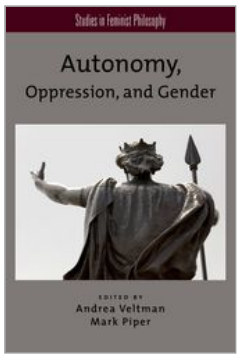

Autonomy, Oxpression, and Gender
Andrea Veltman and Mark Piper

Print publication date: 2014
Print ISBN-13: 9780199969104
Published to Oxford Scholarship Online: August 2014
DOI: 10.1093/acprof:oso/9780199969104.001.0001

Autonomy and Adaptive Preference Formation

Natalie Stoljar

DOI:10.1093/acprof:oso/9780199969104.003.0011

Abstract and Keywords

This chapter examines the consequences for autonomy of the phenomenon of adaptive preference formation. Oppressed agents are said to have *deformed desires*, that is, desires that are adapted to conditions of oppression. Feminists often suggest that deformed desires are incompatible with an agent's autonomy. The chapter identifies two models of adaptive preference formation. It argues that, although adaptive preferences are not incompatible with autonomy by definition on either model, many adaptive preferences of concern to feminists *count as* autonomy impairments. Preferences adapted to oppressive conditions fail tests introduced by both procedural and substantive theories of autonomy. The chapter also addresses the position that adaptive preferences are not deformed but rather are rational choices under conditions of oppression and deprivation in which agents have very limited options. It argues that, even if many adaptive preferences are the products of a rational cost–benefit analysis, this does not entail they are autonomous.

Keywords: Autonomy, adaptive preferences, deformed desires, oppression, deprivation

> At a Sunday Bible-study group I attended for teenage girls, the mother who was teaching had the girls hold hands, march in a circle and say: “My husband will treat me like the princess that I am. He will be the head of my household.” But the girls’ own ambitions seemed at odds with that vision. One girl earlier confessed that her biggest earthly temptation during her college years was likely to be “pursuing too many higher degrees.” Another was known to her friends in the group as the “future president.” I got the sense that relying on a man was not what they considered their best option.[1]

Many philosophers think that the inculcation of gendered norms is harmful to *autonomy*: for instance, gendered socialization may damage the skills and competencies required for autonomy; oppressive stereotypes may block the ability to imagine alternatives thereby limiting and constraining life plans; hostile and sexist responses to women’s aspirations may impair ambition and damage self-esteem so that even if nonsexist options are institutionally available girls do not pursue them.[2] It is encouraging that some of the young girls described at the outset of this paper articulate ambitions for themselves that do not require deference to men, but it remains to be seen whether any of them find the resources to resist their socialization and put these ambitions into effect.

In this chapter, I focus on an argument claiming that the inculcation of oppressive norms damages autonomy in a particularly insidious way. Agents who are oppressed come to internalize their oppression: they come to believe in the ideology of oppression and to make choices, and form preferences and desires, (p.228) in the light of that ideology. Ann Cudd says that oppression creates *deformed desires*, “in which the oppressed come to desire that which is oppressive to them…[and] one’s desires turn away from goods and even needs that, absent those conditions, they would want.”[3]

Sandra Lee Bartky describes an example of this process in her analysis of the phenomenology of oppression. Through what she calls the “interiorization of the fashion–beauty complex,” agents come to believe oppressive norms that tie appearance to self-worth.[4] Bartky says that “repressive satisfactions”—that is, deformed desires—“fasten us to the established order of domination, [to] the same system…produces false needs” and that false needs are produced by the “denial of autonomy.”[5] These remarks suggest that desires that reinforce one’s own oppression are morally problematic because they are formed by agents with

impoverished autonomy. The oppressive conditions are responsible for the desires, not the agent herself.[6]

Theorists of oppression propose that oppression is distinctive because it a group harm. Oppression occurs when a group suffers systematic injustice due to institutional structures or background social practices.[7] Paradigm examples of oppression are systematic injustices suffered by groups whose members share social identities, for instance those of class, gender, race, sexuality, or disability. Although oppression employs the notion of group disadvantage, agents who are members of oppressed groups are also harmed as individuals. Cudd identifies both material and psychological harms of oppression.[8] For instance, racial segregation in the United States is a material and economic injustice that (p.229) victimizes a group. Black Americans suffer systematic disadvantage through residential, job, and school segregation. Individual agents who are black are harmed in virtue of their group membership; they are prevented from attending better schools, doing better jobs, and living in the more affluent areas. Similarly, the unjust stereotyping of a social group can cause members of the group to suffer psychological harms such as humiliation and shame.

In addition to direct material and psychological harm, Cudd describes *indirect* psychological harms that originate within the psychology of the oppressed themselves and hence are often responsible for the self-perpetuating and entrenched nature of oppression. On her account, deformed desires are indirect harms that can be mistaken for "legitimate expressions of individual differences in taste" when in reality they are "formed by processes that are coercive: indoctrination, manipulation and adaptation to unfair social circumstances."[9] If Cudd is right, we have a neat argument for the conclusion that oppression undermines autonomy: gender oppression leads to deformed desires and deformed desires constitute autonomy impairments.

One version of the argument is that desires for one's own oppression are ipso facto deformed due to their contents. For example, Anita Superson claims that deformed desires fail a properly fleshed-out "informed desire test." They are irrational and impair autonomy because desiring one's own oppression is inconsistent with the agent having an appropriate sense of her own moral worth.[10] For the purposes of this paper, I set aside Superson's position and focus on a second possible argument that deformed desires are autonomy impairments, namely, deformed desires are *adaptive preferences* and for that reason "paradigmatically nonautonomous."[11] In Jon Elster's canonical example of sour grapes, a fox, after finding that he can't reach some grapes, decides (p.230) that he doesn't want the grapes

after all. Elster takes this to be an unconscious and nonautonomous process that he calls “adaptive preference formation.”[12] The unconscious accommodation of desires to feasible options often occurs in conditions of oppression. For example, a girl raised in a patriarchal household may come to prefer domestic chores because other nontraditional options are not feasible for her.[13]

Scholars of development ethics also employ the notion of adaptive preference formation.[14] Martha Nussbaum and Amartya Sen focus on agents’ habituation to conditions of severe deprivation.[15] On the face of it, deprivation and oppression are not the same injustice. Oppression occurs in affluent societies in which people are not typically victims of severe economic deprivation. The young girls described in the opening quote are members of a group that is subject the patriarchal oppression but they live in conditions of relative affluence.[16] It is plausible, however, that agents in conditions of severe deprivation in the developing world are members of a group that is subject to economic oppression. In their own cultural context, they may also suffer oppression as women or as members of particular ethnic minorities. As we will see, there are sufficient similarities among the two sets of examples to treat them together for purposes of an analysis of autonomy.

There are two main challenges to the position that preferences that are adapted to the circumstances of oppression are deformed and constitute autonomy impairments. The first, articulated recently by Serene Khader, claims that even if desires for oppressive conditions are morally problematic in some sense, this is *not* because they are “autonomy deficits.”[17] Khader identifies a category of “inappropriately adaptive preferences” that are morally problematic mainly because they are inconsistent with the flourishing of the agent who forms the preference.[18] However she argues that adaptive preferences are not autonomy deficits on either “procedural” or “substantive” accounts of (p.231) autonomy. Procedural theories claim that critical reflection is sufficient for autonomy: agents are autonomous when they critically reflect on their motivations, beliefs, and values in the right way. Procedural approaches seek to define autonomy using morally neutral conditions such as reflective endorsement or nonalienation.[19] Substantive theories build in normative conditions, for instance moral attitudes to oneself or background moral conditions.[20] Khader considers examples of adaptive preferences including women’s preferences to malnourish themselves to feed their male relatives or their endorsement of harmful practices like genital cutting. She argues that neither the tests employing critical reflection required by

procedural accounts nor the tests employing moral conditions required by substantive accounts show these adaptive preferences to be autonomy impairments.

A second challenge claims that preferences that are adapted to the circumstances of oppression are not deformed; rather, they are a rational response to difficult and distressing circumstances in which options are curtailed. These preferences have been described as the products of a rational process of "bargaining with patriarchy"[21] or one in which they reflect a rational compromise because they correspond to the second best alternative that is available to the agent under the circumstances.[22] The challenges are united by various common underlying concerns, for instance that treating oppressed and deprived people as lacking autonomy in effect extinguishes or disrespects their agency, their choices of values and conceptions of the good; that it amounts to an empirically inaccurate generalization about their psychologies; or that it potentially licenses coercive government intervention in their lives.

