DE-MORALIZATION AS EMANCIPATION: LIBERTY, PROGRESS, AND THE EVOLUTION OF INVALID MORAL NORMS *

By Allen Buchanan and Russell Powell

Abstract: Liberal thinkers of the Enlightenment understood that surplus moral constraints, imposed by invalid moral norms, are a serious limitation on liberty. They also recognized that overcoming surplus moral constraints — what we call proper de-moralization — is an important dimension of moral progress. Contemporary philosophical theorists of liberty have largely neglected the threat that surplus moral constraints pose to liberty and the importance of proper de-moralization for human emancipation. This essay examines the phenomena of surplus moral constraints and proper de-moralization, utilizing insights from biological and cultural evolutionary thinking.

KEY WORDS: de-moralization, surplus moral constraints, invalid moral norms, social experimentation, evolution, Environment of Evolutionary Adaptation (EEA), moral progress

I. INVALID MORAL NORMS AS CONSTRAINTS ON LIBERTY

Liberal thinkers of the Enlightenment believed that throwing off the mental shackles of superstition was crucial for human emancipation. They also believed that human emancipation was central to moral progress. They understood that liberty can be curtailed, not just by coercion or other external forces, but also by misguided internal constraints of conscience. They also understood that although morality necessarily involves internal constraints on liberty, people can mistakenly believe that morality constrains them when it does not. In other words, they possessed the idea of **surplus moral constraints**; and they were aware that surplus moral constraints are a major limitation on liberty and one that ought to be overcome. Enlightenment thinkers tended to take liberty seriously, but they were also rightly attentive to other ways in which surplus moral constraints can be detrimental to human well-being. The focus of this essay is on emancipation from surplus moral constraints, and hence on the expansion of liberty.

---

1 The fact that abandoning a moral norm would increase liberty does not, of course, show that this change constitutes moral progress. Abandoning valid moral norms might increase liberty, but would not be progressive. The topic of this essay is proper demoralization — abandonment of invalid moral norms. So far as invalid moral norms constrain liberty, they do so without justification, and removing this constraints counts as moral progress, other things being equal, for two reasons: first, because it is a case of remedying a defective understanding about what morality requires; and second, because (at least from a liberal standpoint), unjustifiable constraints on liberty are to be avoided.
human liberty, but in the recognition that proper de-moralization can contribute not only to greater liberty but to greater well-being.

A. Surplus moral constraints

By “surplus moral constraints” we mean limitations on liberty in the form of invalid moral norms. The term “surplus” is chosen to indicate that although morality essentially involves constraints, there are some constraints imposed in the name of morality that exceed those that morality properly imposes.

For those who regard invalid moral norms as valid, surplus moral constraints have an internal aspect: they are limitations on an individual’s liberty imposed by conscience. Such internal constraints may be accompanied by external constraints, including not only various sanctions for violating the moral norms in question (such as punishment), but also more subtle but nonetheless powerful forms of social pressure. And insofar as individuals tend to care greatly about what others think of them, and in particular about whether they are regarded as having morally good character and whether their behaviors are perceived as socially acceptable, such external pressures can reinforce the constraints of conscience. Such internal constraints may persist even when external sanctions and pressures have abated.

Contemporary classical liberals tend to focus chiefly on external constraints and in particular on curtailments of liberty wrought by the coercive power of the state, whether through law or acts of raw despotism. Liberal Enlightenment thinkers had a more balanced view: they understood that both internal and external constraints on liberty matter.

In a previous work, we have noted another lacuna in much contemporary liberal thought (in this case spanning the spectrum from classical liberalism to various forms of welfarist liberalism): the concept of moral progress. There has been little systematic thinking in these literatures about the concept of moral progress, in spite of the fact that it used to lie

2 Morality constrains liberty in two ways. First, one of the most distinctive and important features of morality is that it constrains the satisfaction of desires and imposes limits on the pursuit of interests, especially, but not exclusively self-interest. Second, “ought” judgments, even positive ones as opposed to prohibitions, entail limitations on liberty: If I ought to do X, then I ought not to refrain from doing X and I ought not to do that which prevents me from doing X.

3 These “social desirability” effects are pronounced and impose substantial biases in moral psychology research paradigms that rely on self-report. For a review of this effect and its implications for ethics research, see Donna Randall and Maria Fernandes, “The Social Desirability Response Bias in Ethics Research,” *Journal of Business Ethics* 10, no. 11 (1991): 805–817.

4 Focusing only on external constraints not only obscures the fact that invalid moral norms, if internalized, can unnecessarily limit liberty; it also abets a failure to see that false factual beliefs can limit liberty and at great cost. Allen Buchanan, “Prisoners of Belief,” *Oxford Handbook of Freedom*, David Schmidt, ed. (New York: Oxford University Press, forthcoming).
at the core of liberal thought from the Enlightenment through the nineteenth century.\(^5\) Perhaps if contemporary classical liberals thought more about moral progress they would pay more attention to surplus moral constraints as limitations on liberty. Why? Because emancipation from surplus moral constraints — what we term “proper de-moralization” — is an important type of moral progress, at least for any conception of moral progress that values liberty.\(^6\) Proper de-moralization occurs when people come to see that behavior thought to be morally impermissible is morally neutral or even good. Because surplus moral constraints are unnecessary limitations on liberty, escaping their thrall is a form of emancipation. To the extent that surplus moral norms significantly interfere with liberty, welfare or other important moral values or duties (such as those associated with justice or beneficence), the de-moralization of those norms, all else being equal, constitutes moral progress.

This essay explores the notion of surplus moral constraints as an important type of limitation on liberty, and in doing so corrects both the contemporary classical liberal tendency to focus chiefly if not exclusively on external limitations on liberty and the tendency of liberals of all stripes to neglect the topic of moral progress. Our goals are first, to show that proper de-moralization is an important form of moral progress as increased liberty or emancipation, and second, to begin to develop an evolutionarily informed account of how improper moralization, in the form of surplus moral constraints, arises and is sustained, and how it can be overcome.

B. Proper de-moralization as emancipation from surplus moral constraints

There are many forms or types of moral progress. These include not just better conformity to valid moral norms of which we are already aware, but also improvements in understandings of the moral virtues, of responsibility, and of moral standing and moral status. For example, consider a conception of honor that largely reduces it to chastity for females and the willingness to respond with violence to perceived insults for males. Surely, it is an instance of moral progress when that conception of honor gives way to one that emphasizes fidelity to one’s principles, and dignity as involving self-restraint.\(^7\) Similarly, the shift from conceptions of responsibility that reduce responsibility to causality, toward those that take into account an agent’s mental state (including beliefs and intentions) is also


progressive. Notice that this shift will involve (proper) de-moralization: people who embrace the new understanding of responsibility will no longer regard certain behavior as morally wrong — for example, cases where one person accidentally and non-negligently kills another while acting with perfectly acceptable intentions. One important question not pursued in this essay is whether de-moralization precedes and fosters improved moral concepts or vice versa.

Likewise, the recognition that at least some nonhuman animals have moral standing — that they matter morally in their own right — is surely an important instance of moral progress, too, as is the recognition that all persons have the same basic moral status, including the possession of certain moral rights. These shifts involve proper moralization: coming to recognize moral constraints that were previously overlooked. By focusing in this essay on proper de-moralization, we do not wish to suggest that it is the only or even the most important form of moral progress.

