

# UCLA

## UCLA Previously Published Works

### Title

The role of disgust in norms, and the role of norms in disgust research: Why liberals shouldn't be morally disgusted by moral disgust

### Permalink

<https://escholarship.org/uc/item/4kp86089>

### Authors

Clark, Jason A.  
Fessler, Daniel M.T.

### Publication Date

2014

Peer reviewed

**The role of disgust in norms, and of norms in disgust research: Why liberals shouldn't be morally disgusted by moral disgust**

Jason A Clark, PhD  
Postdoctoral Researcher  
Institute for Cognitive Science University of Osnabrueck  
28 Albrechtstrasse #417  
49069 Osnabrueck  
Germany [jasonanthonyclark@gmail.com](mailto:jasonanthonyclark@gmail.com) [sites.google.com/jasonanthonyclark](http://sites.google.com/jasonanthonyclark) +49 0541 878 7902

Daniel M.T. Fessler  
Department of Anthropology  
341 Haines Hall  
Box 951553  
University of California, Los Angeles  
Los Angeles, CA 90095-1553 [dfessler@anthro.ucla.edu](mailto:dfessler@anthro.ucla.edu)  
<http://www.sscnet.ucla.edu/anthro/faculty/fessler/>

*PAPER ACCEPTED FOR PUBLICATION IN  
TOPOI: AN INTERNATIONAL REVIEW OF PHILOSOPHY*

**The role of disgust in norms, and the role of norms in disgust research:**  
**Why liberals shouldn't be morally disgusted by moral disgust**

*The proper course for liberalism is not to obliterate disgust, but to reform its objects so that we come to value what is genuinely high, to despise what is genuinely low... It's no surprise that legal moralizing of this sort has been, and continues to be, an instrument of "brutal and indefensible regimes." But why should the proponents of defensible regimes declare a unilateral cease-fire rather than fighting the indefensible ones on their own terms? Erecting a liberal counter-regime of disgust, I've tried to show, is exactly the aim behind "hate crime" laws, which seek to make the proponents of illiberal species of hierarchy the object of our revulsion. It seems unlikely that a philosophical abstraction as malleable as "liberalism" is conceptually incompatible with this form of legal moralizing. But if it is, so much the worse for liberalism. (p. 1653)*

*...The kind of hierarchic rankings characteristic of [illiberal] disgust are too durable to be driven from the scene by the morally antiseptic idiom of liberalism. Those who believe otherwise are fooling themselves. If we let them fool us, those of us who oppose brutal and indefensible hierarchies in law risk becoming their unwitting defenders. (p. 1656)*

--Dan Kahan (1998), "The anatomy of disgust in criminal law", *Michigan Law Review*, vol.96:1621.

**Introduction**

Recently, a growing number of philosophers, legal scholars and scientists have evaluated the role of disgust in moral judgments, and have found it wanting. They argue that features intrinsic to the emotion itself lead to immoral judgments, or unethical treatment of its objects. Nussbaum (2009, 2010) argues that shame and disgust have no positive role to play in moral and legal judgments, while guilt and anger do. Similarly, Russell and Giner-Sorolla (2011a, 2013) argue that disgust is an "unreasoning" emotion, lacking the cognitive, behavioral, and situational flexibility of anger, with negative moral consequences.<sup>1</sup> Kelly (2011) has also raised concerns about disgust, arguing that the features of disgust which make it well-suited for its primary pathogen-avoidance tasks render it ill-suited for moral judgment; e.g., disgust purportedly has a "hair trigger" and is highly susceptible to false positives, and "the nature of the emotion itself, the slope from moralization to demonization and dehumanization is just too slippery to endorse" (p. 178), even in the service of morally justified goals. Bloom (2013) asserts that "the intuitions associated with disgust are at best unnecessary...and at worst harmful in that they motivate irrational policies and license savage behavior" (p. 155). He argues that, unlike other moral emotions (including empathy) which can *sometimes* go wrong, "[d]isgust is different" in that its verdicts at best

---

<sup>1</sup> Russell and Giner-Sorolla also argue that disgust is more likely to produce "moral dumbfounding", and that people are generally much worse at providing reasons for their moral disgust than they are for other moral emotions. This is a vital issue in evaluating disgust, however, for reasons of space, we must set aside this aspect of their challenge.

contribute no more than a moral “coin toss”, and that if the capacity for moral disgust were suddenly exterminated in humans, this would disrupt society very little, and could only have a beneficial effect.