My goal in this paper is to argue against both these challenges and to dispel some of the worries underlying them. I agree with Khader that adaptive (p.232) preferences should not be *defined* using the notion of autonomy. They are formed through psychological processes and should be defined using psychological criteria. Nevertheless, I argue that on several prominent theories of autonomy, both procedural and substantive, the adaptive preferences of concern to feminist theorists (deformed desires) will count as impairments of autonomy.[23] I distinguish two models of adaptive preferences, both of which originate in Elster's work. The first, the *psychological processes model*, proposes that adaptive preferences are produced by a distinctive causal and psychological process. On this model, it is the flawed process that is responsible for the autonomy impairment. The second, the *freedom to do otherwise model*, characterizes adaptive preferences as adjustments that occur when options that agents would choose under other or better conditions are excluded from their feasible set. On this model, it is the limitation of free agency or freedom to do otherwise that is responsible for the autonomy impairment.

Sections 1 and 2 address the claim that adaptive preferences are not autonomy impairments. I first consider adaptive preferences construed on the psychological processes model and argue that historical, procedural theories of autonomy have the prima facie resources to count deformed desires as autonomy impairments through the device of *procedural independence*.[24] However, I claim that procedural independence cannot fully explicate why deformed desires are

autonomy impairments without importing a background moral theory into the procedural account. Thus, deformed desires are counterexamples to procedural accounts of autonomy that employ a morally neutral test. Section 2 addresses the second model that claims that deformed desires are adaptive in a problematic sense because they fail a freedom to do otherwise test. I argue that this test is a moral one and moreover that agents whose preferences fail this test are precisely those agents who count as having impaired autonomy on well-known substantive accounts of autonomy. (p.233) I conclude therefore that Khader is wrong to claim that the adaptive preferences of concern to feminists are not autonomy deficits.

Sections 3 and 4 turn to the second challenge claiming that what are called *deformed desires* are not deformed but instead are rational accommodations to bad options. I argue in Section 3 that, even if many adaptive preferences are the products of a rational cost–benefit analysis that is endorsed by the agent, this does not settle the question of whether the decision is autonomous. Autonomy theorists often distinguish between the competency and the authenticity dimensions of autonomy[25] or the mental capacities required for autonomy and additional conditions such as adequate significant options.[26] The ability to engage in a cost–benefit analysis may be sufficient to establish some version of rational competency, but it does not follow that the product of this process satisfies the further criteria necessary for autonomy.[27] Section 4 attempts to dispel the worries underlying arguments against classifying adaptive preferences as autonomy impairments.

1. The Psychological Processes Model

As Elster points out, “sour grapes” is a common psychological phenomenon: “people tend to adjust their aspirations to their possibilities.”[28] For example, a wage laborer may initially resist a move from agricultural work in the country to factory work in the city but may come to prefer city life as a result of the experience of factory work. Although the outcome of this process is an adjustment of the laborer’s preferences to his current feasible options (those offered by city life), Elster claims that if the preference for city life is a stable and irreversible preference due to a process of learning and experience it should be distinguished from the problematic phenomenon of adaptive preference formation that is reversible and probably involves “habituation and resignation.”[29] Similarly, adaptive preference formation is to be distinguished from the *deliberate* adjustment of desires to possibilities in character planning:

> Adaptive preference formation differs from deliberate character planning. It is a causal process that takes place “behind my back” not the deliberate shaping of desires...The psychological state of wanting to do a great many things that you cannot possibly achieve is very hard to live with. If the escape (p.234) from this tension takes places through some causal mechanism, such as Festinger’s “reduction of cognitive dissonance,” we may speak of adaptive preference change. The process then is regulated by something like a drive, not by a conscious want or desire.[30]

On this model, adaptive preferences are formed by a distinctive psychological process. The agent unconsciously turns away from a preference that they would otherwise have had to avoid the unpleasant sense of frustration or cognitive dissonance that accompanies having preferences for inaccessible options.[31] The process can occur both in cases of preference *change*, such as that of the fox, and preference *formation*, such as girls in patriarchal households who unconsciously turn away from inaccessible nontraditional options. The model does not employ the concept of autonomy in the definition of adaptive preference formation. Rather, adaptive preferences are thought to be nonautonomous because a “blind” and “unconscious” causal mechanism appears to be incompatible with autonomy; it is not intentional or under the agent’s control.[32] The psychological processes model also does not justify the claim that feminists often seem to endorse, namely that desires for oppressive conditions *automatically* count as deformed and nonautonomous. If a girl who is raised in a patriarchal household comes to prefer domestic roles over other options, her preference could be the result of learning and experience (autonomous) or the result of adaptive preference formation (nonautonomous): “one cannot tell from the preferences alone whether they have been shaped by adaptation.”[33] The question therefore is whether the process of formation of the preferences corresponds to an autonomy-undermining process.

Critics of Elster point out that it is implausible that blind and unconscious processes of preference formation are always incompatible with autonomy. Donald Bruckner considers an agent whose spouse dies. After a period of mourning, she gradually relinquishes the preference to have significant experiences with the spouse and acquires a preference to have these experiences with a new spouse.[34] It seems that, preferences that are the products of unconscious causal mechanisms are not necessarily nonautonomous despite the fact that they are formed behind the agent’s back. Further, Bruckner employs empirical evidence to argue that adaptive preferences can help to

promote a valuable life because adapting to one's circumstances can be conducive to (p.235) subjective well-being, which itself is partially constitutive of a valuable life.[35] He concludes that preferences that are formed through a process of unconscious accommodation to feasible options should be considered presumptively rational and autonomous. The presumption will be defeated if the agent on examination fails to endorse (repudiates) the preference or, if she did not examine it, would fail to endorse it had she examined it.[36] Bruckner thinks that the causal mechanism of formation of a preference is irrelevant to its autonomy (or nonautonomy); rather, the key is the way a preference is retained.[37]

John Christman's notion of the socio-historical self also suggests that unconscious processes of preference-formation are compatible with autonomy.[38] There are many aspects of the self that are not the product of voluntary choice or deliberation. Consider a passage from Shane Phelan: "I was a lesbian. I experienced that moment partially as discovery: so this was the difference I had always felt but never had a name for."[39] For Christman, preferences or desires will be nonautonomous only if they fail either a competency condition or a "hypothetical reflection" condition. Competency corresponds to the capacity of the agent to form effective intentions relative to a desire as well as to reflect critically about the desire. The hypothetical reflection condition employs the notion of nonalienation to characterize authenticity and hence autonomy.[40] Alienation is a "combination of judgment and affective reaction. To be alienated is to feel negative affect, to feel repudiation and resistance…"[41] On Christman's account, if one is (or would be) alienated from a desire but does not succeed in repudiating it, one is inauthentic and heteronomous with respect to the desire. In Phelan's case, her unconsciously formed preferences about her sexual identity are autonomous because she is not (or would not be) alienated from them.

Neither the test suggested by Bruckner nor Christman's hypothetical reflection condition would classify deformed desires as nonautonomous. Consider agents whose preferences for traditional feminine roles are deeply ingrained through oppressive socialization. Due to the effects of the oppressive ideology, the agent treats false stereotypes as natural and formulates desires and plans based on the stereotype. Such agents are unlikely to experience alienation from either the norms that they have internalized or the preferences formed on the basis of those norms. In the absence of actual alienation, is it plausible to conclude that alienation would obtain were the agent to reflect on how her (p.236) preferences were formed? No doubt young women who adopt traditional female roles would think it entirely appropriate that they

were taught to do so by their mothers. As Paul Benson argues, in cases in which oppressive norms are very deeply ingrained it is not plausible to think that the agent would repudiate or feel alienated from desires based on the norms even after reflecting on their process of formation.[42]

Moreover, on Christman's account, forming a different preference in a situation of reduced options to avoid or eradicate actual (or counterfactual) cognitive dissonance may well be a paradigm case of autonomous preference formation. By hypothesis, the fox's desire for the grapes, combined with their unavailability, led to alienation. Hence, in resolving the cognitive dissonance by repudiating the desire for the grapes, the fox eradicates his alienation and seems to achieve autonomy on Christman's account. Similarly, the adaptive preference of girls for traditional female roles (by hypothesis) is the result of the inaccessibility of nontraditional options in the patriarchal context. Girls turn away from inaccessible nontraditional options in part to avoid the frustration of desiring the inaccessible. Hence, alienation would accompany the inaccessible, rejected preference, not the one the agent actually adopts. On Christman's view, the preference for the traditional, feasible, option is the autonomous one.