Instances of proper de-moralization abound and play a prominent role in the catalog of morally progressive developments. Profit-seeking, lending money at interest, premarital sex, homosexual behavior, interracial marriage, masturbation, refusal to die “for King and country,” and virtually all instances of resistance to government authority were once widely thought to be immoral, but are no longer so regarded by many people.

We will take it for granted, because we are assuming a broadly liberal moral perspective, that these are all cases of proper de-moralization — that, at least from a secular liberal point of view, beliefs that these behaviors are morally wrong per se or that they warrant institutionalized punishment were unjustified, and that coming to realize the falsity of these beliefs is an instance of moral progress. Later, we will suggest that one of the major points of contention between liberals and conservatives is a disagreement about how reliably one can determine when a given instance of de-moralization is a case of proper de-moralization. If one is to develop an account of the relationship between liberty and de-moralization, one must be able to determine when internalized moral norms, and external sanctions for violating those norms, are instances of surplus constraints and when they are not.

For now, it suffices to emphasize that internalized improper moralization can significantly limit an individual’s options for acting, and that it
does so in ways that entail great material and psychological costs, even when external sanctions have been removed. For example, if one is homosexual but has internalized a norm that brands all homosexual behavior as sinful and morally abhorrent, and one adheres to this norm, then one may experience great psychological suffering due to the self-inflicted frustration of one’s most basic needs. These needs include not just sexual satisfaction, but also the intimacy and deep attachment of partnership — needs that can be frustrated even after such behavior has been decriminalized and other formal external sanctions have been removed, and, if the internalized norms of conscience have sufficient psychological inertia, even in the absence of any external constraints at all. Alternatively, if one violates the moral prohibition of homosexuality and engages in homosexual acts, then one may experience severe shame and guilt.

Similarly, if one believes that any perceived insult to one’s honor requires violent retaliation, one may put oneself at lethal risk by initiating a duel, or feel compelled to engage in other violent behavior that runs contrary to one’s basic values, thereby risking self-alienation, unnecessary trauma, and guilt. Likewise, if people in a society refrain from profit-seeking or from lending money at interest on the grounds that these vital economic behaviors are immoral, the result may be the perpetuation of a state of economic underdevelopment, with disastrous consequences for human welfare, liberty, and justice. One final example: in a society in which it is widely believed that it is wrong for women to engage in independent economic activity outside the home, women will not only be barred from important paths to flourishing; they will also remain so dependent upon men and have so little influence on the political process that there may be little prospect of eliminating the grosser abuses of patriarchal society, including domestic abuse and honor killings. In short, the costs of surplus moral constraints, and accordingly the benefits of proper de-moralization, can be extraordinarily high. That is why proper de-moralization is an important form of moral progress.

C. How does proper de-moralization occur?

One cannot assume that the limits of proper de-moralization have already been reached. Assuming that further instances may turn out to be warranted, it is important to understand how proper de-moralization comes about. Our aim here is not to supply such an account, but rather to indicate what some of its core elements are likely to be. What we say will be speculative, but should help to guide empirical research into how proper de-moralization occurs. For this limited purpose, we will consider three

---

cases in which widely accepted moral prohibitions were eventually largely abandoned in some societies: prohibitions on sex between unmarried people, on gay and lesbian sex, and on interracial marriage. “Abandonment” here covers several distinct but causally interrelated phenomena: the abolition of institutionalized punishment for violating the norm, a significant decline in compliance with the norm, and a widespread willingness to acknowledge that the norm is not valid. Abandonment in this sense is, of course, compatible with some segments of the population continuing to adhere to the norm and to affirm its validity.

Abandonment of the norm against sex between unmarried individuals, that is, against fornication, appears to have been fueled by the confluence of several major social changes. These include (1) the decline of religious authority (along, perhaps, with the emergence of more permissive interpretations of religious duties that de-emphasize the importance of chastity), (2) the growth of urbanization and with it the increase in privacy that the anonymity of life among large numbers of strangers brings, and (3) the availability of effective contraception. All three of these factors correlate with abandonment of the norm, and there is reason to believe that they are causative, though their relative causal powers are unclear.

Items (2) and (3) are likely to have played significant causal roles because both involved a significant lowering of the tangible costs of violating the norm. Where effective contraception is available, single women can have sex and yet avoid the economic costs and social stigma of bearing children outside marriage. Insofar as people adhere to moral norms mindfully, that is, because they recognize the negative consequences that flow from violating them, a significant reduction in the costs associated with casual sex would have tended to attenuate compliance with and enforcement of anti-fornication norms, other things being equal. Urbanization greatly reduced the costs of violating the norm in another way: in an urbanized environment, individuals are not as vulnerable to reputational loss and other forms of social stigma as they are when they live and work in smaller groups, under conditions in which defection from norms is easily detected and can be effectively sanctioned.

In addition, it may have been the case that a growing tendency to question norms, or at least norms that are onerous to comply with, rather than to follow them blindly, played a role in the demise of the norm against fornication. More specifically, the abandonment of this norm may have been due in part to a more general tendency in modern as opposed to traditional cultures to question norms that are costly to comply with but which do not obviously confer any off-setting benefit in terms of human welfare.

Finally, the shift away from anti-fornication norms can also be attributed to the decline of religious authority and the rise of more permissive interpretations of religious strictures on sexual activity, which likely resulted from multiple causes. Given the extraordinary strength of sexual desires, once the costs of noncompliance with the prohibition were dramatically
lowered by factors (2) and (3), systems of moral or religious beliefs that required adherence to the prohibition would have been seriously disadvantaged in the competition for people’s allegiance. In other words, there may have been cultural selection pressures for attitudes and beliefs that rejected the prohibition outright or at least downplayed its importance, further reducing the costs of noncompliance.

The point is that, as much empirical research has shown, norm compliance, especially when it involves thwarting strong desires, can often only be achieved if there are effective costs for noncompliance. And the stronger the motivation for noncompliance, the greater the need to impose significant costs on noncompliance if the norm is to be effective. Given that human sexual desire is among the strongest of motivations (constraining it within the bounds of marriage has always been difficult and has never been fully achieved), and given that Homo sapiens is at best a “nominally monogamous” species, once the costs of noncompliance were significantly reduced, fornication would have become pervasive. And once fornication became pervasive, and incentives for costly enforcement became weak, this is likely to have further shifted public opinion regarding the norm, which in turn would have resulted in the repeal or refusal to enforce anti-fornication laws.