There is a long history of condemning particular emotions not merely as cognitively inferior, but also as being psychologically harmful, either to those who experience them, or to those toward whom they are directed. Virtually every emotion has been singled out for such treatment at some point. Aside from assessments that emotions in general represent an inferior means for making moral decisions, condemnations of particular emotions typically involve negative comparisons with other emotions that are seen as morally positive. Disgust is frequently compared unfavorably to emotions such as anger, and is embedded in a network of normative assessments and comparisons to other emotions with which it is associated. In this regard, the literature has focused on the relationships among five emotions: anger, disgust, guilt, shame and empathy. Anger is paired with guilt insofar as guilt is held to be the predominant response to being the target of anger, whereas disgust is paired with shame for the same reason (Giner-Sorolla and Espinosa 2011). Empathy is also often seen as an unqualified good in morality, one that is selectively compromised by disgust and shame (Ehrlich and Ornstein 2010; Rifkin 2009). In general, scholars evince a bias in favor of “prosocial”, approach-related emotions, and against “antisocial” and withdrawal-based emotions. Guilt, for example, is held to be superior to shame not only because it is directed at acts rather than at the more global self, but also because, while shame leads to social withdrawal, guilt leads to reparative approach (Giner-Sorolla and Espinosa 2011). Some have also argued that shame is more closely linked to mental health problems (Tangney et al. 2007)

In such comparisons of emotions, differing valuations of specific emotions by individual researchers often result in self-cancelling circles of blame. For example, anger is seen by many as superior to disgust in part because, while anger is associated with guilt, disgust is associated with shame (Giner-Sorolla and Espinosa, 2011), and shame is seen as inferior to guilt (Deonna et al. 2011)—a hierarchy that, while common among Western liberal academics, is decidedly culturally parochially. However, one of the most frequently cited reasons for believing that shame is considered a destructive emotion is its close ties to anger (Tangney and Dearing 2003). As this example illustrates, we believe that pronouncements as to the merits or hazards of the use of particular emotions to regulate social behavior are often based on an incomplete assessment of the relevant considerations. Thorough evaluations are likely to reveal that all emotions can have positive or negative social effects, depending on numerous factors external to the operation of the emotion itself. This includes emotions that are sometimes lauded as unequivocally positive, such as empathy or compassion, and, conversely, those, such as disgust, that have heretofore been largely condemned by Western liberal academics, as is illustrated by the heated 2011 “Tiger Mother” debate in the U.S. about the relative values of Asian parental styles, which leverage shame, high demands and unflinching criticism to drive achievement, and American parental styles which emphasize self-esteem and “trophies for everyone” (Chua 2011).<sup>2</sup>

---

<sup>2</sup> Tracy and colleagues' account of “hubristic” and “authentic” pride is one further example of this phenomenon (Tracy and Robins 2007; Tracy et al. 2009). They argue that these “forms” of pride are differentially associated with a wide range of psychological variables, with hubristic pride being associated with negative characteristics, and authentic pride with positive ones. This split mirrors the conservative/liberal split insofar as many of the characteristics of hubristic pride that are condemned are related to those we discuss below; e.g. hubristic pride is

We believe that research into disgust has been hindered by the importation into empirical work of a variety of normative assumptions, assessments and associations, reflecting moral, typically meta-ethical, norms. Here we challenge both the degree of certainty we should have in the correctness of such normative preferences, as well as the use of these norms in interpreting empirical data. We are not interested in defending conservatism or liberalism, or any of the variables associated with them; rather, our goal is only to show that there is more doubt and controversy surrounding such (meta) ethical assumptions than the literature suggests, and that such assumptions therefore oversimplify the debate. Furthermore, we do not allege that researchers are intentionally “cooking the books” in their experiments in order to support a liberal political or moral agenda, but rather that the broad cultural context in which research is conducted is shaped by (often unnoticed) liberal assumptions, and various related culture-bound factors.

Finally, we believe that an understanding of disgust’s role in morality has been hindered by a reliance on general theories of disgust that we think are mistaken. These accounts have been increasingly challenged, and many investigators are now adopting alternative approaches (Chapman & Anderson 2013; Curtis et al. 2011; Kelly 2011; Oaten et al. 2009; Tybur et al. 2013). Such theories are rich in insights that need to be incorporated into any alternative, more inclusive theory. However, while the existing literature (much less the present paper) does not provide a conclusive refutation of them, here we focus more on inducing theory change via the development of our own and others' positive theories as preferable alternatives, rather than attempting to provide definitive reasons to reject such theories. In the final section, we present portions of our own general theory. We believe that when questions about the moral appropriateness of disgust are reframed within these theories, a stronger case can be made that, under the right circumstances, disgust can be a morally beneficial emotion.

### ***Conservatism, liberalism and disgust***

#### *Moral disgust's associations with conservative authors and values*

Contraposing authors who subscribe to the liberal values that dominate the larger intellectual environment of philosophy and the social sciences, conservative authors have sought to affirm a positive role for disgust in morality. The rejection of conservatism and conservative values, and the association of disgust with them, appears to have played a role in arraying liberal theorists against the emotion itself. For example, the opening salvo in the current debate on the moral status of disgust was fired by Kass (1997). Kass is a conservative bioethicist who chaired President George W. Bush’s Council on

---

associated with pride deriving from character- or identity-based traits, while authentic pride is associated with specific, effort-driven achievements. Holbrook, Piazza, and Fessler ([in press]a, [in press]b) have argued that scales designed to measure authentic versus hubristic pride are measuring not two distinct affective states; rather, the latter scale measures the normative assessments that are imposed by observers on expressions of pride. To this we add the further claim that these assessments themselves reflect the kind of liberal bias that runs throughout discussions of disgust. Similar considerations apply to shame and other negatively regarded emotions (see Deonna et al. 2011 for a defense of shame against these charges).