One option for Christman would be to bite the bullet: since alienation is not or would not be experienced by agents whose preferences are unconsciously adapted to oppressive conditions, these preferences are not autonomy impairments. However, there may be further resources available on his historical and procedural account that will help to demarcate autonomous adaptive preferences from nonautonomous ones.[43] The hypothetical reflection condition, although it purports to offer a historical condition of autonomy, actually requires only that the agent reflect in the present about the historical formation of her desires. Procedural theorists have noticed that critical reflection in the present is not sufficient for autonomy and claimed that present reflection must not be the product of a distorted causal mechanism leading to its formation.[44] As Khader herself points out, an "independence of mind" or "procedural independence" condition has been considered a necessary condition of autonomy on standard procedural accounts.[45] For instance, Gerald Dworkin argues that (p.237) certain influences—such as hypnotic suggestion, manipulation, and coercive persuasion—can subvert agents' critical faculties and undermine their procedural independence.[46] Christman himself mentions that hypothetical critical reflection must not be constrained by "reflection-distorting" factors such as "constriction, pathology, or manipulation" or "being denied minimal education and exposure to

alternatives."[47] Indeed, in early work, Christman compares "happy slave" cases in which the slave has "expunged her desires for freedom only as a result of the oppressive presence of the restraints she faces" with slaves who continue to resist their circumstances and desire their freedom.[48] He claims that if preference formation "resulted from the very presence of a…restraint, bearing down on the agent and forcefully causing a change in the desire," it is implausible that the desire change is an increase in freedom.[49] In other words, the accommodation of a happy slave to her circumstances does not make her autonomous; rather, the presence of coercive, external restraints that inhibit the free formation of her preferences constitutes an autonomy impairment.[50]

Thus, if we understand adaptive preferences using the psychological processes model, the test of procedural independence could be used to identify adaptive preferences that are also nonautonomous. The question then becomes whether the causal and psychological mechanisms that produce deformed desires render them autonomy impairments. Khader considers several examples of the adjustment of options to possibilities that have been labeled "deformed" in development ethics. One is that of intrahousehold food distribution.[51] Some women in South Asia starve or malnourish themselves to feed their husband and male children. Khader points out that such preferences are related in complex ways to religious and gender norms that link self-deprivation and self-discipline. She describes the case of an "elderly Javanese woman [who] recalled being told as a child that women needed to discipline themselves, because they were superior to men who could not control themselves."[52] Khader (p.238) notes that this woman seems committed to her values and reflective about the ways her values inform her decision to malnourish herself. Hence, Khader thinks there is no autonomy deficit on procedural accounts that employ a critical reflection condition. However, on the account I am now considering, being reflective in the present is not sufficient for autonomy. The question is whether the preference to deprive oneself of food due to the internalization of misogynistic norms is procedurally independent—that is, whether the influence on the formation of the preference constituted coercive persuasion or was reflection-distorting due to insufficient exposure to alternative possibilities or some other factor.

Whether or not a preference fails a test of procedural independence is in principle a case-by-case question. However, if theorists of oppression are correct, we can make some general claims about the desires of members of oppressed groups: they are often "formed by processes that are coercive: indoctrination, manipulation and adaptation to unfair

social circumstances."[53] Thus, prima facie, the desires of members of oppressed groups for their own oppression fail a test of procedural independence such as that of coercive persuasion.

Procedural accounts of autonomy therefore potentially employ conditions that would classify adaptive preferences as autonomy impairments. The problem for such accounts is whether the processes that violate procedural independence, like coercive persuasion, can be spelled out in a satisfactory way without importing background moral conditions. Cudd argues that a purely empirical account of coercion is unsatisfactory because it cannot distinguish between a hard choice and a forced choice.[54] Similarly, to flesh out the reflection-distorting influences that correspond to lack of minimal education or inadequate exposure to alternatives, we need some normative account of what counts as minimally satisfactory education or adequate alternatives. Therefore, although procedural theorists can explain how desires for oppressive conditions are nonautonomous, the explanation comes at the expense of maintaining the moral neutrality of their own theory.

2. The Freedom to do Otherwise Model

In the last section we considered a psychological processes model of adaptive preference formation that was derived from Elster's distinction between adaptive preference formation and the similar processes of learning and deliberate (p.239) character planning. On that model, adaptive preferences are accommodations to feasible options that lack autonomy due to a distorted underlying causal mechanism. Elster also offers an analysis of adaptive preferences that employs the notion of freedom to do otherwise:

> We can exclude operationally one kind of non-autonomous wants, viz adaptive preferences, by requiring freedom to do otherwise. If I want to do x, and am free to do x, and free not to do x, then my want cannot be shaped by necessity.[55]

On this second model, preferences are adaptive and nonautonomous because they fail the freedom to do otherwise condition and are shaped by necessity. The Javanese woman forms a desire to malnourish herself only because she is not free to flout the cultural and religious norms that lead to the desire. The girl raised in a household in which she is always expected to help her mother with the traditional chores comes to prefer these chores only because other options are not available to her.

Nussbaum applies the freedom to do otherwise model in her analysis of the preferences of poor working women in India. One of her examples is that of Vasanti who chooses to remain in an abusive marriage because she thinks that although the abuse "was painful and bad, but, still, a part of women's lot in life, just something women have to put up with as part of being a woman dependent on men, and entailed by having left her own family to move into a husband's home."[56] Nussbaum thinks that Vasanti's preference is adaptive; it is accommodated to her feasible options and lacks the condition of freedom to do otherwise.[57] Nussbaum points out, however, that there will be many examples of the accommodation of preferences to feasible options that fail the test of freedom to do otherwise yet are reasonable and conducive to the agent's well-being. Being a basketball player is not a feasible option for a short person, and being an opera singer is not a feasible option for someone with a weak singing voice. Adjusting one's expectations to accommodate the feasible options in these cases seems perfectly reasonable and a good thing to do.

In Vasanti's case, the desire to remain in an abusive relationship is by hypothesis not a good thing for her. Her choice is the result of her freedom being constrained in a morally problematic way. Thus, Nussbaum claims that Elster's morally neutral account of adaptive preferences is inadequate. To distinguish the basketball player and the opera singer from Vasanti and to explain why agents like Vasanti have adaptive preferences that lack autonomy, we need a "substantive theory of justice and central goods": "People's liberty can indeed be measured, not by the sheer number of unrealizable wants they have, but by (p.240) the extent to which they want what human beings have a right to have."[58] The freedom to do otherwise condition must therefore be understood as a moral condition. If agents' preferences are accommodated to external options only because the options are limited by unjust social circumstances, their preferences are shaped by necessity in a way that restricts their freedom.

For Nussbaum, adaptive preferences are morally problematic in situations in which agents face a morally inadequate set of options. In such circumstances, when preferences are accommodated to the options, and agents do not desire what they have a right to have, their autonomy is impaired. Thus, for Nussbaum, adaptive preferences are morally problematic precisely because they are (substantive) autonomy deficits. In response to this kind of position, Khader proposes a perfectionist definition of adaptive preferences. A preference is inappropriately adaptive when it is inconsistent with a person's basic flourishing, formed under conditions non-conducive to basic flourishing,

and would not have been formed under conditions conducive to basic flourishing.[59] Agents like Vasanti satisfy the definition and have inappropriately adaptive preferences. But Khader claims that such preferences do not correspond to autonomy impairments even on substantive accounts of autonomy.[60]

Let me address Khader's argument by focusing on her discussion of Joseph Raz's substantive theory.[61] Raz proposes that autonomy requires adequate external options in addition to mental ability and independence from coercion and manipulation by other agents.[62] For example, he describes a "hounded woman"—a woman on a desert island who is hounded by a wild animal and has to spend all her time and energy planning for survival. The woman has a variety of options in addition to the mental abilities necessary for planning, but the options she has are inadequate. For one thing, they are "dominated by her one overpowering need and desire to escape being devoured by the beast."[63] For another, having adequate options is understood by Raz as a moral (p.241) requirement: "autonomy requires that many morally acceptable options be available to a person."[64]

Khader interprets Raz's test as follows: "autonomous preferences are preferences that reflect an agent's own preferences and are *consistent with her flourishing.*"[65] She considers the case of a person who chooses to be a bullfighter and argues that, since this a dangerous sport, it is not compatible with the agent's flourishing. Khader suggests that choosing bullfighting would count as an autonomy impairment on Raz's account and further that, since choosing bullfighting is not intuitively an adaptive preference, the category of adaptive preferences cannot coincide with the category of autonomy impairments delivered by Raz's account.