The abandonment of the moral prohibition on gay and lesbian sex seems to have followed a somewhat different pattern. The decriminalization of consensual same-sex sexual activity appears to have been a tipping point, because it dramatically lowered one important cost of gays and lesbians acknowledging their sexual identities. In an environment in which homosexual relations were no longer subject to legal penalties, gays and lesbians were freer to express their sexuality in socially acceptable ways and to integrate their sexual lives with the rest of their activities. This change fundamentally altered the experience of heterosexual people: they came to see that some of their relatives, friends, co-workers, and coreligionists were gay or lesbian. The result was that increasing numbers of heterosexual people came to understand that gays and lesbians are “just like us,” only homosexual. Once this transformation of experience occurred on a sufficiently large scale, the erosion of the prohibition gathered momentum. None of this is to deny, however, that a shift toward more tolerant attitudes toward gays and lesbians played a role in mobilizing support for de-criminalization. The point, rather, is that once de-criminalization occurred, a different kind of social experience emerged that accelerated the transition toward greater acceptance of gays and lesbians.12

12 The role of changing social experience in proper-demoralization is complex and warrants further investigation than we can give it in this essay. Liberal societies allow initially local “experiments of living” that can challenge current norms and also provide conditions of freedom of expression and association that can enable norm change to spread.
DE-MORALIZATION AS EMANCIPATION

The abandonment of the norm against interracial marriage shares some features with the previous two examples, but also appears to differ significantly from both. Here the abandonment of the prohibition — or perhaps more accurately its relaxation, since interracial marriages are still relatively rare — seems to have depended on a more profound prior alteration: the recognition of the equal basic moral status of blacks and other non-white minorities. In other words, if there had not already been moral progress in the form of rejecting an understanding of basic moral status that excluded blacks and other minorities, it is inconceivable that people would have come to see that interracial marriage was in fact morally permissible. As long as blacks and other minorities were regarded as so inferior by nature as to occupy a lower rung of moral status — something akin to that of the more intelligent nonhuman animals — they could not be seen as suitable marriage partners for whites. So, in this case, proper de-moralization seems to have depended on the prior achievement of a different, even more momentous kind of moral progress.

II. Evolutionary Accounts of Surplus Moral Constraints

A. Why do unjustifiable and destructive moral norms arise and persist?

So far, we have appealed to widely shared intuitions (at least among liberals broadly defined) to make the point that there are genuine instances of proper de-moralization; and we have argued that proper de-moralization is an important form of moral progress, at least from the perspective of any morality that highly values liberty, because surplus moral constraints can seriously and yet unnecessarily curtail liberty. At this point, a puzzle looms: Some supposed moral constraints are so costly to obey and yet so clearly irrational, destructive or bigoted, that the question arises as to how they came about and how they could persist. Consider, for example, Biblical prohibitions on planting more than one kind of crop in a field or wearing garments with more than one kind of fiber, on women trimming the edges of a man’s beard, on simmering a young goat in its mother’s milk, or prohibitions in some cultures on eating fish that result in avoidable malnourishment, or norms that require men to gorge on protein-rich foods while depriving women of them.

Such norms seem to be irrational limitations on liberty at best and destructive of human welfare at worst. How did these norms come to be institutionalized and internalized by large numbers of people, and to persist despite their costs and apparent lack of sound grounding in morality or prudence? Having a theory of the origination and persistence of improper moralizations may help us to develop effective strategies for overcoming them. Evolutionary explanations of morality may provide resources for constructing such a theory, and thus it is to these explanations that we now turn.
B. The evolution of morality

Understanding the evolutionary origins of surplus moral constraints first requires an understanding of the evolutionary origins of moral constraints in general. Evolutionary theorists have long regarded key aspects of human morality as adaptations to ecological conditions of the Middle to Late Pleistocene, the so-called Environment of Evolutionary Adaptation, or EEA — the time when core components of human moral psychology supposedly crystalized. An adaptation is a trait that came to exist because it contributed to the reproductive success (fitness) of an organism, group, or system at some point in the past. Adaptations may be maladaptive or fitness-neutral in the modern environment. Further, a trait may be currently adaptive and yet immoral or otherwise undesirable, at least in any situation in which maximizing reproductive fitness is not of supreme value. For example, a norm that treats stealing as a capital offense may, in certain environments, be adaptive at the individual or cultural group level, but nonetheless immoral.

In brief, the evolutionary explanation of morality goes as follows. Human morality, including evaluative tendencies and dispositions to acquire and adhere to cultural moral norms, was selected in the EEA because it promoted cooperation within the group and successful competition with other groups. Moralities achieved this in at least two ways. First, they fostered altruism, coordination, and peaceful conflict-resolution within the group, and in some cases they directly contributed to reproduction, either by encouraging certain types of mating behavior or by reducing risks to those who could reproduce. Second, they encouraged members of the group to act toward out-group members in aggressive, distrustful or exploitative ways, which reduced the threat of biological and social parasites, defended against out-group aggression, and facilitated the expropriation of the resources of other groups. Evolutionary risk management suggests that in the EEA, where an infrastructure of political institutions and social practices for facilitating peaceful cooperation among groups was lacking, the risks of biological and social parasitism would have often outweighed whatever benefits might be gained from engaging in welcoming behavior toward out-group members.

13 The Pleistocene is the period lasting from approximately 1.7 million to 10,000 years ago, terminating with the beginning of the Neolithic Period, in which agriculture began to be established.
14 For a review of the theoretical and empirical evolutionary link between human social psychology and parasite stress, see Corey Fincher and Randy Thornhill, “Parasite-stress promotes in-group assortative sociality: The cases of strong family ties and heightened religiosity,” Behavioral and Brain Sciences 35 (2012): 61–119.
15 For a more detailed discussion of the ecological conditions under which “thick” moral considerations toward in-group members and “thin” moral considerations toward out-group members were likely to evolve, see Buchanan and Powell, “Toward a Naturalized Theory of Moral Progress,” supra note 4. See also Buchanan and Powell, The Evolution of Moral Progress: A Biocultural Theory (New York: Oxford University Press, forthcoming 2018).
In short, in the EEA there were strong selective pressures for the coevolution of a psychology and culture that supported “thick” moral relations among in-group members, but xenophobia, distrust, and other aversive attitudes toward out-group members. This “thin” moral consideration of out-group members culminated in a refusal to accord them equal moral status and, in extreme cases, the denial of any moral standing whatsoever. Individual and cultural susceptibilities to bigoted, xenophobic, and unjustified discriminatory norms may thus have adaptive roots.

Not all adaptive moral norms are so straightforwardly functional. Some moral norms that appear irrational or arbitrary are in fact mechanisms for delineating group membership, coordinating group action, signaling cooperative intent, and/or maintaining group cohesion. This is, in fact, how religious rituals are now widely understood.16 Norms requiring distinctive attire, body modification, hair growth, or ritual participation draw a boundary around the group in such a way as to reduce the risk of exposure to out-group biological parasites and “alien” ideas, norms, and behaviors that could destabilize cooperation within the group. The Biblical requirement of male circumcision and the prohibition on simmering a young goat in its mother’s milk, mentioned earlier, may be instances of this phenomenon: circumcision is a costly device to signal a distinctive group identity, and the prohibition on simmering a goat in its mother’s milk banned participation in a ritual of the competing Canaanite religion. Various cultural dietary restrictions may also serve to demarcate group boundaries and serve as similar costly signals of cooperative intent.

Other seemingly irrational moral norms have an even less direct effect on the fitness of cultural groups. Because cooperation, at least on a fairly large and complex scale, requires coordination through the following of norms, and because internalization of norms improves compliance and reduces the costs of achieving it, it is important for individuals to develop the disposition to follow moral norms automatically, as it were. Given that this is so, the internalization of some norms may be functional only insofar as they contribute to the habit of obedience to authority, a habit that has significant fitness payoffs in the case of other, directly functional norms.

