



Butchering Benevolence Moral Progress beyond the Expanding Circle

Hanno Sauer¹ 

Accepted: 18 February 2019 / Published online: 25 February 2019
© The Author(s) 2019

Abstract

Standard evolutionary explanations seem unable to account for inclusivist shifts that expand the circle of moral concern beyond strategically relevant cooperators. Recently, Allen Buchanan and Russell Powell have argued that this shows that evolutionary conservatism – the view that our inherited psychology imposes significant feasibility constraints on how much inclusivist moral progress can be achieved – is unjustified. Secondly, they hold that inclusivist gains can be sustained, and exclusivist tendencies curbed, under certain favorable socio-economic conditions. I argue that Buchanan and Powell concede too much to the evolutionary conservative, because their second point shows that conservatives are right about the first: inclusivist shifts are unrealistic where it matters most, namely under harsh social, political and economic conditions. I suggest two promising strategies for solving this problem. One is to focus on different *forms* of moral progress to secure the same moral gains. The other is to look beyond possible extensions of our psychological capacities altogether, by providing institutional support that renders them irrelevant. We should bypass, rather than further stretch, the constraints of our evolved psychology to make moral progress possible.

Keywords Moral progress · Conservatism · Expanding circle · Evolutionary psychology · Moral disgust

1 Introduction

When it comes to furnishing an explanation of the building blocks of our moral psychology, evolutionary theory holds a lot of promise. At the same time, there is an inherent tension between an evolutionary perspective, which looks for the distal causes of deeply entrenched features of cognition and behavior, and the possibility of moral progress, which concerns how human

✉ Hanno Sauer
h.c.sauer@uu.nl

¹ Universiteit Utrecht, Utrecht, Netherlands

thinking and action can be changed for the better. Evolution is about how the way we used to be informs how we are *today*; progress is about how we can be *tomorrow*. The Pleistocene environment we were selected for and the environment we inhabit today are utterly different from each other. The environment we will inhabit, if advocates of moral progress get their way, will be even more different. For evolved beings like us, genuine moral progress may be out of reach.

To counter this suspicion, Buchanan and Powell (2015, 2016) have recently proposed the outlines of a naturalistic theory of moral progress. They claim, firstly, that standard evolutionary explanations cannot account for inclusivist shifts that expand the circle of moral concern beyond strategically relevant cooperators. This is supposed to show that *evolutionary conservatism* – the view that our inherited psychology imposes significant feasibility constraints on how much inclusivist moral progress can be achieved – is unjustified. Secondly, they hold that inclusivist gains can be sustained, and exclusivist tendencies curbed, under certain favorable socio-economic conditions.

I argue that Buchanan and Powell concede too much to the evolutionary conservative, because their second point shows that evolutionary conservatives are right about the first: inclusivist shifts are unrealistic where moral progress matters most, namely under harsh social, political and economic conditions. I suggest two ways around this problem concerning the feasibility of morally progressive gains. One is to focus on different *types* of moral progress that are underwritten by psychological capacities which are not subject to comparable evolutionary constraints. The other is to look beyond possible extensions of our *psychological* capacities, but for *institutional* support that renders them irrelevant. We need to bypass, rather than further stretch, the constraints of our evolved psychology.¹ When Buchanan and Powell (2018) talk about *inclusivist* and *exclusivist* moralities, they clearly have individual human beings' inclusivist or exclusivist psychological "responses" (219) or "dispositions" (199) in mind. This focus in individual psychology, I aim to show, is misleading, and should be replaced with a focus on how to secure social cooperation institutionally. These two strategies complement each other, and together offer a more promising response to the challenge to moral progress posed by evolutionary conservatism than the one suggested by Buchanan and Powell.

My paper has five sections. In (1), I briefly outline the basic thrust of the challenge from evolutionary conservatism. In (2), I offer a diagnosis of the central problem with Buchanan and Powell's progressivist response. (3) develops my first objection, according to which a more promising reply to the evolutionary conservative draws on moral attitudes that can secure the desired inclusivist gains without being subject to comparable psychological constraints. My main example here will be moral disgust, which I will describe as a suitably malleable candidate for satisfying these desiderata. In (4), I sketch an institutionalist solution to the overall problem that is based on the idea that in many cases, smart institutional kludges allow us to economize on moral motivation in a way that bypasses issues of psychological feasibility altogether. (5) situates the argument developed here in a wider context of cumulative moral learning and its evolutionary role.

¹ The position of the evolutionary conservative is perhaps best instantiated in Asma (2013) and Haidt (2012). On a descriptive level, Greene's (2013) argument offers a similar diagnosis, but with very different normative conclusions. See Persson and Savulescu (2017) for a more pessimistic take on the possibility of moral progress without the aid of human enhancement.

2 Evolutionary Conservatism

Evolutionary conservatism is the view that certain progressive social developments are at odds with our evolutionarily inherited psychology. This fact, in turn, is supposed to have non-trivial morally and politically conservative implications.²

According to the dominant account of moral progress, ethical improvements consist in an “expanding circle” (Singer 2011) of moral concern. Buchanan and Powell describe this as a series of “inclusivist shifts” (Buchanan and Powell 2015, 55). Not long ago, (full) membership in the moral community was restricted to an astonishingly narrow group of people. Over time, the moral franchise was extended from wealthy men of the respectively dominant ethnic group to a wider selection of people. Sex and gender, race and ethnicity, wealth and education, religion and nationality, health and even species membership came to be recognized as essentially morally arbitrary distinctions that carry no independent moral weight. In order to count as genuinely progressive, the expanding circle must, of course, consist of the right kind of inclusion. Moreover, the respective expansions can occur at the level of the individual or the majority of a society, and can result in informal moral as well as formal legal recognition.³ This process has frequently been imperfect, or hasn’t happened at all (yet?) in many places. But across a wide variety of social contexts, the expanding circle of inclusivism has been an important aspect of the trajectory of moral development.