Bioethics. In “The Wisdom of Repugnance”, he argued that cloning was part of a broader erosion of traditional values, also involving birth control, gay marriage, feminism and other forces that he takes to have undermined conservative values concerning the proper place of sex and reproduction, the natural structure of human families, and the natural order itself. Kass argues that disgust towards these trends reveals the truth of certain deep values, and argues that those who have lost the ability to respond with repugnance to such practices are lacking a fundamental form of human wisdom.

Recent experimental evidence also suggests that disgust is more closely associated with conservatism, as well as specific traditionally conservative values, and more conservative modes of moral reasoning, such as deontological (as opposed to utilitarian) decision making. Politically conservative individuals are more likely than politically liberal individuals to endorse moral codes that emphasize purity (Graham, Haidt, & Nosek, 2009), and tend to be more easily disgusted (Inbar, Pizarro, & Bloom, 2009). Those more prone to disgust are also more likely to express xenophobia, ethnocentrism (Navarette and Fessler 2006), and negative attitudes toward homosexuals and other stigmatized out-groups. Helzer and Pizarro (2011) also found that reminders of physical purity (physical cleaning) influenced specific moral judgments concerning behaviors in the sexual domain as well as broad political attitudes. For example, participants reported being more politically conservative and made harsher moral judgments toward violations of sexual purity. Inbar et al. (2012) found that a disgusting ambient odor led to a selective reduction of warmth perceptions of gay men.

We challenge Kass’ notion that the intrinsic validity of any moral position can be exclusively assessed, or even indexed, by disgust in particular, emotions in general, or, indeed, any form of exclusively intuition-based reasoning or argumentation. It appears that a central feature of the evolved mental mechanisms that acquire, process, and deploy culturally-constituted moral rules is their capacity to transform even moral positions that are historically, socially, and geographically parochial into seemingly self-evident universal facts about the world, such that it then seems “natural” to experience strong emotions when they are violated. Accordingly, what appears obviously and intrinsically disgusting to one observer need not elicit such responses from members of a different cultural tradition, leaving the accidental fact that one was born into a particular group as the principal basis for holding the given moral position. Indeed, such a pattern is patently evident in the case of the position that Kass advocates with regard to family structure and the natural order, as the anthropological literature clearly indicates that a majority of human societies allow – and value – polygynous marriage (Ember et al. 2007). Likewise, employing precisely the language of purity that is often strongly linked to disgust-based argumentation, anti-miscegenation laws were once the norm in the United States, yet today even the vast majority of self-identified conservative Americans approve of interracial marriage.<sup>3</sup>

As we will argue at length later, while Kass is fundamentally mistaken about the validity of disgust as an index or cue of the violation of an objective moral fact, he is nonetheless correct in asserting that disgust

---

<sup>3</sup> <http://www.gallup.com/poll/149390/record-high-approve-black-white-marriages.aspx>

can reveal the perceiver's commitment and sensitivity to deeply held values, independently of the content of those values. On this view, disgust is "norm-neutral", but does have the important positive affordance that its display and attendant behaviors communicate to third parties that the emoter recognizes and disapproves of a norm violation (or an actor having a history of such violations). Hence, in some qualified sense, we agree with Kass that those who have lost the capacity to feel repugnance are morally disadvantaged – not because, as Kass would have it, they therefore lack the ability to intuit putative objective moral facts, but rather because they suffer a reduced capacity to convey their moral disapproval in a social arena populated by judgmental actors.

### *Focus on violations of liberal norms*

Authors criticizing the moral status of disgust have focused primarily on its role in facilitating moral violations that most readers would condemn (as opposed to its role in enforcing norms with which most would agree), including genocide, racism, persecution of sexual minorities, etc. However, disgust has also been shown to support moral practices that most readers would endorse, such as condemnations of hypocrisy, lying, unfairness, betrayal, disloyalty, and even towards the very behaviors that are sometimes seen as selectively reinforced by disgust, such as disgust towards racism, homophobia, and ethnic persecution (Borg et al. 2008 Chapman et al. 2009; Danovitch & Bloom 2009; Olatunji et al. 2012; Rozin et al. 1993; Sherman et al. 2007). Indeed, whereas anti-miscegenation laws were once bolstered by disgust, today such laws are themselves the target of disgust. Anti-smoking efforts and moral vegetarianism are also associated with liberals, and not only has disgust been deployed in the service of these goals (Rozin, Markwith, & Stoess, 1997; Rozin & Singh 1999) but, moreover, has become so central to such efforts that it plays a role even in professional debates on the subject (Alderman et al., 2010). Likewise, there are emerging indications that a similar process of moralization is occurring with regard to obesity, again with a liberal-disgust linkage (see Vartanian, 2010). Furthermore, while disgust may indeed be more closely connected to conservatism, even Pizarro and colleagues present results demonstrating that the judgments of political liberals involve disgust, and they suggest that, at least when it comes to the issue of disgust, the conservative/liberal split may not run as deep as has been suggested.