Here an analogy may be helpful. In basic military training, considerable time is devoted to what American soldiers used to call “monkey drills” — learning to execute rather complex movements on the parade ground that are of no use in combat. The standard explanation for why military authorities devote so much time and energy to such apparently functionless behavior is that it helps form the habit of immediate, unreflective obedience to orders. Similarly, some of the seemingly excessive and nonfunctional moral rules found in the Bible or in the taboos of pre-modern

societies recorded by anthropologists may be only indirectly functional: they may serve chiefly or exclusively to cultivate the disposition to follow supposedly authoritative norms, a disposition that can have considerable fitness benefits. In addition, as Norbert Elias has emphasized in his monumental book, *The Civilizing Process*, compliance with some apparently nonfunctional norms may promote cooperation and even reduce the incidence of violence if they serve as social signals of self-restraint, readily observable proxies for “prosocial” dispositions.\(^\text{17}\)

Given this *functional* understanding of the origins of morality as an evolutionary adaptation to the conditions of the EEA, we can begin to construct explanations that appear to illuminate the emergence and persistence of at least some instances of surplus moral constraint. In particular, as we have seen, we can begin to understand why many seemingly irrational and even outright bigoted moral norms are adaptations, or modern expressions of adaptive propensities, that evolved in the prehistoric ecological context of competing cultural groups. Further, the persistence of certain moral norms that are destructive and de-stabilizing in the modern human environment can be understood as (1) spillovers from an earlier environment in which they were evolutionarily (and perhaps morally) non-surplus, (2) the result of selection conflicts between levels of organization (e.g., social castes) in multilayered human societies, (3) the result of recalcitrant evolutionary collective action problems to which solutions are not readily found, and/or (4) the result of the susceptibility of cultural transmission to the spread of deleterious variants. We discuss each of these in turn.

**C. Surplus moral norms as the result of evolutionary mismatch**

Some apparently invalid and indeed destructive moral norms can be explained as “Pleistocene hangovers” — remnants of moral responses that made sense in the EEA but are unnecessary in and discordant with the current environment. Modern human ecology is far removed from the hunter-gather life that characterized the vast majority of the history of the human species and in which core components of human moral psychology evolved. Thus, as Pinker notes, some of “our ordeals come from a mismatch between the sources of our passions in evolutionary history and the goals we set for ourselves today.”\(^\text{18}\) Evolutionary moral mismatch can take two forms. The first is a mismatch between adaptive innate moral psychological dispositions and the modern environment; the second is a mismatch between evolved cultural moral norms and the modern environment.


An example of the first type is the mismatch between evolved leadership-followership psychology and the modern human environment. Leadership-followership psychology was critical to coordinating activities of prehistoric human groups, such as hunting, foraging, making war, and resolving internal disputes — particularly as human societies expanded to sizes that make strictly egalitarian decision-making cumbersome and inefficient in real-time intergroup conflicts. However, leadership-followership psychology may pose grave risks in the modern world of interstate conflicts involving powerful weaponry that can have profound consequences for generations far into the future. For instance, there is a well-documented human tendency to gravitate toward authoritarian, hawkish, masculine, and charismatic leaders in times of actual or perceived intergroup conflict.

In the EEA, such aggressive posturing and “rally-round-the-flag” effects may have been adaptive; but in the modern world of state-level brinkmanship with incredibly destructive weaponry in play, hawkish virtues may have devastating costs. Moreover, elites may exploit this prehistorical moral psychology by provoking intergroup conflicts or engendering perceptions of intergroup threat in order to consolidate power (see discussion below on within-group fitness conflicts).

Other evolutionary moral mismatches have more recent cultural evolutionary roots. For example, there are profound differences in cultural conceptions of honor between American populations in the North and those in the American South — in particular, Southerners are far more likely to respond to insults or affronts to their honor (especially indignities to their masculinity) with violence than are their northern counterparts, and they are more likely to believe that such violent reprisals are justified. These well-established populational differences are not due to innate psychological differences; they result, so far as we can tell, from the differential persistence of cultural moral systems that can be traced to particular ecologies and migration patterns. In particular, Scotch-Irish livestock herders were the predominant settlers of the South, whereas peasant farmers from Germany, England, and the Netherlands were the chief settlers of the North. Livestock herding is robustly associated with hyper-masculine, honor-based cultures around the world because, so the cultural evolutionary hypothesis goes, it typically occurs in rugged, lawless regions of countries where theft

---


20 Although early human societies are generally thought to have been rather egalitarian, subordination to the temporary authority of a powerful male (so-called “Big Men”) in times of armed conflict also seems to have been common. See Van Vugt et al., supra note 19.


and other forms of predation are commonplace. In such contexts, having a reputation for resorting to violence in case of attack is adaptive, and this remained so in the American South until fairly recently in U.S. history. As we have seen, in modern environments cultures of honor impose significant surplus moral constraints (and involve improper moralization as well). Imported into grand conflicts between powerful states, and combined with prehistoric leadership-followership psychology discussed above, a culture of honor can cause spiraling, destructive intergroup conflicts, and make peaceful resolutions harder to come by.

An example of a costly evolutionary moral mismatch that may implicate both innate dispositions and cultural moral norms concerns the treatment of homosexuality. It has been suggested that the prohibition on homosexual sex and, even more so, on homosexual partnership, may have been selected for because of its contribution to higher fertility rates in small, vulnerable societies whose survival depended upon achieving high fertility. In addition, this prohibition may have facilitated an efficient division of labor between men and women (big game hunting/warfare on the one hand, child care/local foraging on the other), and this specialization may have been crucial for human survival for the vast majority of human evolutionary history. If these explanations are correct, then prohibitions on homosexual sex and the enforcement of stereotypical gender roles are adaptations or expressions of adaptations that are no longer fitness-enhancing at the cultural group level due to changes in the selective environment (though they may enhance the fitness of subgroups within the larger population). In today’s world, population increases can be detrimental in some cases to a group’s economic success, and the ability of women to enter the workforce and to be able to compete for desirable positions has a significant positive impact on economic development. In modern human ecology, therefore, prohibitions on same-sex partnerships and the enforcement of strict gender roles do more harm than good — and this harm falls disproportionately on women and sexual minorities. Moreover, there may be no good moral reason to tolerate the net harm they produce. Such norms are therefore not only surplus evolutionary constraints, but also surplus moral constraints, from the standpoint of morality properly conceived.

Although norms can sometimes change rapidly, evolutionary investigations of honor culture and other norm-environment mismatches show that formerly adaptive norms can have substantial inertia, even when societies find themselves in ecological circumstances to which the norms are quite ill-suited. What explains this cultural staying power? One explanation is that maladaptive norms will not be winnowed out by selective

---

pressures, or will be winnowed out only very slowly, if they do not reduce fitness substantially. Yet even if maladaptive norms are drastically fitness reducing, they can still be highly resistant to selective shaping if they are “culturally entrenched” — a cultural analogue of one important form of developmental constraint in biology. A norm N is culturally entrenched if it is casually connected to other aspects of a cultural tradition web, such that N cannot be altered without the costly alteration of many other aspects of the web, resulting in N’s selective preservation in an evolving cultural system. Like genes that lie “downstream” in a developmental cascade that leads to the phenotype, norms that lie on the periphery of a cultural web (such as driving on the right side of the road) can readily be modified without collateral developmental costs, and thus they are not culturally entrenched or only weakly so. In contrast, just as developmentally “upstream” genes have numerous downstream effects on other genes in the developmental cascade, norms that occupy a more central place in the cultural web (such as those implicating group identity) will be more recalcitrant to adaptive modification across generations of cultural users — and further adaptive mechanisms may canalize these norms, or make them robust against environmental inputs. The concept of cultural entrenchment can help to explain why certain cultural traditions die hard, even when they are evolutionarily superfluous. The conservative’s mistake noted above is to think that virtually all existing norms are culturally entrenched or that we can never know which ones are.