This argument is too quick. Choosing bullfighting may be dangerous and incompatible with flourishing (in some sense), but if the agent has a variety of morally adequate options in addition to bullfighting the agent and her choice to pursue bullfighting are nevertheless autonomous on Raz's account. The agent must have the capacity to choose the good, that is, a choice among morally adequate options, but there is no requirement that particular choices are consistent with flourishing. Suppose we assume that bullfighting constitutes a morally bad option because it is dangerous and undermines the agent's well-being. (Note that this is not a very plausible assumption because not all dangerous life choices are morally impermissible.) Even this assumption does not negate the agent's autonomy but only the value of the exercise of autonomy in the particular case. Raz argues that "autonomy is valuable only if it directed at the good.…Autonomy is

consistent with the presence of bad options, [but] they contribute nothing to its value."[66] Choosing a bad option in a situation in which the agent has a variety of morally acceptable ones is an autonomous choice, though not a valuable one.[67] Conversely, agents like Vasanti who are living in conditions of severe deprivation lack morally acceptable options and hence are obvious cases of nonautonomous agents on Raz's theory. Coming to (p.242) prefer one's severely limited options (i.e., adapting one's preferences to one's feasible option set) corresponds to exercising impaired autonomy.

In general, adaptive preferences as Khader defines them would count as autonomy deficits on Raz's adequate options test. On her account, basic flourishing or well-being is defined using a "minimal, vague and cross-culturally acceptable conception" of the good.[68] It is clear that, if an agent lacks the options required for minimal flourishing, she would also be deprived of morally adequate options. Thus, a preference that is formed only because better options are inaccessible to the agent—that is, a preference that would be classified as adaptive on Khader's definition—also would be classified as nonautonomous on Raz's approach to autonomy.

Khader acknowledges that adaptive preferences on her model are likely to be ruled nonautonomous on Raz's account because they are shaped by inadequate options. But she claims that "calling [adaptive preferences] autonomy deficits and incorporating a conception of the good into autonomy leads us to policies that are decidedly illiberal."[69] She worries that if we characterize adaptive preferences in response to oppression, poverty, and deprivation as lacking autonomy, this will license coercive governmental policies that will override the voluntary choices of such agents, thereby compounding their oppression and deprivation. Thus, Khader's main objection to characterizing adaptive preferences as autonomy deficits is a moral one.[70] I postpone discussion of moral objections to Section 4.

Up to now, the discussion has focused on the intersection of Raz's account of autonomy and the adaptive preferences of agents in conditions of severe deprivation. It may be more difficult to invoke Raz's theory to account for the deformed desires of agents in affluent societies. Notice first, however, that while oppressed people in Western countries often live in conditions in which minimal flourishing is possible, it does not follow that the conditions required for minimal well-being rise to the level of adequate options. Second, as Raz and others point out, autonomy comes in degrees. The options of members of oppressed groups—for education, housing, social services, medical

care, employment—are often compromised by comparison with the options of members of groups that are not oppressed.

Consider a phenomenon that Cudd describes as “oppression by choice.”[71] Cudd observes that in a labor market in which the average wage for women is significantly less than it is for men, it might be rational for a mother rather than (p.243) a father to choose to be the primary caregiver of children, that is, to choose either to work part-time work or not to work at all. At the individual level, it is a rational choice because overall the family will be financially better off. But Cudd points out that when the choice is made by many individuals it leads to a vicious cycle in which structural inequalities are reinforced. By choosing to stay home and become the primary caregiver of children, the mother will become a domestic specialist—a specialist in work that is unpaid and undervalued as a contribution to society—whereas the father will acquire experience and seniority in whatever area of paid employment he takes on. When many women make the same individual rational choice, this in turn will reinforce social perceptions that women are primarily unpaid domestic workers rather than potential wage earners. The cycle is oppressive because as a result of the individual women’s choices, “employment opportunities are continually degraded both for the individuals and for women as a group.”[72]

Cudd argues that although women in affluent Western societies are thought to have occupational free choice, in fact their options are morally constrained (relative to those of men); moreover, their options become progressively more constrained as a result of the choices that are initially made *rationally* in response to the structural inequalities. The argument suggests that women—and members of oppressed groups in general—have fewer morally adequate options and hence less autonomy than comparable members of nonoppressed groups. The phenomenon of oppression by choice is not equivalent to adaptive preference formation because the former involves deliberate planning whereas the latter does not. Nevertheless, in both cases agents’ freedom to do otherwise is compromised relative to that of members of groups who are not oppressed.

3. Choosing among Bad Options

I argued in sections 1 and 2 that prominent theories of autonomy, both procedural and substantive, have the resources to classify many adaptive preferences as autonomy impairments. In this section I turn to a differently framed challenge to deformed desires, namely, the position that preferences adapted to the circumstances of oppression are *not* deformed. It is claimed that, rather, they are a rational (and therefore autonomous) response to circumstances in which agents have limited options. For example, Uma Narayan argues that they are the outcomes of a rational process of "bargaining with patriarchy, "[73] and Harriet Baber proposes that deformed desires correspond to the second (p.244) best alternative that is available to the agent under the circumstances. When the best alternative is inaccessible, it is rational to choose second best.[74]

The different versions of the challenge diverge in an important respect. Baber, whose focus is to provide a defense of preference utilitarianism against Nussbaum's critique, argues that many preferences that are called "adaptive" are not in fact genuine preferences at all. Choices in response to what she calls a "raw deal" may not be what agents really prefer. On the other hand, feminist critics like Narayan (and Khader), whose focus is to defend women's agency, claim that preferences for sexist norms or bad circumstances are genuine preferences. In many cases these preferences are reinforced by other goals that the agent takes to be valuable. Khader discusses women who opt for genital cutting or clitoridectomy. She points out that often women are motivated by the wish for community belonging and social recognition, in particular by the goal to promote the marriageability of their daughters.[75] Similarly, the aforementioned preference for malnourishment of the Javanese woman is reinforced by religious conviction. Women's preferences seem genuine when considered in conjunction with these other factors.

Narayan and Baber do, however, offer parallel arguments for the rationality of putative deformed desires. Agents living under oppression or deprived conditions are capable of utility maximization; they are capable of engaging in cost–benefit analyses that weigh up the options available and of making a choice among options on the basis of this analysis. In particular, they note that most agents rank *bundles* of options, not options in isolation. Consider Narayan's example of a community of women in India, the Sufi Pirzadi, who "live in relative purdah (seclusion) within the home and are expected to veil when they are in public."[76] These women acknowledge that purdah severely limits their education and mobility and has the effect of making them dependent on male members of the community. But they also explicitly

recognize benefits, for instance, that veiling signifies “womanly modesty and propriety” and their “superior standing vis-à-vis other Muslim women.”[77] She argues that due to their cultural and religious context these women cannot separate out preferences that limit their movement and promote their dependence from those that promote the piety and modesty that they value: the options come as a bundle.

Baber makes the same point in her analysis of Vasanti. She points out that Vasanti may prefer a bundle of options including abuse and having a roof over her head over a bundle that contains no abuse and being on the street. It is not irrational to choose the former bundle but rather an example of rational (p.245) calculation under oppressive conditions in which the agent gets some advantages and forgoes others:

> We might with equal justification understand Vasanti’s decision as the result of a utility calculation given a reasonable assessment of her options and the probabilities of various outcomes. Vasanti recognizes that given her circumstances, staying in an abusive marriage is her best bet if she wants to have a home and basic necessities: even if she would rather avoid getting beaten, she is prepared to take on that cost in order to avoid her least preferred outcome: homelessness and destitution.[78]

Thus, for both Baber and Narayan, putative deformed preferences are not deformed but rather instances of the exercise of rational agency in which agents rank bundles of options under difficult circumstances.

Baber offers a second argument as well: although Vasanti and other agents may appear to have preferences for oppressive conditions, in fact we cannot infer that they prefer what they choose: “making the best of a raw deal when no other alternatives are available is not the same as preferring it.”[79] Consider the fox. According to Baber, the fox’s preferences have not changed; he is only pretending to himself that he does not value or want the grapes. If a bunch of grapes suddenly became accessible to him, he would “jump” at them, and hence, his preference for the grapes persists.[80] The fox engages in a rational process of settling for second best. The fox’s possible options are (1) grapes and no felt frustration, (2) no grapes and no felt frustration, and (3) no grapes and felt frustration.[81] Although the fox cannot have the best option (1), he “prefers serenity over felt frustration” and thus eradicates the felt frustration that comes with continuing to want the grapes. Similarly, Vasanti does not really prefer domestic abuse. She would jump at a better situation were it accessible to her. But staying under her husband’s roof and putting up with the abuse is a rational

choice because it is better than the alternative, namely, no abuse and homelessness.[82]

These arguments contain two basic ideas. The first denies that the desires of agents who endorse apparently harmful practices such as clitoridectomy or practices in which their rights are curtailed such as purdah are deformed. On the contrary, these agents are rational choosers with complex motivations who are making the best of the circumstances in which they find themselves. The second idea is that there is a difference between what the agent chooses and (p.246) seems to prefer, and what they really prefer. Since many agents like Vasanti, if offered better options, would jump at them, what they choose is not what they really prefer. It is a compromise in the face of a terrible situation in which what they really prefer is inaccessible to them.