Evolutionary conservatives think that these shifts, however desirable they may seem in theory, are simply not feasible. More precisely, they argue that at the very least, this dynamic of inclusion cannot be continued indefinitely, and will ultimately run up against decisive psychological obstacles that we owe to our ancestors’ evolutionary trajectory. In its stronger forms, evolutionary conservatism is the view that we have recently reached the limits of progressivism or, stronger still, have already begun to *overstretch* the carrying capacity of our moral psychology.

The argument goes, roughly, as follows: in order to be sustainable, progressive social developments crucially depend on a certain psychological environment. For racism and speciecism to become a thing of the past, people need to be able to have moral concern, and extend moral consideration, to beings beyond their nearest and dearest, kin and kith. Otherwise, the aforementioned progressive developments will remain unstable.

The problem, as conservatives are quick to point out, is that our ability to care about others at all has an evolutionary rationale. In small, closely genetically related tribal groups, certain cooperative dispositions such as kin altruism or reciprocal altruism can be adaptive. Internally cooperative groups will enjoy a selective edge in intergroup competition for scarce resources in the environment of evolutionary adaptedness. Individuals who are disposed to care for their offspring and engage in reciprocal chains of sharing and helping increase the comparative frequency of their genes in the next generation. Unfortunately, this also means that our capacity to care for others is, and must be, limited. Natural selection operates on genes, and will tend to favor cooperative strategies that promote the proliferation of copies of alleles. This result cannot be achieved through cooperation that goes beyond relatives or reciprocators, because

² Buchanan and Powell cite Asma (2013) and Haidt (2012), among others, as representatives of the evolutionary conservative position. For a critical discussion of Haidt’s “moral foundations” theory, which forms the basis of his conservative argument, see Sauer (2015). For a general response to Haidt’s account of moral reasoning, see Sauer (2017).

³ I am grateful to an anonymous reviewer for forcing me to clarify these points.

such forms of cooperation would result in costs that are not outweighed by any adaptive benefits, and are thus selected against.⁴

This means that evolutionary pressures have equipped our minds with a recalcitrant tendency to carve up the world in terms of an in- and an out-group. Benevolence exists, but *universal* benevolence is evolutionarily unstable. The very dispositions that make us cooperative also make us tribalistic. Buchanan and Powell do not deny this.⁵ Since progressive moral developments require psychological support, evolutionary conservatives claim that this imposes substantial limits to how much inclusivist progress can be achieved. Progressives may fantasize about a world with open borders and full-blown animal rights. But a world without *us* and *them*, conservatives suggest, is problematically utopian. It is an ideal that should be given up.

That there are limits to our empathic concern can be empirically corroborated. Recent studies suggest that we suffer from steeply declining marginal empathy (Västfjäll et al. 2014, 2015). We can muster concern for one (identifiable) individual. Two – not so much. (This is called “compassion fade”.⁶) And even though there are cases in which people do make sacrifices to help strangers – either in an emergency or by engaging in charitable giving – on average empathic concern tends to drop quite rapidly.

This is exacerbated by the fact that the impossibility of helping everyone makes us less likely to believe that we can help *anyone*. (This phenomenon is called “pseudoinefficacy”.) It is worth mentioning that even if this is true, which markers the ingroup/outgroup distinction latches onto is actually somewhat arbitrary (Heath 2014a). There is no reason why our exclusivist tendencies cannot be channeled towards morally arbitrary but *harmless* features. (The boundaries of group membership could be drawn, for instance, not on the basis of race or ethnicity, but in terms of which inherently arbitrary group one identifies with, such that no morally and politically relevant forms of discrimination and disadvantage remain attached to these group divisions. Some people may support Manchester United, some may root for Liverpool. But it isn’t particularly likely that these forms of group membership are tied to significant differences in wealth or social status.) If we also grant, for the sake of the argument, that it is not ideal when social institutions encounter individual minds whose psychology is rigged against them, this gives us a powerful argument for thinking that certain proposed progressive goals ought to be abandoned, or at least reconsidered.

Let me emphasize that the selectionist explanation of altruistic attitudes sketched above does not mandate conservative conclusions. Joshua Greene (2013), for instance, argues that we evolved a certain type of cognitive machinery to deal with an evolutionary “tragedy of the commons” and the fact that free-riding and selfishness remain individually rational for everyone, thereby undermining cooperative relations for mutual gain. But the solution that was selected for then ends up trading one problem – the tragedy of the commons – for another, the “tragedy of commonsense morality” (Greene 2013, 1ff.). This second problem is a direct result of the first: for morality to evolve at all, it had to remain restricted to group members, thereby replacing the “me vs. us” structure of the former tragedy with the “us vs. them” structure of the latter. According to Greene, however, we can still search for a common normative currency that can constitute a “metamorality” which will allow us to mitigate intergroup conflict and reconcile competing intragroup moralities. Moral progress is feasible if we emphasize our common moral ground.

⁴ Whether this is obvious is actually an open question, see Bowles and Gintis (2011).

⁵ See de Lazari-Radek and Singer (2012) and Greene (2013).

⁶ For the role this undersensitivity can play in debunking the reliability of moral judgment, see Sauer (2018a).