D. Surplus moral norms as special interest adaptations

Because selection can work simultaneously in different directions across levels of organization, a trait that is adaptive at one level can be deleterious at other levels. In the case of paradigmatic biological individuals comprised of lower-level individuals — such as multicellular organisms comprised of cells, or colonies comprised of organisms — there are mechanisms in place to ensure that lower-level adaptations deleterious to higher-level function will be selected against. Such mechanisms include, for example: the division of labor between germ and somatic components, which effectively transfers fitness to the higher level; the specific targeting of cancerous cell lines by the immune system in multicellular organisms; and eusocial insect nest-mates killing workers that attempt to reproduce.


25 A key problem for future cultural evolutionary research is to identify what William Wimsatt, supra not 24, calls “escape mechanisms” that allow for deep modifications of entrenched cultural structures whose alteration would otherwise send devastating ripples across a cultural system.
The emergence of new evolutionary individuals — such as complex eukaryotic cells, multicellular organisms, and eusocial colonies — is only possible because of solutions to replication conflicts between hierarchical levels of organization.26 Human societies do not rise to this level of evolutionary individuality, however, in part because they have more limited means of addressing inter-level replication conflicts. As a result, in human societies, adaptations of lower-level components (e.g., elite castes) can emerge despite their deleterious consequences for the larger human collective.

For instance, norms that benefit an elite subset of individuals or a privileged class within a society can persist despite the fact that they are deleterious for most individuals within the society and even for the society as a whole. We will call such cultural adaptations special interest adaptations. In the long haul, special interest adaptations may undermine cultural group stability and thus be selected against; in the short term, however, there is historically little by way of social mechanisms to stop the evolution of special interest adaptations in complex human societies. Some surplus moral norms may thus persist as special interest adaptations because human societies — in particular, large, hierarchical post-Neolithic populations — lack adequate enforcement mechanisms at the group level to guard against them. In the small, kin-based hunter-gatherer bands that dominated the history of human evolution, cooperation conflicts were solved through the evolution of a robust egalitarian ethos.27 These mechanisms were insufficient, however, to prevent exploitation in the comparably recent and far more complex post-Agricultural societies — at least not until the advent of the rule of law and constitutional democracy, which are plausibly viewed as cultural innovations “designed” to check special interest adaptations and thus place limits on state and elite class power.

Traits that were originally selected for performing one fitness-enhancing function may come to perform a new function, including a special interest function that enhances the fitness of a subset of a collective at the expense of other individuals. The complex of norms that constitute the Indian caste system, for example, may have reduced the risk of biological parasites that aboriginal peoples of the subcontinent posed to their Vedic conquerors.28 But later, when the conquered and the conquerors came to comprise one larger society, these same norms may have functioned to consolidate the power of the conquerors’ descendants through their control over the state apparatus (coercion) and religious authority (ideology), by preventing the dilution of power

through intermarriage with descendants of the conquered, and by re-
severing valued social positions for themselves.29

Another example of deleterious norms generated and perpetuated as special interest adaptations concerns the profoundly incompetent, ineffective and unjust criminal justice systems of many developing countries that have gained independence from colonial rule. The norms that underpin criminal justice institutions in many post-colonial developing countries were originally designed to protect the property and power of colonial rulers and their elite allies at the expense of the general population — in essence, they served as mechanisms of popular suppression. After independence, rather than reforming these norms, many post-colonial regimes preserved and benefited from them, with elite groups coming to occupy the powerful positions held by their former colonial rulers.30

E. Destructive moral norms sustained by failures of collective action

Some destructive moral norms can persist due to collective action problems even when they confer no reproductive advantage on anyone. Duelling, footbinding, and female genital mutilation may be examples. Duelling may have originated through a combination of sexual selection (in an earlier environment in which females put a premium on physical courage in males) and selection for relatively constrained forms of violence (where conflicts are settled on the field of honor by the actions of two individuals, as opposed to the continuing, uncontainable conflict of intergenerational blood feuds between groups). But even when these original functions became otiose due to cultural innovations that provided less physically destructive outlets for competition among males — and even though most people, including most participants in the practice, recognized how destructive it was — duelling persisted. Similarly, careful investigations of female mutilation norms, such as footbinding and female genital cutting, show that these practices arose initially as special interest adaptations (in particular, as paternity confidence measures for wealthy elites), and then dissipated into the general population where they were perpetuated by false empirical beliefs (e.g., about the health benefits of the practices).31

A central explanation for the persistence of such maladaptive norms is that abolishing them required solving difficult collective action problems.

29 It is worth noting that while in the past, socioeconomic and political advantages may have been conducive to reproductive fitness, this is no longer true in many societies, where the better-off tend to have lower rates of reproduction.
For instance, even if each potential duelist believes the practice to be irrational and even immoral, any defecting individual will face debilitating social stigma or, in the case of footbinding and genital mutilation, severely reduced marital prospects. Similarly, as Kim Sterelny has suggested, even if the initial victims of female genital mutilation enjoyed an advantage in the mate selection market (given the cultural context of a patriarchal, deeply sexist society in which female chastity was inordinately valued), once the practice became widespread, the advantage evaporated: if virtually every woman has mutilated genitals, having them confers no comparative advantage. In other words, undergoing female genital mutilation came to be a matter of horrific costs with no offsetting fitness benefits for the individuals undergoing it. The practice persists, Sterelny argues, because abolishing it, as with the case of dueling, requires solving a difficult collective action problem. The first defectors from the practice will suffer a reproductive penalty, because they will be viewed as inappropriate mates in societies in which unmarried females have grim economic prospects, and they will be subject to moral condemnation and stigmatization.

Effectively counteracting special interest adaptations, such as those discussed above, may also include solving a collective action problem not only for the society as a whole (through, e.g., anti-infibulation or anti-footbinding pledge societies), but also for the disadvantaged subset of society. A powerful elite or privileged contingent has at its disposal impressive resources for blocking the collective action necessary for the masses to emancipate themselves from surplus norms that favor special interests. These resources include coercion (attaching material costs to noncompliance with such norms) and ideologies that obscure the fact that the norms are nothing more than instruments of class domination. Participation in a revolution involves a cost to the individual participator, but whether the revolution will succeed depends upon whether enough people participate. Even though emancipation would be best for all members of the oppressed class, it may be rational for each oppressed member to refrain from participating in the revolution. Indeed, the tendency of the worse-off to rationalize special interest norms may be a mechanism for avoiding the costs of challenging prevailing special interest norms, given the likelihood that such challenges will not succeed due to problems of collective action.