One can agree with both of these ideas. Yet neither establishes that agents like the Sufi Pirzadi, Vasanti, or for that matter women in Western contexts who appear to adapt to oppressive conditions are autonomous. Engaging in a cost–benefit analysis may be sufficient for formal rationality yet insufficient for autonomy.[83] As we have seen, theories of autonomy typically distinguish between competency conditions and additional conditions such as authenticity or availability of adequate options. Thus, it is possible that the agents under consideration satisfy competency and formal rationality conditions yet nevertheless fail to be autonomous. For the same reasons, Baber's argument that agents like Vasanti make rational choices that might enhance their subjective well-being given their options is also not incompatible with attributing lack of autonomy to Vasanti and agents like her. The ability to improve one's well-being under conditions in which one's options are curtailed is not equivalent to exercising autonomy.

Indeed, not only is Baber's analysis not incompatible with attributing an autonomy impairment, it also provides an implicit argument that the choices of agents like Vasanti are not autonomous. In effect, Baber is distinguishing between apparent and true preferences.[84] She thinks that reversible adaptive "preferences" are not true preferences and that the true preferences are revealed by the behavioral dispositions to jump at alternative better options. Hence, the fox and Vasanti are deceiving themselves about what they truly prefer. Many theories of autonomy, however, would consider self-deception and autonomy incompatible. For instance, Diana Meyers argues that one of the skills required for agents to be autonomous is that of self-discovery; agents who are blind to their own true preferences are not exercising this skill and are not autonomous.[85] The reversibility of a preference therefore

would be evidence that it is not autonomous because reversibility suggests that it is merely apparent not real. (Indeed, this may be the intuition behind Elster's original classification of adaptive preferences as reversible and nonautonomous and hence unlike the irreversible—and autonomous—preferences that come about through learning.)

Other theories suggest that apparent self-deception is compatible with autonomy, but only if self-deception occurs in the service of right reasons. For example, Henry Richardson argues that an agent is autonomous if she (p.247) acts on a conception of objective moral reasons.[86] He compares the fox's reasoning with that of the character Bully Stryver in Charles Dickens's novel *A Tale of Two Cities*. Mr. Stryver selects "the beautiful and kind Lucie Manette" as his future wife. When she refuses, he persuades himself and others that he never loved her or wished to marry her. In other words, he adopts the fox's strategy of self-deception. He denies that he ever wanted to marry Lucie; he "shields himself from significant loss" and eliminates the frustration and embarrassment at being refused. Richardson claims that nevertheless Mr. Stryver's adaptive preference formation in this case is autonomy preserving because it promotes self-respect and thereby corresponds to acting on a conception of (actual) moral reasons.[87] Richardson's analysis could be applied to Khader's examples of adaptive preference formation, such as that of the Javanese woman who undernourishes herself in part because she thinks it promotes spiritual enlightenment. Given the cultural context, the preference may be an instance of acting on actual moral reasons such as self-respect. But in other cases, including perhaps that of Vasanti, if agents' preferences are not consistent with acting out of self-respect they would count as autonomy impairments.

Baber's discussion raises a further issue. It is not typical that preferences for oppressive conditions can be shown to be unstable and hence for that reason to be merely apparent. It is not the case that women, if given the opportunity, would always jump at the chance of alternative nonsexist roles. For example, there is evidence that the norm that only men are appropriate heads of households is remarkably resistant to shifts of context. A recent article in the *New York Times Magazine* discusses the situation of families in a small town in the United States in which due to a factory closure many men who had been employed and relatively affluent lost their jobs.[88] In several of these families, the wives, although they had not previously worked, had become successful primary breadwinners for the family. At the same time, the men struggled with adjustment to being financial dependent on their wives, and both husbands and wives clung to traditional roles

including the designation of the husband as head of household. Hence there is evidence that cases in which agents adopt oppressive norms (or social roles that are formed by oppressive conditions)[89] are instances (p.248) of stable preferences in which agents have adapted to patriarchal social structures. Baber considers an analogous case of a Cambodian prostitute, Srey Mom, whose freedom was bought from the brothel owner by a *New York Times* journalist.[90] Although subject to appalling conditions in the brothel, Srey Mom had come to prefer some if its material advantages such as jewelry and a cell phone. When given the opportunity to return to her village, she does so, but after a while she returns to prostitution. Baber proposes that the preference for a life of prostitution is a stable and genuine adaptive preference; moreover, given the other options available to Srey Mom, it is too quick to conclude that returning to prostitution is not in her interests or that it fails to promote her well-being.[91]

These examples reinforce the point that although the reversibility of a preference may be evidence that it is merely apparent, it does not follow that stable and irreversible preferences are always autonomous.[92] As we saw sections 1 and 2, even stable and irreversible preferences might fail the tests of procedural independence or adequate options. Similarly, even if a choice promotes an agent's subjective well-being because it is the best of the bad options available (as in the case of Srey Mom), this does not imply that the choice is autonomous.

4. The Possibility of Coercion and Other Worries

Many of the objections to treating deformed desires as autonomy impairments are not conceptual but moral, pragmatic, and empirical. Khader often claims that characterizing agents with adaptive preferences as nonautonomous would license coercive policies. For example, she writes that "if people whose preferences do not manifest a value for their own independence are not autonomous, public institutions may reasonably coerce those people into changing their preferences."[93] She also thinks that the characterizations of agents' desires as deformed and nonautonomous are problematic empirically because it tends to overlook the complexity of their motivations and that this may result in policies with an "ineffective focus."[94] Baber emphasizes the point that oversimplified accounts of the motivation of agents in deprived circumstances can (p.249) be empirically inaccurate and morally problematic. For example, she criticizes Nussbaum's reference to Vasanti being in a "slumberous state" before being exposed to programs in which she came to be aware of her rights.[95]

Indeed, many theorists are concerned that denying that women's adaptive preferences are autonomous is equivalent to denying their agency. In this section, I attempt briefly to dispel these worries.

I do not believe that denying autonomy is tantamount to denying agency. *Autonomy* and *agency* are not equivalent concepts. As we have seen, agents that are judged to have autonomy impairments often retain the full gamut of complex mental capacities. They retain the mental abilities required for planning and weighing up competing options. They retain abilities that are closely related to autonomy such as forms of rationality and the capacity for self-control.[96] Khader articulates a notion of adaptive preference that does not rely on the notion of autonomy. Yet she acknowledges that various aspects of agency may be damaged in agents who have adaptive preferences and that noncoercive interventions may be justifiable to attempt to improve agents' deliberative capacities.[97] Her argument itself suggests that agency and autonomy are distinct ideas.

The wish to defend a robust conception of women's agency under conditions of oppression is underwritten by a very legitimate concern. Focusing on deficiencies in agents' psychologies may end up "psychologizing the structural"—that is, putting causal responsibility for oppression inappropriately on the agent rather than appropriately on the unjust social conditions.[98] Remarks such as that women's agency is "pulverized" by patriarchy or that Vasanti was in "slumberous state" or even that desires for one's own oppression are "deformed" are rhetorical forms of words that potentially reinforce the problem.[99] I agree that these labels are misleading and better avoided.

The negative labels also obscure the subtleties of the autonomy analysis of adaptive preferences. For instance, theories of autonomy distinguish between local and global autonomy. When theorists of autonomy claim that a preference or desire is nonautonomous, this usually means that the conditions for local autonomy—namely, what is required for particular choices, preferences, or desires at particular times to count as autonomous—do not obtain. Vasanti's decision to stay in an abusive marriage may well be locally nonautonomous. If it is, her degree of overall autonomy is reduced, but her agency is relatively unaffected. Global autonomy corresponds to an agent's ability to lead (p.250) an autonomous life. Global autonomy theorists typically introduce external conditions such as inadequate moral options or socio-relational conditions.[100] For instance, Marina Oshana argues that autonomy is a temporally extended, global condition of agents in which they have "de facto power and authority over choices and actions significant to the direction of [their lives]."[101] Severely constraining external conditions

remove the de facto power required for autonomy. Hence, global theories do not usually focus on agents' psychological capacities at all; they do not impugn agency or psychologize the structural. Indeed, on global theories, it is precisely the structural conditions that impair autonomy.