3 What Conservatives Get Right

What is Buchanan and Powell's solution to the conservative feasibility challenge? In their first of the aforementioned papers, Buchanan and Powell argue that the reality of the "inclusivist anomaly" shows that our psychology is not, or at least not thoroughly, rigged against progressive shifts. Modern forms of "subject-centered" (Buchanan and Powell 2015, 48) morality are purportedly not available to orthodox selectionist evolutionary explanations. The main reason for this is that subject-centered moral outlooks grant moral status on the basis of the mere fact that someone is a person, or a sentient being, rather than being a strategically relevant (potential) cooperator. The latter feature excludes infants, the severely disabled, or future people, and probably excludes most (if not all) non-human animals. Moreover, it can easily be lost, while the former cannot (Buchanan and Powell 2018, 57 f.). I will argue that the claim that such moral outlooks are not amenable to evolutionary explanation is misleading, or at any rate unnecessary. We do not need our onboard psychological resources to support progressive change. We need cleverly designed institutions that harness the power of our evolved dispositions for counterintuitive cooperative arrangements.

Buchanan and Powell do not just argue that there is a certain fact – the existence of inclusivist moral developments that lead people to recognize the moral status of strategically irrelevant entities such as animals or the severely disabled – that contradicts evolutionary conservatism. They also develop a positive account of the conditions under which inclusivist moral progress is possible. Here, their core claim is that exclusivist tendencies are *adaptively plastic traits*. Under favorable conditions, these traits need not be expressed. It is only when conditions are harsh, or when enough people believe that they are harsh, that exclusivism will manifest, creating a hostile environment for a more generously drawn moral circle.⁷

However, remember that evolutionary conservatives argue that inclusivist shifts are psychologically infeasible or at least so psychologically unrealistic that they threaten the sustainability of progressive change. And note that achieving progressive shifts in the direction of extended cooperation and less uncooperative behavior (such as violent conflict between groups) is by far most urgent in societies that suffer from unstable cooperation and lots of violent conflict due to the morally arbitrary constrictions of the moral circle prevalent in them. These two claims together entail that even on Buchanan and Powell's optimistic story, evolutionary conservatives are proven right where it matters most: places ridden with conflict and poverty are also most likely to struggle with exclusivist tendencies towards minorities, from discrimination to genocide. These places are thus where the need for inclusivist shifts is greatest. But due to socially, economically and politically harsh conditions, inclusivist moral progress *is* psychologically infeasible in those places, which is to say where it matters most. For most intents and purposes, evolutionary conservatism is correct.

Inclusivist and exclusivist tendencies, of course, admit of degrees. Individuals and societies can be more or less strongly inclusivist or exclusivist. The main advantage of the alternative I will sketch is that where any given individual or society falls on this spectrum is largely irrelevant for the prospects of moral progress *if* the right institutions are in place.

Consider, for the purpose of illustration, that dense urban environments can foster liberal attitudes (Campbell 2017). Evolutionary conservatism thus seems strictly speaking false. Stable progressive societies are possible. The problem is that in order to get to those dense

⁷ Actually it seems that the second condition is sufficient, because when circumstances are objectively harsh but people somehow don't appreciate this fact, exclusivism may not manifest either.

urban environments in the first place, many of the most pressing problems of intertribal conflict must already be sorted out. Large, prosperous cities are *facilitated* by liberal attitudes as much as they promote them. If Buchanan and Powell's admission that inclusivist attitudes are an evolutionary "luxury good" (Buchanan and Powell 2016, 996) is correct, evolutionary conservatives are proven right about the fact that expansions of the moral circle are psychologically unrealistic in the very contexts where they are most needed.

The goal of moral progress is to eliminate the wrongs caused by discrimination, conflict, and poverty. But if inclusivist attitudes can only thrive under conditions of economic prosperity and social stability, expansions of the moral circle will be out of reach in most places, and remain psychologically feasible only where the cooperative conditions they are supposed to result in already obtain. Now, it needs to be emphasized that the fact that progressive attitudes both require and facilitate favorable social conditions does not entail that there is no way out of this apparent vicious circle, as long as there are ways to get the progressive ball rolling without relying on those very attitudes. Below, I will argue that various forms of institutional support can indeed perform this function of unleashing moral progress.⁸

4 The Wrong Kind of Progress

In what follows, I wish to suggest ways around this problem. For one thing, empathy and altruism may be particularly unsuitable for achieving moral progress (Bloom 2017). For another, it may be that we should not expect our individual moral psychology to play an important role in promoting or maintaining moral progress to begin with.

The good news is that the problem with Buchanan and Powell's naturalistic account of moral progress is homegrown. They look at the wrong kind of moral progress, and the wrong kind of mechanism to support it. This is at least somewhat surprising, since they clearly appreciate the general importance of such other types of progress and the importance of institutional support. What they fail to see is their potential when it comes to countering the challenge from evolutionary conservatism. In the following two sections, I will elaborate on these points.

The main task is to find a form of moral progress that satisfies three conditions:

- (i) it is able to deliver the same kinds of moral gains as the expanding circle of moral concern is (supposedly) able to while
- (ii) not being subject to comparable evolved psychological constraints and
- (iii) not being an evolutionary "luxury good".

Here, it is worth mentioning again that the expanding circle of moral concern is only one type of moral progress among many. Other types include processes of proper demoralization, proper moralization, or indeed contractions of the moral circle (Moody-Adams 2017; Evans 2017; Summers 2017; Arruda 2017, and Rønnow-Rasmussen 2017). Life insurance used to be considered morally repugnant. So have early forms of vaccination, or flushing a toilet on screen (Rebello 1990). Nowadays, these things have become demoralized, and rightly so. Conversely, workplace sexism used to be seen as a morally neutral fact of life. It constitutes progress that it is now considered morally problematic.