F. Surplus moral norms can result from cultural transmission dynamics

Finally, some harmful norms may proliferate and be sustained in a society simply due to the intrinsic dynamics of cultural transmission that allow maladaptive variants to spread rapidly in a cultural population. Unlike the clean lines of vertical descent exemplified by genetic transmission,

---

cultural variants can be acquired from and transmitted to any member of a population within a single lifetime; this allows cultural variants to spread much more rapidly than genetic variants, but it also makes cultural transmission uniquely susceptible to the spread of maladaptive variants. As Peter Richerson and Robert Boyd’s modeling work has shown, cultural copying biases — such as tendencies to copy cultural variants that are common, to emulate prestigious individuals, and to identify transparently successful strategies — can allow for cumulative cultural adaptation; but these are far from fail-safe heuristics, as deleterious variants can become common in a population, are often adopted by prestigious individuals, and can often be mistaken for successful strategies, particularly in cases where causation is complex.33

In summary, seemingly arbitrary, irrational or bigoted moral norms may be sustained in a society even though they are deleterious in modern selective environments, even though they are adaptive for only small subsets of human populations (such as powerful elites), even though they actively harm large segments of society, and even though they no longer (or never did) confer a fitness advantage on anyone. The fact that a norm is maintained in a society does not, therefore, provide persuasive or even prima facie evidence that the norm has a salutary function.

G. Evolutionary explanations and the validity of moral judgments

Evolutionary accounts of apparently unjustifiable moral norms can explain why people have strong intuitions that these norms are valid in the absence of any good reason, whether moral or prudential, for sustaining them. In that sense, evolutionary explanations are said to be “debunking” of morality: they imply that, in the cases they explain, intuitive plausibility in itself provides no evidence for the justifiability of a norm. This is not to say that moral norms that people on reflection consider to be valid are unjustified simply because there are adequate evolutionary explanations for their differential proliferation and persistence. Norms against killing, stealing, lying or freeriding may be adaptations at the level of cultural groups; this does not, however, imply that these norms are invalid or unjustified. The reason that evolutionary explanations of such norms are not debunking in the sense noted above is that we can offer persuasive reasons in support of their validity, reasons that do not appeal to reproductive fitness. In contrast, plausible evolutionary explanations of a norm, in the absence of any reasoned justifications, may give us good reasons to doubt its validity.

The fact that evolutionary explanations can be given for supposedly valid moral norms may constitute a good reason to doubt strong versions of epistemological moral realism.34 But the availability of an evolutionary explanation of a given moral norm is clearly not sufficient to show that it is a surplus moral constraint and hence a proper target for de-moralization. This is because moral constraints can be valid on broadly coherentist or constructivist views of moral validity regardless of whether they now or ever were conducive to reproductive fitness. Showing that some moral norm is or was adaptive does nothing whatsoever to establish that it is morally justified — unless we are warranted in assuming that conditions are so dire that maximizing reproductive fitness trumps all other values, which will rarely if ever be the case. By the same token, the fact that a moral norm is nonadaptive or even maladaptive (from a biological standpoint) does not entail or otherwise indicate that the norm is unjustified. What then is the relevance of evolutionary explanation to moral justification?

Evolutionary explanations serve a more modest, but still valuable function: they help us to identify potential targets for proper de-moralization because they give us reason to discount intuitions about the validity of certain norms we have hitherto intuitively regarded as valid. In this sense, they simply reinforce the realization that basic intuitions are too untrustworthy to count as evidence for the validity of moral norms; instead, they are only starting points for the moral reasoning needed to determine whether a norm is justified.35

Some might respond that all moral reasoning is in effect confabulation that serves to justify pre-theoretical evaluative intuitions whose existence can be explained by recourse to evolutionary principles. If one looks at paradigmatic cases of moral confabulation, however, such as those elicited in Jonathan Haidt’s “moral dumbfounding” experiments,36 the signature of confabulation is discernable: it involves a robust insensitivity to information that is logically inconsistent with pre-existing moral beliefs and other symptoms of highly motivated reasoning. There is simply no reason to think that all moral judgments have this characteristic. Such heterogeneity of the moral realm is particularly clear if we compare cases of moral dumbfounding — which typically implicate sexual morality, religion, social identity and other deeply entrenched cultural norms — with moral judgment processes involved in putative instances of moral progress, such as the shift toward the protection of basic human rights, the anti-abolition,

35 Intuitions may also be unreliable if they were formed by processes that are distorted by defective social epistemic practices, including unwarranted epistemic deference to people wrongly thought to be moral experts. Allen Buchanan, “Social Moral Epistemology,” Social Philosophy and Policy 19, no. 2 (2002): 126–52.
decolonization, animal welfare and environmental movements, and improvements in norms that govern international relations. As we shall show, sometimes the moral reasoning required to justify a norm is quite complex and presupposes the resolution of disputed issues in moral and political philosophy.

Understanding how moral norms that involve surplus moral constraints arise and persist, and identifying the functions they are now performing, can provide guidance for how to achieve emancipation from them. For example, if a norm no longer performs any valuable function but entails significant human costs and persists solely due to a collective action problem, techniques for solving collective action problems can and ought to be brought to bear, as has been the case with successful efforts to stop female genital mutilation. If the norm can be shown to function so as to advantage only a privileged elite or tyrannical majority, it may be possible to mobilize social and political pressure to abandon it, or to implement legal-institutional safeguards to prevent the concentration of power.

But first, one must be able to distinguish reliably between necessary and surplus moral constraints: that is, one must reliably determine when de-moralization would be proper de-moralization and when it would amount to the abandonment of a valid moral norm. Answering this question turns out to be more difficult than might first appear, and significantly complicates the task of developing a theory of moral progress that gives a prominent role to increased liberty in the form of proper de-moralization.

III. CAN SURPLUS MORAL CONSTRAINTS BE RELIABLY IDENTIFIED?

A. Improper de-moralization: The conservative challenge

This essay began with a list of relatively uncontroversial cases of surplus moral constraints and hence proper targets for de-moralization. It is worth emphasizing, however, that de-moralization can go awry: people can and often have come to regard as morally permissible behaviors that are in fact morally wrong. For example, in the thrall of Nazi ideology, many ordinary Germans came to believe that behavior they previously viewed as immoral when directed at any person was permissible when the target was a Jew. Similarly, eugenic propaganda in the United States and elsewhere convinced many people that policies of compulsory sterilization that would otherwise have been rejected as immoral were not only permissible but obligatory, given the false assumption that the human gene pool and even civilization itself was imperiled by the rampant proliferation

---

of “defective germplasm.” How can one reliably ascertain whether abandonment of a moral norm and the constraints it entails is a case of proper de-moralization — and hence of moral progress — and when it is not?

One might think that the solution to the problem is simple: any moral norm is likely to be a surplus moral constraint, and hence a proper target for de-moralization, if no sound justification can be given for it. Assuming that we know how to identify sound moral justifications, what more is needed? If we adopt a liberal perspective, then we know that justifications must appeal ultimately to the freedom and welfare of individuals and that brute appeals to religious authority or tradition do not suffice. If compliance with some supposed moral norm exacts significant human costs and there is no justification for it in terms of its contribution to individual well-being and freedom, then isn’t one justified in thinking that it is a surplus moral constraint?

Unfortunately, things are not so simple, as generations of conservatives have emphasized. Whether an accepted moral norm ought to be disregarded as being a case of improper moralization cannot be determined unless we can reliably ascertain its role within a complex web of norms, institutions, and social practices. The justification of moral norms must be holistic or more precisely, ecological, then; but given the limitations of our knowledge about the social wholes within which norms operate, we are at risk of failing to see the true value of certain norms. For example, some moral norms may fit the “monkey drill” analogy discussed above: compliance with them may produce no particular substantive good, but they may nonetheless be valuable because they cultivate and sustain the disposition to follow those moral norms that are important for human flourishing. Other norms may in fact contribute to some substantive good or to the preservation of valuable liberties, but in complex ways that are not likely to be captured by widely understandable, and to that extent simple, moral justifications.