An example of the use of disgust to support both extreme conservatism and extreme liberalism can be found in the conflict between Dan Savage and Rick Santorum. In 2003 Rick Santorum, then a U.S. Senator, vocally opposed homosexuality, making statements such as ""In every society, the definition of marriage has not ever to my knowledge included homosexuality. That's not to pick on homosexuality. It's not, you know, man on child, man on dog, or whatever the case may be" (Farah 2008)). Dan Savage is a liberal writer, sex-advice columnist, and podcast host, who, together with his husband, created the "It Gets Better Project", a web site on which those who have overcome struggles with their gender can post video messages offering encouragement to those who still struggle with gender issues. In response to Santorum's comments, Savage held a contest to redefine the word 'santorum'. The winning definition was "the frothy mixture of lube and fecal matter that is sometimes the byproduct of anal sex". Savage subsequently encouraged his followers to make this definition the top result on search engines. Liberals

gleefully obliged, and it shot to the top of Google and other search engines, where it has hovered ever since, including the period during which Santorum was a presidential candidate in the 2012 elections.

When Savage was asked whether he was concerned about the effects of this campaign on Santorum's children, he replied:

"My response generally is / have children. There are millions of gays and lesbians in this country who have children, and our children have to listen to the Rick Santorums ... of the world compare us to dog-fuckers and suggest that gay marriage is akin to terrorism. And what about *our* children? Why am I required to be civil to a man who compares my relationship to incest and bestiality and terrorism? And where's the concern for children when gays and lesbians *are* the children? ... The only people who come at me wringing their hands about Santorum's children are idiot lefties who don't get how serious the right is about destroying us." (Mencimer 2010)

Of course, many critics of disgust would condemn Savage's actions (or any other use of moral disgust, even in the service of liberal goals), but there are also many liberals who revel in such a move, which at least makes the moral status of disgust a more complicated question than is suggested solely by the illustration of disgust's role by those who advocate the kind of illiberal values expressed by Santorum.

### ***Disgust and Dehumanization***

#### *Experimental and theoretical support for the disgust-dehumanization link is limited*

In addition to an almost exclusive focus on the illiberal uses of disgust, there has been a selective focus on *disgust's* role in atrocities, to the neglect of other emotions that are clearly involved. Here we focus on dehumanization, which is at the core of most critiques of disgust. Those who criticize the role of disgust in morality argue that disgust is intrinsically "dehumanizing", a consequence that is condemned even when it occurs in the service of defensible norms. Liberal scholars' somewhat selective association of disgust with dehumanization is no doubt due in part to Rozin et al.'s "animal reminder" account of disgust, which has been the dominant theory of disgust employed by critics of moral disgust. According to this theory, disgust originated as a reaction to orally borne toxins and other pathogens, but has come to play an expanded role in human life, now including disgust towards nonoral stimuli such as blood, sex, insects, etc., as well as towards out-groups and moral violations (Rozin et al. 2008). Rozin et al. argue that this expansion was driven in large part by the uniquely human capacity to experience "existential terror" at the prospect of one's own death. On their view, disgust operates to direct our attention away from such paralyzing fear, and thereby performs a Terror Management function, mediated by disgust's ability to remind us that we are mere mortal animals. However, while it is clear that the conjunction of issues of animality and disgust are extensively used to symbolically mark social boundaries, this is more revealing of the cultural deployment of disgust than it is of any essential feature of disgust itself (see De Block & Cuypers 2012). Indeed, while Rozin and colleagues' theory usefully directs attention to such deployments, to the extent that it is read as claiming that coping with existential anxiety is a biologically evolved ultimate function of disgust, the theory is subject to

numerous cogent criticisms with which we agree (see Chapman & Anderson 2013 for review). Correspondingly, recent evolutionary theories of disgust reject the animal reminder account. Of relevance here, such revisions call into question the general theoretical basis for connecting disgust to dehumanization – yes, disgust *can* be linked with dehumanization via animality, but, given that the latter is not an inherent focus of disgust, so too is this linkage not an inherent consequence of the operation of this emotion. We take up such theories below; however, it is worth mentioning at this point given that some of the problems posed for Rozin’s theory apply equally to the concept of dehumanization itself.