⁸ For an account of how institutional scaffolding can support competent moral judgment, see Sauer (2018b).

Buchanan and Powell (2017) are of course aware of this. However, in developing their naturalistic theory of moral progress, they nevertheless decide to focus on inclusivist moral gains and the psychological capacities such as altruism or empathy that underwrite them. This is a mistake, because it deprives them of a plausible and normatively satisfying response to the challenge from evolutionary conservatism. Empathy and altruism *are* limited. Fortunately, there are other emotions beyond empathy, and other forms of moral progress beyond expansions in the moral circle, that can deliver the same moral goodies of moving us towards more inclusive and less discriminatory practices. I will use disgust and demoralization as my main examples, but the point generalizes to other attitudes as well.

Dynamics of demoralization can replace expansions of the moral circle. It is misleading to suggest that progressive developments towards inclusivism share an elective affinity with empathy and altruism. Exclusivism is the historical default, but it doesn't result from a lack of empathy as much as revulsion and aggression towards members of other groups. Recent accounts indicate that one of the main reasons sexist oppression prevails in, say, India isn't that powerful men don't care about or don't particularly like women, or that they don't consider them to be members of the moral community.⁹ It is that they act with complicity in burdensome and coercive norms that dictate the mistreatment of women on the basis of dehumanizing attitudes of disgust, in combination with having the power to put those attitudes into practice. This makes redirecting or mitigating (moral) disgust towards marginalized people at least as promising and potentially powerful as expanding empathy and altruistic dispositions towards a larger group of people.

A frequently overlooked advantage of disgust is that the mandates of empathy and other inclusive attitudes frequently remain within the realm of what's praiseworthy rather than obligatory. "Imperfect duties" of aid or assistance require judgment; the demands of disgust, on the other hand, yield claims of impermissibility and obligatoriness that admit of much less leeway. What is disgusting psychologically presents itself as something that *must* not be done, even if such an attitude is not necessarily endorsed by those who harbor it; what is helpful *may* (or may not) be (Kahan 1999). Moreover, disgust is motivationally powerful. Empathy usually is not (Prinz 2011; Bloom 2017) or, more precisely, empathy is often surprisingly motivationally weak even when something of great moral significance is at stake (think of global poverty). Its motivational strength remains biased in favor of one's ingroup of friends and family. Indeed, altruistic tendencies and disgust have complementary virtues and vices: empathy may be inherently benign, but is weak and rigid; disgust may be inherently problematic, but it is forceful and malleable. Empathy mainly picks up on distress cues, which are processed in a biased fashion. What people are disgusted by, on the other hand, is highly culturally variable (beyond a core of markers for pathogen and toxicity vectors).¹⁰

Now, the problem with disgust is that it seems like a paradigmatically *non*-progressive emotion, and thus an awkward bedfellow for those who want to accelerate moral progress. There is extensive evidence linking disgust to concerns about bodily purity and generally conservative political views (Graham et al. 2009). It is also often thought to be particularly unreliable. Due to its origins as a safeguard against contamination, it follows a hypersensitive "better safe than sorry" logic that yields an excessive number of false positives (Kelly 2011).

⁹ See, for instance, <https://www.theguardian.com/global-development/2015/dec/22/india-menstruation-periods-gaokor-women-isolated>

¹⁰ Disgust is frequently cited as a main example for so-called "recalcitrant" attitudes that can persist despite being irrational. For a recent discussion of the merits of moral disgust, see Plakias (2018).

When disgust is coopted to police social norms, it can have all kinds of pernicious effects (Nussbaum 2009).

In order to enlist disgust for progressive causes, one needs to get people to be disgusted by the right things and, perhaps even more importantly, prevent them from being disgusted by the wrong things. This is where the notions of demoralization and (proper) moralization come into play. Demoralization happens when individuals and/or groups get rid of unjustified pseudomoral prohibitions. Vivid examples for such “surplus moral constraints” (Buchanan and Powell 2017) are the stigmatization of menstruating women in India or discriminatory attitudes towards the disabled. Disgust, or lack thereof, plays a crucial role in eliminating such harmful prohibitions. Proper moralization, on the other hand, happens when people start to disapprove of practices whose moral odiousness had previously gone unnoticed.¹¹

The prospects for a convincing response to the challenge from evolutionary conservatism would greatly improve if disgust could be “appropriated” (Kahan 1999) for such progressive causes. Historically, disgust has frequently played a role in sustaining unjustified moral norms, such as rules against “miscegenation”, religious satire, or various medical innovations (Kass 1997). But we don’t have to be disgusted by interracial marriage, provocative music videos or stem cell research. People can get rid of their intuitive revulsion towards such actions. After a while, people cease to develop it in the first place. Instead, they can become disgusted by such instances of harmful moralization themselves. Disgust towards sacrilegious art can turn into disgust towards those who want to regulate art for the sake of religious ideology (Kahan 1999, 65ff.). Disgust towards homosexuality can turn into disgust towards “Don’t ask, don’t tell” advocates. And when none of this works, steps can be taken to engineer one’s external environment to prevent the undesirable but recalcitrant response from being triggered. (i) thus seems to be satisfied: there are some normative attitudes that can secure the moral gains typically associated with extended empathy, disgust being one example.

Indeed, some authors have recently tried to save disgust from its dubious reputation. Alexandra Plakias (2013), for instance, argues that due to its properties as a detector of vectors of disease, disgust is suitable for tracking social contagion as well. Immoral behavior can spread through populations in a way that is more or less analogous to disease and contamination. Disgust detects social contamination. In a similar vein, Kumar (2017) writes: “that disgust is implicated in important moral norms and values that are shared by liberals and conservatives. Disgust is repurposed in ways that support these norms and values, by motivating an important form of punishment, tracking the spread of moral violations, and expressively coordinating collective action. Disgust accurately detects the nature of certain wrongs that commonly elicit moral revulsion” (13).