Consider, for example, a set of norms concerning sexual morality that includes a prohibition on unmarried women bearing children and that requires stigmatization of those who violate the norm. Compliance with these norms inflicts serious psychological, social, and economic costs on unmarried women who have children. Yet the norms may in fact be beneficial overall and in the long run, for the class of women as a whole and perhaps even for disadvantaged women, in an environment in which social support for unmarried mothers is lacking and in which marriage is unattainable for many disadvantaged women (or, if attainable, does not constitute an economic improvement). In such an unjust social order, a norm that imposes severe costs on unmarried women who bear children

---

might, depending upon the factual particularities, make moral sense, provided that we give significant weight to the well-being and opportunity of women over the long-run. And if that is so, then concluding that it is a surplus moral constraint — and striving to abolish it — might not be morally progressive, all things considered. Whether the norm is justifiable will depend upon complex moral reasoning that includes among its premises highly disputable empirical predictions about the consequences of adhering to it or abandoning it.

To such reflections on the difficulty of knowing the unintended consequences of de-moralization, conservatives typically add two ideas — and then conclude, too quickly, we shall argue, that one should err in the direction of adherence to the moral status quo. The first idea is that if a moral norm has persisted over a long period of time, then there is some reason to believe it is beneficial. When fleshed out, this idea turns out to rest on a vulgarized, inaccurate view of cultural evolution, namely, that unhealthy social practices will eventually be winnowed out, and that hence that the longevity of a practice is evidence of its salutary nature. We have already seen why this is not so: highly destructive and bigoted social practices can persist for long periods of time because they are adaptations, because they are evolutionary hangovers, because they are entrenched and thus refractory to modification due to their causal connections to other cultural norms, because they serve the special interests of powerful elites at the expense of other segments of the population, because abolishing them requires solving difficult collective action problems, and/or because cultural transmission is highly susceptible to the spread of maladaptive variants.

B. Pessimistic conservative metaphors are inapt

The second idea that conservatives typically invoke to support their claim that judgments about surplus moral constraints are so unreliable as to warrant a strong bias for the moral status quo is the supposed aptness of a metaphor: society is said to be like a seamless web (or, more hysterically, a house of cards). Snipping one apparently insignificant fiber may unravel the whole thing (alternatively, making what one thinks is a minor adjustment in the position of one card may cause the whole edifice to collapse). The operative notion here is simple: there are dense interconnections among moral norms and the social practices that support them, and given how little we know about the particulars of these dense connections, it is hubristic — and morally irresponsible — to abandon a norm simply because we can’t produce a convincing justification for it.

As we have argued at length elsewhere, the seamless web and house of cards metaphors exaggerate the core conservative insight that we ought to take seriously the risk of unintended bad consequences when we abandon
what we have previously regarded as a valid moral norm. Societies, like individual organisms, are not seamless webs and they are certainly not like houses of cards, because neither type of entity is a plausible product of evolution. For an entity to evolve through natural or cultural selection, some features of that entity must be able to change without altering (in a countervailing way) many of its other characteristics. Biological theorist Richard Lewontin has called this the “quasi-independence” criterion for adaptation — and since we know that adaptation exists at both individual and cultural levels, quasi-independence must obtain at both levels as well. Therefore, if societies or individuals were as fragile as the seamless web and house of cards metaphors suggest, they would not be resilient enough to survive and adjust to changes in the environment, including competition from other societies or individuals. The very phenomenon of adaptation is thus not compatible with the hyper-dense internal connectedness that conservative metaphors imply.

If one insists on textile or architecture metaphors, then it would be better to say that individuals and societies are like seamed webs or complexly modular houses. Any entity that is subject to selective shaping is likely to feature a good deal of modularity and redundancy. Modules are functional units that have denser connections among their own constituents than between themselves and other functional units. Modularity is conducive to adaptability because it allows for incremental (intra-modular) changes, with the boundaries between modules akin to seams in a web. Functional redundancy is also conducive to adaptability because it allows for changes that undercut a function in one system or organ to occur without complete loss of that function. For example, many genetic innovations in evolution occur via gene duplication that results in functional genetic redundancy, which in turn frees up one of the duplicates to assume a novel evolutionary function.

The inaptness of the conservative metaphors is further confirmed by the research of Boyd and Richerson, whose modeling work shows that individual norm compliance is much more developmentally autonomous than the conservatives’ favorite metaphors would suggest: almost any norm, including one that requires abandonment of a preexisting widely accepted norm, can enjoy robust compliance if there is effective punishment for noncompliance. Norms are thus not as densely interconnected

as conservatives assume; and if they are not, then the risk of improper
de-moralization is not as great as they assert. Conservatives are fond of
saying that we don’t know enough about society to disregard long-standing
moral norms; but they ignore the fact that we now know that societies are
not like seamless webs or houses or cards.43

If there is clear evidence that continued compliance with a supposed
moral norm entails serious suffering, curtailment of important liberties, or
loss of opportunities for welfare for significant numbers of people — and
especially if those whom the norm particularly disadvantages are already
among the most disadvantaged — then the abstract possibility that aban-
doning the norm may result in some bad unintended consequence is not a
sufficient reason against de-moralization. None of this is to deny that the
risk of unforeseen bad consequences is a serious problem for the reliable
identification of proper targets of de-moralization. The point rather, is that
the standard conservative stance on this problem is one of overreaction
and undue pessimism.

Nevertheless, any theory of moral progress or of liberty that takes seriously
the need for emancipation from surplus moral constraints must develop a
plausible strategy for managing the risks of bad unintended consequences
of de-moralization. “Managing” is the right term here, because it would
be both unfeasible — and undesirable if feasible — to reduce the risks of
bad unintended consequences to zero. Here, as elsewhere, risk reduction
is not costless and the marginal costs of risk reduction are likely to rise at
some point within the feasible set. Instead, the goal is to achieve cost-effective
risk management, where costs are construed quite broadly.

C. Retrospective versus prospective judgments about surplus constraints

It may be, generally speaking, that people are better at making reliable
retrospective judgments as to whether the abandonment of a norm is a case
of proper de-moralization, for two reasons. First, in retrospect (at least if
enough time has elapsed) we may have reason to conclude that abandoning
the norm did not in fact have serious unintended bad consequences. Here it
is worth noting that there are many cases in which conservatives have pre-
dicted dire consequences of de-moralization that have not occurred — for
example, that if same sex marriage is permitted the institution of marriage

43 Of course enforcement only works if it is employed. It might be the case that a norm N1
could be abandoned without bad consequences, including the undermining of a valid norm
N2, but only if another norm N3 were enforced. Suppose, however, that the fact that the
enforcement of N3 is necessary to prevent the abandonment of N1 from causing damage to
N2 is not known and a consequence N3 is not enforced. This possibility lends support to a
moderately conservative thesis with which the author of this paper agrees, namely, that any-
one proposing or welcoming the abandonment of a norm ought to take seriously the risk of
unintended bad consequences of doing so. It does not support the assumption of extremely
dense interconnections among norms suggested by the seamless web metaphor.
will be damaged, or that if interracial marriage is permitted it will lead to the degeneration of the “white race” or to social chaos, or that if consensual homosexual acts are decriminalized, fundamental values will be eroded and the social fabric will unravel. Or consider the extremely pessimistic if not hysterical predictions of the social and psychological consequences of allowing in vitro fertilization thirty years ago.