Aside from theoretical support derived from the animal reminder theory, most of the evidence used to support the dehumanizing effects of disgust is qualitative in nature, consisting of case studies and historical evidence (e.g., Nussbaum 2009, 2010). Such studies are valuable, but need to be supplemented with empirical research. Only a handful of studies actually purport to establish a direct link between disgust and dehumanization. Bastian and Haslam (2010) distinguished two types of dehumanization: denials of characteristically human features involving animalization, “Human Uniqueness” dehumanization (henceforth “animalistic” dehumanization), and a mechanistic form of dehumanization in which people are seen as objects, “Human Nature” dehumanization (henceforth “mechanistic” dehumanization). Animalization is closer to the type of dehumanization associated with disgust in both the perceiver and the target. Noting the lack of research on the psychological responses of targets of dehumanization (as opposed to perceivers), Bastian and Haslam (2011) tested such responses, using minor everyday slights. Animalistic dehumanization was associated with judgments that targets are seen as animal-like and lacking autonomy, resulting in targets feeling lowered, debased, contaminated, and disgusting, and was associated with shame, guilt, self-blame and self-directed rumination. Buckels and Trapnell (2013) also presented some of the first direct empirical evidence that disgust causes dehumanization. Using arbitrarily generated groups (over- and under-estimators in a guessing task), they found that experimentally induced disgust increased associations of out-group members with animals, and of the in-group with humanity.

Interestingly, however, Buckels and Trapnell found that *all* participants showed such a shift, regardless of disgust induction, and despite the arbitrariness of the groups. In addition, they note that the implicit association task that they used is incapable of distinguishing between the role of disgust in strengthening associations between out-groups and animals, and strengthening associations between the in-group and humanity, as participants showed both changes. This suggests that “dehumanization” may be a more general and fundamental part of our group psychology, rather being disgust-specific, and that exclusion from our moral circles might be a secondary effect of heightened in-group bias rather than out-group negativity. Therefore, it is unclear that these experimental results should be interpreted as unique to issues of dehumanization. Instead they may reflect more general psychosocial dynamics relating to, e.g., in-group/out-group boundaries. Indeed, utilizing a similar minimal-group method – with no emotion induction -- Stürmer et al. (2006) document preferential empathy and helping behavior toward members of the in-group. Given the extended nature of human selves to include membership in and identification with groups of varying size and distance from the self, this raises the possibility that in some cases, dehumanization reflects distance from the extended self more generally. The boundaries of our moral circles and in-groups are ever shifting, and rather than always involving a specific comparison

with animals, distance from the self may de facto put out-groups closer to animals insofar as our “self” extends to the human-animal boundary. Insofar as moral disgust reflects these more general patterns, and insofar as such in-group/out-group dynamics can be adaptive, this suggests that phenomena labeled as dehumanization may come in degrees and may not always catastrophically damaging, or even negative.

#### *Disgust is not uniquely involved in dehumanization*

Even if this limited literature suggests that disgust is associated with dehumanization, it fails to establish a *singular* role for disgust in dehumanization, in part because this research has rarely examined whether emotions other than disgust also produce similar results. For example, Buckels and Trapnell (2013) (and many previous studies of disgust) used sadness as a control negative emotion, and so cannot speak to whether anger, fear, or other negative emotions would have the same dehumanizing effect as disgust. Indeed, recent work by Giner-Sorolla and Russell (in prep.) has now shown that anger, disgust, and fear all contribute uniquely to animalized dehumanization. They found that (a) anger, disgust, and fear each contributes to dehumanization, (b) these emotions were differentially involved in different forms of dehumanization, with all contributing to animalistic dehumanization, but only fear showing a connection to mechanistic dehumanization, and (c) anger, fear and disgust mediated animalistic dehumanization in different ways, based on appraisals of threat and action tendencies. Anger and fear were most fully accounted for by perceptions that the out-group had the desire to cause harm, and anger’s effect was fully explained by the participant’s own hostile intents towards the group, while hostile intent only partially mediated fear and disgust. Disgust was most closely connected to perceptions that the group is unnatural or poses a contamination threat, though contamination was also connected with fear (but not anger). Hence, many emotions appear to produce effects associated with dehumanization, including those such as anger that are held to differ from disgust in their potential to generate dehumanization.

#### *Problems with the concept of dehumanization*

In addition to the fact that disgust is not uniquely linked to dehumanization, the very concept of dehumanization is itself problematic. While dehumanization is primarily construed in terms of seeing others as animals, rather than human selves, it ranges over a huge variety of theories of human/animal nature, and many types of comparisons of humans with many types of nonhumans (e.g., “super humanization” involves comparisons with either favorable [e.g., angelic] or unfavorable [e.g., demonic] traits, or degrees of traits, associated with gods or other superhuman entities). Hence, “dehumanization” in this sense depends greatly on which features of the target and animal are in question. Further complicating this picture, cultures differ greatly in their conception of the nature and extent of human-animal boundaries. Even within Western cultures, the human-animal distinction is often a vague and variable one: for example, we often conceive of pets as sentient selves who are a part of our family. Correspondingly, some of the same brain areas involved in social cognition about humans are also active when we mentalize about other animals, such as dogs. Indeed, Barrett (2005) has suggested that the capacity for theory of mind (i.e., understanding what other agents do or do not

know, want, feel, etc.) evolved in part in response to the challenges of dealing with animals, as mentalizing facilitates effective hunting and anti-predatory tactics. Taken together these observations indicate that the phenomena commonly termed dehumanization are unlikely to involve simply a matter of turning off social cognition in response to perceived animality. In addition, animalization is not merely associated with animal-like qualities, but also judgments that targets are childlike, immature, coarse, irrational, backward, incompetent, unintelligent, unsophisticated, uncivilized, hopeless, stupid, and lacking autonomy (Bastian and Haslam 2011, 2010). While some of these may involve direct animalization, they also refer to negative human characteristics as such.