What about condition (ii)? Moral disgust (and other emotions) can play an important role in pushing towards further expansions of the moral circle. But the main advantage of disgust, from an evolutionary perspective, is that it does not suffer from comparable evolutionary constraints. What we are and are not disgusted by is highly malleable (Tybur et al. 2013). Some authors have suggested that, if anything, disgust is *too* malleable to enjoy any normative authority (Knapp 2003).

In many cases, redirecting disgust towards progressive causes is not even necessary. All that is required is for people to shed their revulsion. And clearly, there are virtually no evolutionary constraints on how *few* things people can become disgusted by. Narrowing down what’s

¹¹ Note that frequently, this moral odiousness has of course not gone unnoticed by the *victims* of those practices. See Fricker (2007).

considered gross to feces and food would be a major moral accomplishment. Disgust can also incorporate and update on morally salient information. Famously, people can become disgusted by something, such as the sight and smell of meat, they have found to be morally objectionable on disgust-independent grounds (Rozin et al. 1997). The conservative argument from hardwired evolutionary constraints thus cuts no ice against disgust.

Perhaps most importantly, repurposed disgust is not an evolutionary luxury good. Buchanan and Powell argue that inclusivist attitudes only become psychologically feasible under favorable socioeconomic conditions of economic prosperity and political stability. This is a problem for advocates of moral progress, because it means that societies can only enter the track towards moral gains if they are already on it. Disgust does not suffer from a similar problem. People can easily become disgusted by all kinds of different things, regardless of whether conditions are good or not. Beside a minimal core, what cues disgust is triggered by seems highly cultural variable. Empathy and altruism, on the other hand, are either parochial or inclusivist, depending on the circumstances. (iii) is thus satisfied, too.

One may doubt that moral disgust should count as genuine disgust (Gert 2015). Since we are talking about what is and isn't disgusting, introspective evidence seems admissible. Consider the litany of human atrocity that is Livingstone Smith's *Less Than Human* (2011). The list of disembowelments, cut off breasts, severed limbs, or baby's heads flung against walls is endless. The descriptions of genocide by the Spanish against the Native American population, the Germans against the Jews, the Hutu against the Tutsi, or the Japanese against the Chinese are frequently viscerally disgusting.

It is likely that a fair amount of talk about how "disgusting" a reprehensible action is purely metaphorical. Moreover, immoral actions frequently just happen to be disgusting. On the other hand, there is neuroscientific evidence suggesting that even so-called "pure" moral violations – that is, actions which aren't independently disgusting such as mutilations or unusual sex acts – can elicit reactions of disgust (Kumar 2017, 4 f.). This link is buttressed not just by self-reports, but also facial activity, implicit measures, and behavioral evidence (Chapman and Anderson 2013).

Before I proceed, let me address two possible objections. First, I would like to clear up a potential confusion: I have argued that the expanding circle of moral concern is not the most suitable form of social change for the better for progressives to focus on. But it seems that what I am actually suggesting is not to give up such inclusivist shifts as a valuable end but to look for different *means* of accomplishing the same end.

This impression is, to a large extent, justified. As I have tried to emphasize, one of the virtues of disgust, along with other moral attitudes such as anger or indignation, is that they can be used to secure *the same moral gains*. At the fundamental level, these moral gains consist in reducing discrimination, xenophobia and the gratuitous or instrumental infliction of suffering on members of the outgroup. Such gains can be achieved either by (de)moralizing certain behaviors or by adjusting criteria of membership in the moral community. My point was that the latter strategy is less promising because of the constraints our evolved psychology imposes in such adjustments. By focusing on inclusivism as the main form of moral progress, we are more likely to end up focusing on inadequate psychological means of achieving it. But even if my alternative strategy succeeds, the expanding circle of inclusivity would remain, at the very least, an extremely welcome byproduct.

Second, one may doubt the empirical side of my story. According to this worry, there is simply no reason to believe that disgust fares any better than empathy or altruism do when it comes to how strongly its development and subsequent content is canalized by evolution. Disgust may be equally vulnerable to a challenge from evolutionary conservatism.

It may of course simply be true that the remedy for evolutionary conservatism I have identified falls short because the empirical details of my account do not pan out. This would not entail, however, that, *given what we know about the limitations of inclusivist attitudes*, the general recommendation of looking for other, less evolutionarily constrained ways of securing progressive moral gains would not be worth making. Consider, for instance, the emotion of shame. Like disgust, shame is highly malleable in terms of the sheer variety of (sometimes bizarre) things people can become intensely ashamed of doing or failing to do. Appiah (2010) famously describes the intense pressure, which sometimes persisted for centuries, to participate in gentlemanly dueling or footbinding in Europe or China, respectively. Not doing so would result in enormous shame, social stigma, and ostracism. And yet these centuries-old norms evaporated within a generation or less, never to be encountered again.

The historical record suggests that disgust is malleable in just this way. Homophobic attitudes used to be (and in many cases still are) underwritten by strong feelings of disgust. The same holds for so-called “usury” or “miscegenation”. But the amplification of inclusivist attitudes hardly played a role in ending discrimination against, gay people, Jews, or people in interracial relationships. Disgust, on the other hand, is less constrained by evolution because of the seemingly unlimited range of things people can become disgusted by and the staggering speed with which even whole societies can cease to be disgusted by various actions. This remains true even if one concedes that disgust is subject to some evolved limitations, and that it sometimes does play a role in maintaining ingroup/outgroup divisions.