Second, if sufficient time lapses after the abandonment of what was previously thought to be a valid norm, and if we believe that during the interval the cause of social justice has advanced or at least not been significantly retarded, then we may conclude that the overall effects of abandoning the norm, over the long run, have not been bad. Fortunately, the abandonment of a norm often comes gradually, in stages, as when physician-assisted suicide first becomes permissible only under certain highly constrained circumstances, when medical use is first allowed as an exception to the prohibition on using marijuana, when gay marriage is legalized in certain jurisdictions, or when alternative reproductive practices (such as germline modification) are permitted under limited conditions. In such cases, there will be time to determine whether complete abandonment of a prohibition is advisable or not, and the costs of norm modification will be reversible and contained. Such “moral experimentation,” as it might be called, is a crucial complement to moral reasoning in assessing the justificatory value of existing norms, because the intricacies of cultural casual relations may elude even our best moral reasoning and social modeling. Thus, even if, despite the evolutionary considerations adduced earlier, we take the existence of norm N to be prima facie evidence that N serves some valuable social function that would be vitiated if N were altered, we can conduct controlled experiments in norm modification that allow us to assess the unintended consequences that are likely to flow from N’s alteration.

D. A different evaluative challenge: the problem of distributive justice

As daunting as the task of managing the risks of bad unforeseen consequences of de-moralization is, it may not be the highest hurdle that a theory of moral progress as involving increased liberty faces. A more serious challenge, at least from a moral point of view, stems from the fact that in some cases de-moralization will be beneficial for some, perhaps many, but quite harmful to others. In other words, an account of proper de-moralization must take a stand on some of the most fundamental and disputed issues concerning distributive justice.

The earlier example of a norm against unmarried childbearing and another example also involving constraints on reproduction will serve to illustrate the general point. It is no doubt true that there is nothing morally wrong per se with a woman having a child when she is not married. Similarly, there may be nothing morally wrong per se with having more
than three children. Yet in some circumstances, there can be significant moral benefits if the norms prohibiting these behaviors are sufficiently internalized so as to achieve a high degree of compliance; and abandoning them on the assumption that they are instances of surplus moral constraints may have morally undesirable consequences. Suppose that many of those who would have children when unmarried if the norm were not effective already suffer serious socioeconomic disadvantages. Suppose also, that due to a shameful lack of social support, they will be even more seriously disadvantaged if they become single mothers. Under these circumstances, the norm against unmarried women having children might in fact have significant beneficial consequences for society overall. More importantly, it might improve the condition of children and the prospects for socioeconomic improvement for some of the worst-off women in society. But, noted earlier, the costs of compliance with the norm may be especially high for women who have a strong preference for having children but have dim prospects for marriage or for whom marriage would not be an economic improvement. Abandoning the norm will benefit some women and children but harm others. So, in order to know whether the norm ought to be abandoned, we need first to know which interests count the most. In brief, we need to be able to make a principled judgment about distributive justice. Until that judgment is made, we cannot know whether the norm is a case of surplus moral constraint or one that, all things considered, we ought to adhere to.

Similarly, a norm against having more than three children, if widely complied with over several generations during which time important forms of economic development occurred, might have considerable benefits for many people, including the worst-off. But the costs of limiting the family to three children might be severe for some of the worst-off members of society in the current generation of potential parents, if economic development and social support services are presently so deficient that the worst-off must rely for support in old age on their children, under conditions in which child mortality rates are high. Adhering to a norm prohibiting more than three children might produce great benefits for future generations of the worst-off and for society generally, but might exacerbate the burdens of the current generation of the worst-off. Here, too, there is a fundamental problem of distributive justice — in this case intergenerational distributive justice — and one that must be solved before we can know whether abandoning the norm would be a case of emancipation from surplus moral constraint.

Of course it might be replied that there are other ways of improving the condition of future generations of the worst-off or of single parents, for example by implementing various social welfare policies that improve their health, educate them, provide family leave and other forms of parenting support, and equip them to engage successfully in a modern economy. None of these, so the objection goes, would involve the burdens that
prohibitions on having more than three children or stigmas on single parenting impose on the current worst-off or for capable single parents. True enough, but if in fact no such measures are politically feasible, it might be the case that adhering to these norms is the best way to improve the long-term prospects of the worst-off.

So, there are two difficulties with making a judgment that some norm is a surplus moral constraint. First, there is the problem of predicting the consequences of abandoning it. One may lack the relevant knowledge of causal connections — including effects on other norms. Second, there is the problem of evaluating the predicted consequences. Their evaluation may turn on highly disputed issues of distributive justice — perhaps issues that no current theory would be capable of resolving satisfactorily.

This is not to say, however, that contemporaneous or prospective judgments about surplus moral constraints are never justified. In some cases, the human costs of continued compliance with a supposed moral norm are so horrific, the benefits so arbitrarily skewed toward one group in society, and the lack of a justification so patent, that we may rightly conclude that adherence to the norm is a case of unnecessary, self-inflicted curtailment of liberty. The instances of de-moralization listed at the beginning of this essay seem to satisfy these criteria.

IV. Conclusion

Human beings have frequently been wrong about what morality requires. They have often imposed unnecessary and harmful limitations on their own liberty and that of others by accepting and enforcing moral requirements of dubious validity. Early theorists of liberalism in the Enlightenment recognized that internal as well as external constraints can wrongly limit liberty, and held that genuine human emancipation, and with it moral progress, require emancipation from both types of constraints.

This essay is an attempt to correct for contemporary liberalism’s neglect of both emancipation from unwarranted moral constraints and of the idea of moral progress. First, we argued that proper de-moralization — emancipation from surplus moral constraints — ought to be acknowledged and investigated as an important form of moral progress, especially from the standpoint of any conception of morality that values liberty highly. Second, we showed that evolutionary explanations can account for the origins and persistence of surplus moral constraints, and that such explanations provide valuable, though limited, resources for determining when moral norms constitute unwarranted limitations on liberty, and how emancipation from surplus moral constraints can be achieved.

Third, we showed that a major point of contention between earlier liberals who viewed emancipation from surplus moral constraint as an important form of moral progress and conservatives who are skeptical of progressive cultural change turns on the reliability of judgments that purport to
identify surplus moral constraints; and we argued that conservative skepticism about our knowledge of society, and the deployment of pessimistic (but fallacious) conservative metaphors, exaggerate the unreliability of such judgments. Finally, we argued that in some cases, ascertaining whether a moral norm entails surplus constraint and hence is a proper target for de-moralization cannot be determined without resolving difficult and disputed issues of distributive justice.

Our chief conclusion is that any theory that aims to explicate the importance of increased liberty or welfare for moral progress ought to take the phenomenon of de-moralization seriously, and must develop an account of the conditions for making reliable judgments about surplus moral constraints. Such a theory should make use of knowledge gained from moral experimentation, as well as current research in evolutionary moral psychology and cultural evolutionary theory, to better understand not only the reliability of moral judgments, but also to explain the origins and persistence of invalid moral norms, and to provide practical guidance as to how emancipation from these norms might plausibly be achieved.

*Philosophy and Law, Duke University*

*Philosophy and Biology, Boston University*