Let us now turn to the question of coercion. Khader notes that for liberal political theory, "autonomy [is] the capacity that exempts people from being subject to coercion."[102] She claims that, once agents are judged to be lacking autonomy, liberals do not have the resources in their theory to block policies that override agents' voluntary choices. Let us suppose it is true that for liberals respect for agent autonomy is the most important moral reason for treating coercive interference with agents as illegitimate. This does not entail that coercive policies are always morally permissible in cases of nonautonomy or lack of autonomy: it does not entail that "public institutions may reasonably coerce…people into changing their preferences." Indeed, coercive policies aimed at changing people's preferences may well be self-defeating if they harm or damage agents' autonomy further under the guise of promoting it.

Khader imagines a policy that denies poor women health care unless they stop the practice of malnourishing themselves to feed their male relatives.[103] Putting aside the fact that denying malnourished people health care would be criticizable on the grounds that it would exacerbate the physical harm they are already suffering, such a policy would be self-defeating. Autonomy is a matter of degree and there are different dimensions of autonomy. It would be self-defeating for a policy whose aim is to augment autonomy to further undermine what little autonomy agents have or to undermine one dimension of autonomy while promoting another. On Raz's account, for instance, a policy that subjected people to direct coercion would inhibit an important dimension of autonomy.

The default position for liberals, in part due to respect for the value of autonomy itself, is that coercive policies are not justified unless the goal is to prevent harm.[104] In the West, this issue can arise for example in cases of mental (p.251) illness or religious values. Suppose a person with a mental illness lacks the competencies required for full autonomy. From a liberal perspective, there is no justification for subjecting this person to any sort of coercive policy unless she is likely to harm herself or others. Similarly, preferences against certain forms of life-saving medical intervention can be based on religious beliefs. Jehovah's Witnesses think that blood transfusions are contrary to biblical injunctions. Even if a theory of autonomy judges this religious

conviction to be the result of indoctrination and hence (for instance) to be incompatible with the requirement of procedural independence, there is no justification for subjecting Jehovah's Witnesses to coercive policies to change their convictions. Coercive intervention such as enforced blood transfusions may be justifiable, but only as a last resort to prevent harm.

Khader seems to accept that the harm principle potentially could provide a justification of a coercive policy. She writes that "we should focus on changing 'cultural practices' only when they cause some sort of [serious] harm or wrong to individuals."[105] In the case of genital cutting, for example, there is wide variation among physical practices as well as among social benefits that ensue. The practice does not always constitute serious harm.[106] Khader is right to say that the mere fact that an agent's endorsement of the practice of genital cutting is driven by a cultural belief that is false and misogynistic does not provide a justification for coercive intervention. But there is stronger position running through Khader's critique: namely, that noncoercive policies, if their aim is to change people's cultural beliefs or conceptions of the good, are never permissible because cultural beliefs and conceptions of the good are delivered by autonomous agency.[107]

This position is implausible. Although liberals (especially proponents of procedural theories of autonomy) put a lot of weight on respecting individual agents' conceptions of the good, nevertheless they endorse the value of autonomy and some version of the harm principle. Since some conceptions of the good are harmful or inconsistent with agents' own autonomy, not all conceptions of the good are equally valuable. As Catriona Mackenzie argues, "ruling out coercive political means for promoting autonomy...does not entail ruling out other political means for encouraging citizens to pursue valuable goals—for example, incentive and reward schemes...; health promotion campaigns; funding subsidies for the arts, and so on."[108] Indeed, since autonomy is a value (p.252) liberals endorse, noncoercive policies that attempt to promote the value of autonomy may be a requirement of justice on liberal political theories.[109]

Conclusion

In this chapter, I set out to examine the consequences for autonomy of the phenomenon of adaptive preference formation in which the oppressed come to "desire that which is oppressive to them."[110] Feminists often seem to suggest that internalized oppression ipso facto impairs autonomy and that the notion of adaptive preference offers support for that conclusion. The preceding discussion has shown that the situation is more complicated. Adaptive preferences are not autonomy deficits by definition, and neither is it the case that all unconscious accommodations to feasible options count as autonomy impairments. On the other hand, many of the adaptive preferences of concern to feminists are autonomy impairments. I argued that even if adaptive preferences are not deformed but rather rational and reasonable decisions in the face of oppression, this does not show that they are autonomous. Preferences adapted to oppressive conditions fail the tests introduced by both procedural and substantive theories of autonomy. The key reason that they fail these tests—whether it is a test of procedural independence or an adequate options test—is that the moral constraints faced by members of oppressed groups due to their oppression reduce their psychological freedom.

Acknowledgments

For helpful discussion and comments, I am indebted to Mark Piper, Andrea Veltman, my fellow contributors to this volume, and the other participants in the workshop "Relational Autonomy: 10 Years On," held at McGill University in September 2012. I am also grateful to the Social Sciences and Humanities Research Council of Canada for a grant that supports my research on autonomy and oppression, and for a second grant that made the workshop possible.

Notes:

([1]) Hanna Rosen, "Who Wears the Pants in This Economy?" *New York Times Magazine*, August 30, 2012.

([2]) See, e.g., Diana T. Meyers, *Self, Society and Personal Choice* (New York: Columbia University Press, 1989); Paul Benson, "Free Agency and Self-Worth," *Journal of Philosophy* 91 (1994): 650–668; Catriona Mackenzie, "Imagining Oneself Otherwise," in *Relational Autonomy: Feminist Perspectives on Autonomy, Agency and the Social Self*, edited by C. Mackenzie and N. Stoljar (New York: Oxford University Press, 2000), 124–150 (and in general all articles therein).

([3]) Ann E. Cudd, *Analyzing Oppression* (New York: Oxford University Press, 2006), 181.

([4]) Sandra L. Bartky, *Femininity and Domination: Studies in the Phenomenology of Oppression* (New York: Routledge, 1990), 39.

([5]) Bartky, *Femininity and Domination*, 42.

([6]) Anita Superson, "Deformed Desires and Informed Desire Tests," *Hypatia* 20 (2005): 109–126; Superson, "Feminist Moral Psychology," *The Stanford Encyclopedia of Philosophy,* (Fall 2012 Edition); <http://plato.stanford.edu/archives/fall2012/entries/feminism-moralpsych/>.

([7]) I follow authors such as Iris M. Young, "Five Faces of Oppression," reprinted in *Rethinking Power*, edited by T. Wartenberg (New York: State University New York Press, 1992), 174–95; Sally Haslanger, "Oppressions: Racial and Other," in *Racism in Mind*, edited by M. P. Levine and T. Pataki (Ithaca, NY: Cornell University Press, 2004), 97–123; Cudd, *Analyzing Oppression*. These authors give accounts of oppression that focus on systematic harm to groups; on these accounts, individuals who suffer harms of oppression are harmed in virtue of their group membership. Haslanger points out that oppression can be perpetrated by agents, or it can be *structural*; that is, systematic disadvantage to a group can occur as a result of social practices even in the absence of agents or legal institutions that intentionally perpetrate the oppression. Although here I take systematic harm to a group and concomitant harm to individuals in virtue of their group membership as sufficient for oppression, I wish to leave open the possibility that single agents or nongroups could suffer oppression. Hence, I leave open the question of whether systematic harm to a group is necessary for oppression. I also do not employ the notion of autonomy in the definition of oppression. Cf. Daniel Silvermint's position on which he claims that oppression obtains when "an individual's autonomy or overall life prospects are systematically and wrongfully burdened." Silvermint, "Oppression without Group Relations," unpublished manuscript, Montréal, Canada, 5.

([8]) Cudd, *Analyzing Oppression*.

([9]) Cudd, *Analyzing Oppression*, 183.