I have focused on disgust here, but the same point generalizes to other feelings. Take attitudes of respect/honor. Contractions of the moral circle can also be progressive, such as when excessive belief in and respect for the (alleged) authority of powerful people is curtailed (Huemer 2013). Here, too, there seem to be no evolutionary limits as to how many and which things we can *un*learn to respect. Disgust supports dynamics of demoralization, developments of respect and honor support (proper) contractions of the moral circle. Both of these forms of progress support the kinds of developments we are after when we think about expansions of the moral circle: reductions in violence, genocide, discrimination, or dehumanization. But neither is subject to comparable evolutionarily inherited psychological constraints.

5 Institutional Bypassing

The problem of evolved psychological constraints disappears completely once we start looking beyond psychological solutions to the conservative challenge altogether.

Consider the following analogy: imagine an epistemological view that functions just like evolutionary conservatism does in the field of ethics. Such evolutionary conservatives in epistemology would be in some sense right to say that there are limits to knowledge and information processing, set by our evolved neural capacities. Our brains were not made to appreciate the vastness of the cosmos, or the mathematical properties of prime numbers. Nothing in the environment of evolutionary adaptedness prepared us for gauging the speed of light or figuring out quantum weirdness. Therefore, this epistemological evolutionary conservative would suggest, many forms of knowledge are simply out of reach for simple primates like us.

Yet here we are, flying to the moon, proving Fermat’s Last Theorem, snapchatting and vaccinating our children. How was this possible? The answer, of course, is that we didn’t need enlarged individual epistemic capacities and bigger brains (by analogy: more empathy) to achieve this. Evolutionary conservatism about knowledge on account of our unimpressive

brains is false. To acquire more knowledge, we rely on cumulative cultural transmission, an epistemic division of labor and trustworthy institutions of knowledge-generation and storage such as science. Our onboard epistemic capacities underwent almost no change over the past 10,000 years. Indeed, our superior individual intelligence does not seem to be the “secret of our success” at all (Henrich 2015, 8–22). If anything, this shows that our onboard epistemic capacities are not all that important. They are not what makes us smart.

The same applies to our onboard *moral* capacities such as empathy. To bring about progressive moral change, we don’t need to stretch our moral concern further. We need to work around it with clever institutional kludges.

What are empathy and altruism for? For the most part, their moral value lies in the tendency to induce cooperative dispositions in agents. These cooperative dispositions plausibly face severe limitations, in no small part set by evolution. But with the right institutions, these constraints simply do not matter. We can design institutions that economize on our limited inclusivist attitudes. Adversarial institutions that indirectly induce cooperative behavior via staged competition are a prime example.

Consider the market as a paradigmatic case for such an institution. People are mildly cooperative. Evolutionary conservatives argue that this undermines the prospects of unlimited expansions of the cooperative circle. However, our psychological capacities – which we largely share with other primates – simply cannot account for current forms of human cooperativeness. Modern humans are hypersocial in a way that cannot be attributed to changes in our psychological capacities because these developments are too recent. Cooperation in large-scale societies, if it had to reach through individual minds to work, would impose unbearably heavy motivational burdens on individual people. It is almost impossible to motivate people to cooperate with strangers on the other side of the planet out of the goodness of their hearts. There are evolutionary limits to human inclusiveness.

Does this mean that evolutionary conservatives are right, and large-scale cooperation is not a feasible political ideal? Of course not. There is an institutional arrangement – the market – that provides a workaround. It facilitates extensive chains of cooperation, and incentives to benefit others, without relying on the baker’s (or anyone’s!) benevolence. Joseph Heath (2014b) makes this point eloquently:

When it comes to the limits of benevolence, Smith had what might be regarded as a common-sense view of the matter. The average person exhibits a certain degree of altruism, particularly toward family and friends, and can occasionally be motivated to help a stranger. But this disposition is rather limited and subject to exhaustion. When the limit is reached, the individual can be expected to act in a self-interested fashion. Cooperation, however, to the extent that it is vulnerable to free-rider problems, requires that individuals refrain from pursuing their self-interest. It therefore imposes a *motivational burden* upon individuals, which in turn generates a prima facie limit on the extent of cooperation, and therefore an upward bound on both the scale of a society and the level of social complexity (211)

These are the limits of inclusivism all over again, but this time in an institutional context. Therefore, Heath continues, an institutional arrangement, such as the market, which allows individuals to cooperate without tapping into these motives, is an important discovery. It expands the scope of cooperation by allowing society to economize on moral motivation, that is, to get more out of the level of moral motivation that it can plausibly expect (or non-tyrannically demand) from its members (211).

It is frequently noted that literally millions of people cooperate to produce even the simplest items, such as pens. Markets produce this result of peaceful, mutually beneficial cooperation by working with, rather than against, the crooked timber of humanity. Market competition on the basis of decentrally generated price signals achieves this result by channeling people's existing motivations into socially desirable behavior. It is a form of staged competition through which some market participants (the competitors) are deliberately kept locked into a collective action problem to produce positive external effects for other market participants (the rest of society). People cooperate independent of or even against their will, by acting within an institution whose overall rationale reaches through people's intentions to realize the common good.

Once such indirectly cooperative institutions are established, they benefit from feedback loops that tend to engender more directly cooperative dispositions further downstream. Consider the fact that evidence from behavioral economics suggests that people in market societies tend to make the "fairest" offers in economic games such as *Ultimatum* or *Dictator* (Henrich et al. 2010, 6; Brennan and Jaworski 2015, 96ff.). Contrary to what individual self-interest would recommend, people from, for instance, the US, tend to make relatively high and reject relatively low offers. This effect is likely due to an increased familiarity with norms governing exchange between strangers. Familiarity with such norms, in turn, can foster virtues of trustworthiness and dependability (Anomaly 2017).