The associations between disgust, dehumanization, and maltreatment of others are further complicated by issues of moral valence. Specifically, the notion that attributes associated with humans are uniformly viewed as positive, while those associated with animals are uniformly viewed as negative, is belied by the fact that observers frequently speak glowingly of the characteristics of animals. Granted, sometimes such positive appraisals represent anthropomorphisation (e.g., “The noble lion, lord of the savannah,” etc.). However, often, the positive valence does not rely on analogies with humans (e.g., “The fleet cheetah,” “The graceful gazelle,” etc.). Lastly, similar considerations apply to mechanistic dehumanization, as we often compare ourselves favorably with machines.

While we concur that negative comparisons with animals play an important role in rhetoric and propaganda designed to both justify and motivate social exclusion and exploitation, our key point here is that previous authors have pilloried moral disgust on the grounds that (i) disgust guards the boundary between humanness and animality, (ii) moral treatment of a target is premised on particular psychological attributes, and (iii) dehumanization involves equating the target with animals, thereby denying the target’s possession of the definitional attributes, and thus (iv) via dehumanization, disgust promotes exploitation. Our position is that the process of framing a target as unworthy of moral consideration is not intrinsically linked to the possession of attributes that are viewed as uniquely human. Accordingly, “dehumanization” does not inherently devolve to issues of humanity and animality. Thus, to the extent that disgust guards the human/animal distinction (a proposition itself subject to challenge), it is not inherently culpable in the process of exploitation. Note that we do not contest the fact that propagandists often seek to motivate genocidal actors through the language of hygiene, that is, by comparing out-group members to disease vectors, with all the attendant implications for disgust. Nor is this to say that such excesses cannot occur in the defense of liberal values. Rather, our position is that such tactics constitute a violation of norms concerning the treatment of others who have behaved in a morally disgusting manner, and are not an immediate, inevitable upshot of disgust. Such excesses can occur with respect to any emotion (e.g., moral anger), and have more to do with norms concerning humane treatment of those of whom we disapprove than they do with the judgment that an action or person is morally disgusting.

### ***Meta-ethical Assumptions in the Moral Psychology of Disgust***

#### *Utilitarian versus deontological reasoning*

In addition to empirical characterizations of the role that emotions play in morality, moral psychologists often adopt meta-ethical positions concerning the “proper modes” of moral decision-making, judging participants’ responses as “correct” or as “errors”. One example of this is the growing tendency for moral psychologists to (implicitly or explicitly) accept the meta-ethical judgment that utilitarian moral reasoning is superior to deontological decision making, and to normatively assess emotions and other cognitive processes in terms of whether they facilitate utilitarian or deontological norms (e.g., Baron & Ritov 2009; Greene et al. 2009; Sunstein 2005). However, the superiority of utilitarian versus deontological reasoning is hardly a settled issue within ethics, law, and other explicitly normative disciplines (Bartels and Pizarro 2011), and the importation of such meta-ethical assumptions into ostensibly scientific investigations has the potential to excessively narrow and skew the design and interpretation of experiments in moral psychology.

While we acknowledge that there is no evidence directly linking disgust with deontology, we believe that disgust is embedded in the network of variables that surround the utilitarian/deontology debate and that the evidence for an indirect connection merits a discussion these issues. More specifically, (a) conservatism is linked to deontology while liberalism is associated with utilitarianism, (b) utilitarianism is often connected to harm/fairness reasoning and act-based assessments, while deontology is linked to less flexible norms and character-based assessments, and (c) the interpretation of many of the central experiments concerning disgust and empathy is influenced by such meta-ethical views

To illustrate the bias described above, consider the following: in experiments designed to demonstrate the negative effects of a lack of mentalizing and empathy towards certain groups, the failure to exercise such capacities via deontological reasoning is often seen as morally or psychologically negative, even (or sometimes especially) if it leads to more calculated utilitarian judgments. At the same time, in experiments designed to highlight the advantages of “unbiased” utilitarian reasoning, the comparative lack of emotion-based responses is often presented as a positive trait. In sum, researchers have an uneasy relationship with such conclusions, and it appears that academics generally want people to make utilitarian decisions, but want them to experience difficulty when making them (see Pizarro & Tannenbaum 2011 for further discussion of this issue).

Such uneasiness leads to conflicting interpretations of empirical results. For example, a cornerstone of many arguments against disgust is a set of neuroimaging studies by Harris and Fiske (2006, 2007, 2009, 2011) in which they argue that disgust produces dehumanization of stigmatized groups. They probed the neural correlates of judgments of the warmth and competence of social groups, finding that contemplating groups judged as high in either warmth or competence correlated with increased activity in the medial prefrontal cortex (mPFC), which is closely linked to various forms of social cognition, such as theory of mind and empathy. While contemplating those who scored low in both categories, and who elicited the most disgust (the low-low condition, including the homeless and addicts), participants failed to recruit mPFC, while they did for all other combinations. Low-low groups were selectively associated with heightened activity in the insula and amygdala (which play an established role in disgust processing), and Harris and Fiske construe such activity as reflecting disgust and dehumanization. Similar results were found when participants were asked to mentalize about daily life for individuals from

stigmatized groups, and to rate them on various dimensions of human perception. In particular, anterior insula negatively correlated with “warmth perception” (which Harris and Fiske interpret as indicating less disgust).