([10]) Superson, “Deformed Desires and Informed Desire Tests.” Superson claims that choosing or preferring one’s own oppression is analogous to choosing slavery; it is making a special kind of moral mistake. Although Superson characterizes her position using a Kantian interpretation of an informed desire test, her argument suggests that she is committed to a “strong substantive” account of autonomy in which “the contents of the preferences or values that agents can form or act on autonomously are subject to direct normative constraints.” Paul Benson, “Feminist Intuitions and the Normative Substance of Autonomy,” in *Personal Autonomy: New Essays on Personal Autonomy and Its Role in Contemporary Moral Philosophy*, edited by James Taylor (Cambridge: Cambridge University Press, 2005), 133. As Benson points out, in Natalie Stoljar, “Autonomy and the Feminist Intuition,” in Mackenzie and Stoljar, *Relational Autonomy*, I conflated this strong substantive account with a strong normative competence condition of autonomy. Although deformed desires would be counted as nonautonomous due to their contents on strong substantive accounts, it is less clear whether agents who have deformed desires would fail a normative competence condition. For a fuller explanation of different versions of substantive account, see Natalie Stoljar, “Feminist Perspectives on Autonomy,” *The Stanford Encyclopedia of Philosophy*, (Summer 2013 Edition); http://plato.stanford.edu/archives/sum2013/entries/feminism-autonomy/.

([11]) James S. Taylor, *Practical Autonomy and Bioethics* (New York: Routledge, 2009), 71. Indeed, Cudd says that *adaptive preference* is another term for deformed desire. Cudd, *Analyzing Oppression,* 180–181.

([12]) Jon Elster, “Sour Grapes—Utilitarianism and the Genesis of Wants,” in *Utilitarianism and Beyond*, edited by A. Sen and B. Williams (Cambridge: Cambridge University Press, 1992), 219–238; Elster, *Sour Grapes: Studies in the Subversion of Rationality* (Cambridge: Cambridge University Press, 1983).

([13]) Donald Bruckner, “In Defence of Adaptive Preferences,” *Philosophical Studies* 142 (2009): 307–324, 309.

([14]) Serene J. Khader, *Adaptive Preferences and Women’s Empowerment* (New York: Oxford University Press, 2011), 10.

([15]) See, e.g., Amartya Sen, “Gender Inequality and Theories of Justice,” in *Women, Culture, and Development: A Study of Human Capabilities*, edited by M. Nussbaum and J. Glover (Oxford: Clarendon Press, 1995), 259–273; Martha Nussbaum, *Women and Human Development: The*

Capabilities Approach (Cambridge: Cambridge University Press, 2000); Nussbaum, "Adaptive Preferences and Women's Options," *Economics and Philosophy* 17 (2001): 67–88. Note that the phenomenon of adaptive preference formation was invoked by Elster, Sen, and Nussbaum to critique preference utilitarianism.

([16]) I am grateful to Daniel Silvermint for discussion of this point and to Daniel Silvermint and Daniel Weinstock for helpful conversations about Khader's book.

([17]) Khader, *Adaptive Preferences and Women's Empowerment*.

([18]) Khader, *Adaptive Preferences and Women's Empowerment*, 51.

([19]) For reflective endorsement see Marilyn Friedman, *Autonomy, Gender, Politics* (New York: Oxford University Press, 2003). For nonalienation see John Christman, *The Politics of Persons: Individual Autonomy and Socio-historical Selves* (Cambridge: Cambridge University Press, 2009).

([20]) For a detailed explanation of this distinction, see Catriona Mackenzie and Natalie Stoljar, "Introduction: Refiguring Autonomy," in Mackenzie and Stoljar, *Relational Autonomy*, 3–34; Stoljar, "Feminist Perspectives on Autonomy." There are additional ways of dividing up theories of autonomy. For example, procedural theories may be either *internalist,* requiring only internal psychological conditions to spell out autonomy, or *externalist*, requiring additional historical conditions that are external to agents' present internal psychological states. Alfred Mele, *Autonomous Agents: From Self-Control to Autonomy* (New York: Oxford University Press, 1995). Other externalist theories are not procedural and employ, for instance, the notion of adequate options (Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1988)) or socio-relational conditions (Marina Oshana, *Personal Autonomy in Society* (Aldershot, UK: Ashgate, 2006)). The latter externalist positions may also require a background moral theory to spell out precisely the external conditions that are incompatible with autonomy. Thus, externalist theories are often substantive as well. I leave these complexities aside here.

([21]) Uma Narayan, "Minds of Their Own: Choices, Autonomy, Cultural Practices and Other Women," in *A Mind of One's Own. Feminist Essays on Reason and Objectivity*, second edition, edited by L. Antony and C. Witt (Boulder, CO: Westview, 2002), 418–432.

([22]) H. E. Baber, “Adaptive Preference,” *Social Theory and Practice* 33 (2007): 105–126.

([23]) Khader seems to have recently modified the objection saying that adaptive preferences “sometimes reveal—though need not reveal—compromised autonomy.” Serene Khader, “Must Theorising about Adaptive Preferences Deny Women’s Agency?” *Journal of Applied Philosophy* 29 (2012): 302–317, 312.

([24]) I leave aside structural and internalist forms of procedural theory, such as that offered by Harry Frankfurt, *The Importance of What We Care About* (New York: Cambridge University Press, 1988). For Frankfurt, a sufficient condition of local autonomy is (appropriately understood) endorsement or wholehearted identification at a time with a preference or desire. There are objections to this kind of theory. See, e.g., Stoljar, “Feminist Perspectives on Autonomy.” Moreover, as David Zimmerman, “Making Do: Troubling Stoic Tendencies in an Otherwise Compelling Theory of Autonomy,” *Canadian Journal of Philosophy* 30 (2000): 25–53, 25, points out, Frankfurt’s position risks being committed to an unpalatable consequence, namely, that “acting freely is a matter of ‘making do, ’ that is of bringing oneself to be motivated to act in accordance with the feasible, so that personal liberation can be achieved by resigning and adapting oneself to necessity.” Zimmerman’s interesting point needs further examination, but it does suggest that Frankfurt’s theory is the wrong place to look for an explanation of the nonautonomy of adaptive preferences.

([25]) See, e.g., Christman, *Politics of Persons.*

([26]) See, e.g., Raz, *Morality of Freedom*.

([27]) I argued for this position in Stoljar, “Autonomy and the Feminist Intuition.”

([28]) Elster, “Sour Grapes,” 219. See also David Zimmerman, “Sour Grapes, Self-Abnegation and Character Building,” *Monist* 86 (2003): 220–241; Bruckner, “In Defence of Adaptive Preferences”; Ben Colburn, “Autonomy and Adaptive Preferences,” *Utilitas* 23 (2011): 52–71, for explications of this idea.

([29]) Elster, “Sour Grapes,” 221.

([30]) Elster, “Sour Grapes,” 224.

([31]) Zimmerman, “Sour Grapes,” 221–222, notes that the process could be “sub-personal” rather than unconscious. I note this possibility but do not consider it here.

([32]) If it is not intentional action, it seems that it cannot be autonomous because even if intentional action is not sufficient for autonomous action, it is necessary.

([33]) Elster, “Sour Grapes—Utilitarianism and the Genesis of Wants,” 225.

([34]) Bruckner, “In Defence of Adaptive Preferences”; see also Colburn, “Autonomy and Adaptive Preferences,” 57.

([35]) Bruckner, “In Defence of Adaptive Preferences,” 314–315.

([36]) Bruckner, “In Defence of Adaptive Preferences,” 318–319.

([37]) Bruckner, “In Defence of Adaptive Preferences,” 319.

([38]) Christman, *Politics of Persons.*

([39]) Christman, *Politics of Persons*, 124–125. Christman quotes from Shane Phelan, *Getting Specific: Postmodern Lesbian Politics* (Minneapolis: University of Minneapolis Press, 1996), 52–53.

([40]) Christman, *Politics of Persons*, 155–156.

([41]) Christman, *Politics of Persons*, 144.

([42]) Paul Benson, “Autonomy and Oppressive Socialization,” *Social Theory and Practice* 17 (1991): 385–408.

([43]) I am grateful to Mark Piper for suggesting this possibility.

([44]) Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988); Mele, *Autonomous Agents*.

([45]) Khader, *Adaptive Preferences and Women's Empowerment*, 74. Khader goes on to ask whether various conditions mentioned in procedural theories of autonomy would count adaptive preferences as nonautonomous, but she does not explore procedural independence. She does discuss Elster's idea that adaptive preference formation is an unconscious process that operates behind the agent's back under the heading "life-planning as personal history" (85). However, her objection to this possibility seems off the mark. She writes, "Thinking of [adaptive preferences] this way commits us to a dubious metaphysical position…that there is one authoritative narrative about why a person forms a preference at the moment that it happens" (85). I do not believe that Elster is presupposing this metaphysical commitment. He is claiming that there is a psychological and theoretical difference between unconscious (adaptive) causal mechanisms and deliberate, conscious and planned ones.

([46]) Dworkin, *Theory and Practice of Autonomy*, 18.

([47]) Christman, *Politics of Persons*, 147, 155.