This last point suggests an objection to my argument, which is that some of the institutional arrangements I have mentioned are not really about bypassing our psychological resources at all. Markets and other institutions do not work around as much as *transform* our moral attitudes, and this fact seems to cut in precisely the opposite direction of my argument. Likewise, science has *enlarged* our minds by providing an extensive repertoire of information and technology to download from our culture and peers. If our attitudes can be transformed like this, then why isn't that how we should counter the challenge from evolutionary conservatism?

It is of course true that social institutions can significantly alter and shape our psychological dispositions. However, my point about the merits of bypassing those dispositions altogether stands, because the chains of cooperation and mutual benefit between millions of people achieved by market institutions have not been brought about by such transformations, and do not depend on them. Rather, the most successful (both in terms of outcome and further scalability) modern institutions achieve certain moral gains – securing cooperation and reciprocal benefit between distant strangers – without the involvement of *any motivation whatsoever* to cooperate with or benefit others. Once such institutions are established, they can then create useful feedback loops which engender trust and other forms of social capital. Transformed dispositions play a welcome role; but the heavy lifting in this "expansion" of the circle of cooperation is done by an institutional arrangement that does indeed bypass, rather than modify, our dispositions. And the level of cooperativeness so achieved is clearly beyond the limits of what even the most optimistic assessment of our onboard psychological resources could consider feasible.

The point generalizes. The market is merely one example for how clever institutional design can render evolved psychological constraints irrelevant. Democratic institutions, for instance, also redirect otherwise morally dubious motives towards socially beneficial goals. The hunger for power, prestige, and dominance may have an evolutionary rationale. Under normal conditions, these dispositions typically lead to socially harmful attempts to benefit oneself, one's family, and one's loyal allies. But when the control of the monopoly in (legitimate) violence is tied to the ability to win elections, one's Machiavellian goals have to become at least somewhat aligned with the interests of the majority. This arrangement is arguably far from perfect (Brennan 2016), but it illustrates that frequently, the best way of

dealing with undesirable psychological dispositions (such as selfishness or the will to power) is not to bludgeon them into shape, but to harness their force.

6 Cumulative Moral Learning

To counter the challenge posed by evolutionary conservatives, we need to find, implement and maintain bypassing institutions that solve the problems posed by our evolved psychological constraints. This is a possibility Buchanan and Powell do not take seriously enough. We don't need more empathy or altruism to make human hypersociality work. We need smart institutional support. The good news is that the possibility of such clever scaffolding can, in turn, be given its own naturalistic defense. The prospects for moral progress aren't so bad, given how much human behavior can be shaped by cumulative cultural institution building and self-domestication (Henrich 2015; Richerson and Boyd 2005).

The idea that evolved limitations to individuals' psychological attitudes could impose constraints on moral progress underestimates the extent to which cooperative dispositions can be "offloaded" into an institutional environment. We don't need to reinvent norms of cooperation with each new generation by finding ourselves sentimentally inclined to be nice to others. Indeed, current models of human ultrasociality and beyond-kin cooperativeness place virtually no emphasis on prosocial attitudes, which is readily admitted to being unable to sustain anything beyond cooperation in small groups (Bowles and Gintis 2011, 4ff.). Instead, or late Pleistocene environment set us on a path towards ever greater cooperativeness, over the course of which we constructed a highly specific niche which increased the benefits of cooperation. Human beings accumulate normative capital which becomes embodied in norms of punishment, communication, and internalization (Sterelny 2012). These normative institutions are what creates and sustains expansions of the moral circle.

7 Conclusion

I have suggested two strategies for developing a more promising response to the challenge from evolutionary conservatism than the one proposed by Buchanan and Powell. One has been to focus on different forms of moral progress, such as demoralization, and the corresponding affective attitudes, such as disgust. These progress/attitude pairs, I have suggested, can secure the same moral gains as empathy and/or altruism, which are traditionally thought to be uniquely linked with inclusivist moral shifts. Secondly, I have suggested that to bring about moral improvements, we should move beyond our psychological capacities altogether, finding institutional ways of bypassing, rather than modifying, our evolutionarily inherited constraints. Markets and democratic institutions come to mind.

Finally, I wish to highlight that these two strategies are not independent from, but complement each other: in many cases, adopting the counterintuitive institutions I have recommended as progressive workarounds is forestalled by morally regressive attitudes of revulsion. Under favorable conditions, markets create mutually beneficial positive sum interactions, yet frequently they are opposed for being "noxious" (Satz 2010) or otherwise morally obscene. Thus, the malleability of disgust can be used to get rid of the opposition towards the very institutional structures that allow us to bypass our limited psychological heritage.