On the other hand, Krendl et al. (2013) examined whether impressions of homeless people, and the correlated neural activity, could be altered by perceptions of controllability versus uncontrollability. Stigmatized individuals whose status was presented as *controllable* evoked mPFC activity, whereas networks including the insula, anterior cingulate cortex, and orbitofrontal cortex were engaged in response to those whose condition was perceived as *uncontrollable*. This is surprising in light of the linkages between decreased mPFC activity and dehumanization, and between the insula and disgust, that are proposed by Harris and Fiske. Krendl et al. interpret the activation of the insula as indicating not disgust, but rather emotions such as pity and empathy, together with emotional conflict and subsequent regulation, all of which are also associated with the insula. They hypothesize that uncontrollability required less attention to intentions, and also note behavioral results indicating that uncontrollable conditions resulted in more pity and willingness to help, and less disgust compared to the controllable scenarios. Hence, insula/amygdala activation is not always interpreted as reflecting disgust, and there is a tension in the interpretation of mPFC/insula activity that reflects a variety of factors. This interpretational ambiguity runs through many additional studies (see Clark forthcoming for an extended discussion of such neuroimaging studies).

The problems with an uneasy bias in favor of utilitarian reasoning are compounded by the fact that the kinds of one-off scenarios typically used to test utilitarian judgments are too simple to determine whether participants’ refusals to engage in single actions that maximize local utility reflect deontological reasoning rather than their commitment to utilitarian rules which, while they require local violations of utility value, promote the good in a global sense. Indeed, in the teaching of ethics, such scenarios are commonly used to prime students’ intuitions that act utilitarianism is an inadequate version of the theory, and that in order to best maximize utility, we require more global rules which facilitate overall good, despite requiring acts that fail to maximize utility in a given situation. Such rules are notoriously difficult to formulate, and are plagued by counterexamples, so it is easy to assume that people's inability to offer them reflects a blind commitment to an absolute deontological norm.

Furthermore, preliminary results suggest that utilitarian reasoning is associated with a variety of secondary traits that are viewed negatively in terms of moral or psychiatric norms. For example, psychopaths (and those with “acquired sociopathy” as a result of damage to the mPFC) more consistently apply utilitarian reasoning than do normal controls, as do alcoholics and depressed patients (Bartels and Pizarro 2011). In addition, Bartels and Pizarro found that even in normal subjects, the tendency to endorse utilitarian reasoning in a variety of footbridge-like dilemmas was associated with higher scores on measures of psychopathy, Machiavellianism and meaninglessness, which are in turn associated with emotional callousness and lack of empathy, manipulativeness, cynicism, and detachment – attributes often seen as both psychologically unhealthy and morally condemnable (see below) even by those who endorse utilitarian approaches. We do not adduce such associations in order

to argue against utilitarian reasoning, but rather note them only to illustrate that, whatever the flaws of deontological reasoning, utilitarian reasoning also keeps some bad company.

Finally, the notion that empathy is a universally positive emotion is questionable. Many philosophers and moral psychologists see empathy as lying at the core of what it is to be a moral agent (Rifkin 2009; Ehrlich and Ornstein 2010). Baron-Cohen (2012) identifies evil precisely with “empathy erosion”. However, the centrality of empathy and its potential to produce morally sound behavior has been questioned even among those rooted in the liberal tradition. In one recent critique of empathy, Bloom (2013b) notes that empathy can be “parochial, narrow-minded, and innumerate” and that “We’re often at our best when we’re smart enough not to rely on it”. He points to various biases engendered by empathy, such as the “identifiable victim effect”, in which people are more likely to feel empathy for, and offer help to, particular victims whose presence is made salient, while at the same time failing to exercise empathy and provide assistance to those whose suffering is more abstract. We (even as infants) are also more likely to exercise empathy to those who are like us than those who are not, which can skew objective moral judgments. Hence, empathy can result in differential treatment based on the spatial, temporal, personal, or cultural distance from the self. As Bloom notes:

“A ‘politics of empathy’ doesn’t provide much clarity in the public sphere ... Typically, political disputes involve a disagreement over whom we should empathize *with*. Liberals argue for gun control, for example, by focusing on the victims of gun violence; conservatives point to the unarmed victims of crime, defenseless against the savagery of others ... So don’t suppose that if your ideological opponents could only ramp up their empathy they would think just like you”.