([48]) John Christman, "Liberalism and Individual Positive Freedom," *Ethics* 101 (1991): 343–359, 354.

([49]) Christman, "Liberalism and Individual Positive Freedom," 353.

([50]) A recent discussion of adaptive preference formation explicitly adopts the position that adaptive preference formation is an unconscious process that violates procedural independence. Ben Colburn, "Autonomy and Adaptive Preferences," argues that an independence test fails when the formation of a preference is subjected to *covert influence*. Like subliminal influence, covert influence is a mechanism that produces desires in agents through a process in which the agent is not aware of the causal explanation of her preferences.

([51]) Khader, *Adaptive Preferences and Women's Empowerment*, 78.

([52]) Khader, *Adaptive Preferences and Women's Empowerment*, 81.

([53]) Cudd, *Analyzing Oppression*, 183.

([54]) Cudd, *Analyzing Oppression*, 126–129. I am therefore in broad agreement with Khader when she says that procedural theories need to be "[supplemented] with a theory of the good." Khader, *Adaptive Preferences and Women's Empowerment*, 95.

([55]) Elster, "Sour Grapes," 228.

([56]) Nussbaum, “Adaptive Preferences and Women’s Options,” 68–69.

([57]) Nussbaum, “Adaptive Preferences and Women’s Options,” 68–69.

([58]) Nussbaum, “Adaptive Preferences and Women’s Options,” 79.

([59]) Khader, *Adaptive Preferences and Women’s Empowerment*, 51.

([60]) Khader, *Adaptive Preferences and Women’s Empowerment*, 95–106.

([61]) Khader, *Adaptive Preferences and Women’s Empowerment*, 99–102, also discusses other possible substantive approaches to spelling out the notion of autonomy such as “substantive autonomy as being motivated by good norms.” This label is misleading because *being motivated by good norms* does not name a conception or theory of autonomy but rather a piece of evidence that a theory of autonomy will have to explain. For example, Stoljar, “Autonomy and the Feminist Intuition,” claims that many feminists think that, when agents are motivated by false and oppressive norms, the agent’s autonomy is called into question. I go on to make a conditional claim: *if* we accept this intuition, *then* a strong substantive theory of autonomy will need to be invoked to explain it. As mentioned already, my discussion at the time did not properly distinguish between two possible substantive theories: a strong substantive account that is content based; and a (strong) normative competence account. As I argue here, I still think that a substantive account of autonomy is needed to explain why desires for one’s own oppression seem to be autonomy impairments. But I do not now think this has to be a strong substantive account that is content based.

([62]) Raz, *Morality of Freedom*, 369.

([63]) Raz, *Morality of Freedom*, 374–376.

([64]) Raz, *Morality of Freedom*, 378.

([65]) Khader, *Adaptive Preferences and Women’s Empowerment*, 102. Emphasis added.

([66]) Raz, *Morality of Freedom*, 411.

([67]) I have based my discussion on one possible reconstruction of Khader’s argument. Her remarks are very brief, so I may have misunderstood what she has in mind. Here is another possible reconstruction: she seems to take Raz’s position as equivalent to a strong substantive view that employs normative constraints on the content of preferences. She claims that adaptive preferences cannot be

defined using contents alone. (This is correct; a preference for bullfighting may be or may not be adaptive and nonautonomous.) Hence she may be saying that adaptive preferences do not correspond to the category of nonautonomous preferences on Raz's account because the latter is a content-based account. I think however that the analysis of Raz as offering a content-based account is mistaken. As we saw, he thinks that autonomous agents can have preferences with bad or immoral content. Hence his account is substantive in a different sense from strong substantive accounts. The moral constraints are derived from the notion of adequate options, not from the immorality of the content of agent's preferences.

([68]) Khader, *Adaptive Preferences and Women's Empowerment*, 103.

([69]) Khader, *Adaptive Preferences and Women's Empowerment*, 103.

([70]) Cf. Ann Cudd, "Review of Khader's *Adaptive Preferences and Women's Empowerment*," *Notre Dame Philosophical Reviews*, 2012, http://ndpr.nd.edu/news/27280-adaptive-preferences-and-women-s-empowerment/.

([71]) Cudd, *Analyzing Oppression*, 146–153.

([72]) Cudd, *Analyzing Oppression*, 150.

([73]) Narayan, "Minds of Their Own."

([74]) Baber, "Adaptive Preference."

([75]) Khader, *Adaptive Preferences and Women's Empowerment*, 100–101.

([76]) Narayan, "Minds of Their Own," 420.

([77]) Narayan, "Minds of Their Own," 420–421.

([78]) Baber, "Adaptive Preference," 113–114.

([79]) Baber, "Adaptive Preference," 114.

([80]) Preferences are understood as behavioral dispositions, not occurrent feelings. Baber, "Adaptive Preference," 312.

([81]) Baber, "Adaptive Preference," 111.

([82]) Baber, "Adaptive Preference," 114.

([83]) Stoljar, "Autonomy and the Feminist Intuition."

([84]) Nussbaum, “Adaptive Preferences and Women’s Options,” 73, discusses this line of thought while referencing John Harsanyi’s distinction between *manifest* and *true* preferences.

([85]) See, e.g., Meyers, *Self, Society and Personal Choice*, 47.

([86]) Henry Richardson, “Autonomy’s Many Normative Presuppositions,” *American Philosophical Quarterly* 38 (2001): 287–303.

([87]) Richardson, “Autonomy’s Many Normative Presuppositions,” 292. One could interpret Richardson as claiming that agents who act in the service of right reasons are not *really* self-deceived because acting on right reasons is doing what rational agents *really* want to do. Sarah Buss, “Autonomous Action: Self-Determination in the Passive Mode,” *Ethics* 122 (2012): 647–691, 666, captures the gist of this kind of analysis of autonomy when she writes: “Every rational agent wants to do what she (really) has reason to do; so, insofar as her action is not (adequately) responsive to (the real) reasons for and against acting this way, it does not (adequately) express her defining desire; so there is an important respect in which this action is not attributable to her [i.e., not autonomous].”

([88]) Rosen, “Who Wears the Pants in This Economy?”

([89]) Cudd, *Analyzing Oppression,* 181, points out that there is a difference between desires to subject oneself to oppression and desires for social roles that are the effects of oppression: “It is not that [the oppressed] will prefer oppression to justice or subordination to equality, rather they will prefer the kinds of social roles that tend to subordinate them.”

([90]) The Srey Mom example is taken from Nicholas D. Kristof, “Bargaining for Freedom,” *New York Times*, January 21, 2004.

([91]) Khader, *Adaptive Preferences and Women’s Empowerment*, 130, argues that preferences such as that for malnourishment could enhance the interests or well-being of the agent, for example, “self-depriving behavior often elicits actual rewards.”

([92]) Cf. Elster, “Sour Grapes,” 221.

([93]) Khader, *Adaptive Preferences and Women’s Empowerment*, 98.

([94]) Khader, *Adaptive Preferences and Women’s Empowerment,* 100.

([95]) Nussbaum, “Adaptive Preferences and Women’s Options,” 74; Baber, “Adaptive Preference,” 114, 126.

([96]) Mele, *Autonomous Agents*.

([97]) See, e.g., Khader, *Adaptive Preferences and Women's Empowerment*, 33.

([98]) Khader, *Adaptive Preferences and Women's Empowerment*, 11.

([99]) Narayan, "Minds of Their Own," 422, refers to a remark of Catherine MacKinnon that women's agency is "pulverized."

([100]) For inadequate moral options, see Raz, *Morality of Freedom*. For socio-relational conditions, see Marina Oshana, *Personal Autonomy in Society*.

([101]) Oshana, *Personal Autonomy in Society*, 2.

([102]) Khader, *Adaptive Preferences and Women's Empowerment*, 103.

([103]) Khader, *Adaptive Preferences and Women's Empowerment*, 104.

([104]) See, e.g., Raz, *Morality of Freedom*.

([105]) Khader, *Adaptive Preferences and Women's Empowerment*, 101.

([106]) Khader, *Adaptive Preferences and Women's Empowerment,* 101.

([107]) Khader seems to thinks that noncoercive interventions are permissible to promote, for example, deliberative aspects of agency, but these would fall short of policies whose aim is to change conceptions of the good.

([108]) Catriona Mackenize, "Relational Autonomy, Normative Authority and Perfectionism," *Journal of Social Philosophy* 39 (2008): 529.

([109]) I am grateful to Catriona Mackenzie for many discussions that have helped me think through these ideas.

([110]) Cudd, *Analyzing Oppression*, 181.