Acknowledgements I would like to thank audiences at Utrecht, Potsdam, and Enschede for their helpful feedback. Two anonymous referees for this journal provided many extremely helpful suggestions regarding how to improve the paper, for which I am grateful as well.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Anomaly J (2017) Trust, Trade, and Moral Progress. *Social Philosophy and Policy* 34(2):89–107
- Appiah, K. A. (2010). *The Honor Code. How Moral Revolutions Happen*. W. W. Norton & Company
- Arruda CT (2017) The varieties of moral improvement, or why Metaethical constructivism must explain moral Progress. *Ethical Theory Moral Pract* 20(1):17–38
- Asma ST (2013) *Against fairness*. University of Chicago Press, Chicago
- Bloom P (2017) *Against empathy: the case for rational compassion*. Random House, New York
- Bowles S, Gintis H (2011) *A cooperative species: human reciprocity and its evolution*. Princeton University Press, Princeton
- Brennan J (2016) *Against democracy*. Princeton University Press, Princeton
- Brennan, J., & Jaworski, P. (2015). *Markets without limits: moral virtues and commercial interests*. Routledge
- Buchanan A, Powell R (2015) The limits of evolutionary explanations of morality and their implications for moral progress. *Ethics* 126(1):37–67
- Buchanan A, Powell R (2016) Toward a naturalistic theory of moral progress. *Ethics* 126(4):983–1014
- Buchanan A, Powell R (2017) De-moralization as emancipation: liberty, Progress, and the evolution of invalid moral norms. *Soc Philos Policy* 34(2):108–135
- Buchanan A, Powell R (2018) *The evolution of moral Progress. A biocultural theory*. Oxford University Press
- Campbell R (2017) Learning from moral inconsistency. *Cognition* 167:46–57
- Chapman HA, Anderson AK (2013) Things rank and gross in nature: a review and synthesis of moral disgust. *Psychol Bull* 139(2):300
- de Lazari-Radek K, Singer P (2012) The objectivity of ethics and the unity of practical reason. *Ethics* 123(1):9–31
- Evans J (2017) A working definition of moral Progress. *Ethical Theory Moral Pract* 20(1):75–92
- Fricker M (2007) *Epistemic injustice: power and the ethics of knowing*. Oxford University Press
- Gert J (2015) Disgust, moral disgust, and morality. *J Moral Philos* 12(1):33–54
- Graham J, Haidt J, Nosek BA (2009) Liberals and conservatives rely on different sets of moral foundations. *J Pers Soc Psychol* 96(5):1029
- Greene J (2013) *Moral tribes: emotion, reason, and the gap between us and them*. Penguin
- Haidt, J. (2012). *The righteous mind: why good people are divided by politics and religion*. Vintage
- Heath J (2014a) *Enlightenment 2.0: Restoring sanity to our politics, our economy, and our lives*. HarperCollins Publishers Limited, New York
- Heath J (2014b) *Morality, competition, and the firm: the market failures approach to business ethics*. Oxford University Press, Oxford
- Henrich J (2015) *The secret of our success: how culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press, Princeton
- Huemer M (2013) *The problem of political authority*. Palgrave MacMillan, London
- Kahan D (1999) The progressive appropriation of disgust. In: Bandes SA (ed) *The passions of law*, vol 1999. New York University Press, New York, pp 63–79
- Kass LR (1997) The wisdom of repugnance. *New Republic* 216(22):17–26
- Kelly D (2011) *Yuck!: the nature and moral significance of disgust*. Cambridge MIT Press, Cambridge
- Knapp C (2003) De-moralizing disgustingness. *Philos Phenomenol Res* 66(2):253–278
- Kumar V (2017) Foul behavior. *Philosopher's Imprint* 17(15):1–17
- Livingstone Smith D (2011) *Less than human: why we demean, enslave, and exterminate others*. St. Martin's Press, New York
- Moody-Adams MM (2017) Moral Progress and human agency. *Ethical Theory Moral Pract*:1–16

- Nussbaum MC (2009) *Hiding from humanity: disgust, shame, and the law*. Princeton University Press, Princeton
- Persson I, Savulescu J (2017) Moral hard-wiring and moral enhancement. *Bioethics* 31(4):286–295
- Plakias A (2013) The good and the gross. *Ethical Theory Moral Pract* 16(2):261–278
- Plakias A (2018) The response model of moral disgust. *Synthese* 195(12):5453–5472
- Prinz J (2011) Against empathy. *South J Philos* 49(s1):214–233
- Rebello, S. (1990). Alfred Hitchcock and the making of psycho. Marion Boyars
- Richerson PJ, Boyd R (2005) *Not by genes alone. How culture transformed human evolution*. University of Chicago Press, Chicago
- Rønnow-Rasmussen T (2017) On locating value in making moral Progress. *Ethical Theory Moral Pract* 20(1): 137–152
- Rozin P, Markwith M, Stoess C (1997) Moralization and becoming a vegetarian: the transformation of preferences into values and the recruitment of disgust. *Psychol Sci* 8(2):67–73
- Satz D (2010) *Why some things should not be for sale: the moral limits of markets*. Oxford University Press, Oxford
- Sauer H (2015) Can't we all disagree more constructively? Moral foundations, moral reasoning, and political disagreement. *Neuroethics* 8(2):153–169
- Sauer H (2017) *Moral judgments as educated intuitions*. MIT Press, Cambridge
- Sauer H (2018a) *Debunking arguments in ethics*. Cambridge University Press, Cambridge
- Sauer H (2018b) *Moral thinking, fast and slow*. Routledge, New York
- Singer P (2011) *The expanding circle: ethics, evolution, and moral progress*. Princeton University Press
- Sterelny K (2012) *The evolved apprentice*. MIT Press, Cambridge
- Summers JS (2017) Rationalizing our way into moral Progress. *Ethical Theory Moral Pract* 20(1):93–104
- Tybur JM, Lieberman D, Kurzban R, DeScioli P (2013) Disgust: evolved function and structure. *Psychol Rev* 120(1):65
- Västfjäll D, Slovic P, Mayorga M, Peters E (2014) Compassion fade: affect and charity are greatest for a single child in need. *PLoS One* 9(6):e100115
- Västfjäll D, Slovic P, Mayorga M (2015) Pseudoinefficacy: negative feelings from children who cannot be helped reduce warm glow for children who can be helped. *Front Psychol* 6(616)