#### *Condemnation of particular types of general moral codes*

Most attempts to characterize disgust’s role in norm enforcement have focused on efforts to link disgust to the violation of particular classes of norms or value systems (Rozin et al. 1999; Russell et al. 2013). Initial efforts to categorize value systems and their associated emotions were intended as a descriptive psychological and anthropological enterprise. Normative evaluations of different types of value system were not part of this project, but have been imposed on them after the fact, such that condemnations of disgust have the (often unstated) implication that the types of norms that elicit disgust are somehow inferior to other types of norms. For example, Rozin et al. (1999) proposed the CAD Triad, linking value systems based on community, autonomy or divinity with contempt, anger and disgust, respectively. Disgust has also been alleged to be selectively associated with “purity” violations (a concept related to “divinity” norms) alleged to protect the body and soul from contamination. In the disgust debate, divinity norms are in turn often derided, while community-based norms are frequently ignored altogether.

The mere association of disgust with divinity is likely to turn liberals against disgust, insofar as conservative values are more closely associated with religious norms or justifications, and liberalism is committed to secular forms of government. Furthermore, while the concept of “purity” is meant to include abstract forms of contamination (such as pollution of the soul), the vast majority of studies that

have probed the purity construct have used purity violations that involve the body, many of which are known to independently induce non-moral forms of disgust – indeed, we know of no experiments specifically examining the link between disgust and non-bodily violations, e.g., blasphemy (see Russell & Giner-Sorolla 2013 for a similar criticism within the context of a defense of the purity-norm view). Bodily norms, in turn, are seen within the liberal tradition as being more conventional than other norms. Rejecting or neglecting other types of value system, most critics of moral disgust assume or argue that autonomy must be the primary cornerstone of our moral and legal systems, and that autonomy codes put an emphasis on “harm/fairness” violations, rather than, e.g., purity codes, or valuing the community over the individual. Hence, anger and disgust are respectively lauded and criticized not merely based on the intrinsic properties of the emotions, but also in virtue of their alleged association with value systems perceived by secular Western liberal academics as themselves being superior or inferior.

Despite the predominance of the perspective described above, a closer reading of relevant literatures calls its foundation into question. First, the notion that fairness and harm should be the primary basis of our moral reasoning is contentious even among Western liberals, and is made more so by the enormous range of observable cultural variation in the role of such norms, as well as in different cultures’ conception of what constitutes “harm” or “fairness” (Henrich et al. 2010). Furthermore, as we discuss below, we disagree that moral disgust (or other moral emotions) are selectively associated with any particular class of norms, holding that the occurrence of disgust in norm enforcement is due to structural properties of the emotion that are effective in handling any moral violation that possesses particular contextual and relational features. Insofar as our argument provides a better alternative to the selective assignment of specific emotions to particular value systems, this challenges the inference that disgust is inferior to anger (or other emotions) based on its privileged connection to morally condemned moral codes.

#### *Behavioral, cognitive, and moral inflexibility of disgust*

Disgust is often seen as a cruder form of cognition, involving intuitive versus rational decision making, quick and dirty processing rather than slower deliberative processing, and categorical rather than act-based judgments, whereas other emotions are seen as engaging more sophisticated forms of cognition. Disgust is also held to be less flexible in its attendant behaviors, effects on learning and memory, and sensitivity to contextual features (see Russell and Giner-Sorolla 2013) for a review of these findings). We believe that (a) disgust is more flexible and sophisticated in these respects than has been portrayed, and that (b) to the extent that disgust is less flexible in some of these ways, such inflexibility plays an adaptive function under some circumstances (see Clark forthcoming for an extended discussion of these issues). Even the simplest forms of disgust are more flexible than has been suggested. Disgust is held to be behaviorally inflexible, resulting only in reflexive avoidance. However, while disgust can lead to avoidance, if a disgusting object is on us, or in our space (broadly construed) we will approach it to remove or exterminate it. Additionally, disgust is linked to an information-gathering function that plausibly includes approach independent of disinfecting actions (Eng 2008). Disgust can also lead to aggression-like behavior when the threat is mobile, and must be exterminated or driven away, and can also generate cleansing behaviors. It is also highly sensitive to context, and context is among the most

reliably and strongly coded features of conditioned taste aversion (Reilly & Schachtman 2008), and can change as a result of context or cognitive reappraisal. For example, assessment of the valence of a smell may differ greatly depending on contextual information, e.g., whether one is told that an odor is due to dirty socks or Parmesan cheese (Herz 2003), and the effectiveness of such a reappraisal is demonstrated by the fact that, following reappraisal, people are generally willing to consume the source of the odor. Furthermore, other emotions may be inflexible in their own ways. When anger is not “satisfied” in a particular case, it can persist and spill over onto unrelated judgments or targets, as subjects have been shown to punish subsequent offenders more harshly if justice was not served in the original case (Goldberg et al. 1999).

Furthermore, how we respond to others who have been physically contaminated depends enormously on our relationship to them. Parents approach and clean children, we care for our sick friends and relatives, and so on. Such dynamics also seem to characterize moral disgust, where the proper response to psychological “contamination” is to distance and cleanse oneself before resuming contact. Thus, disgust serves a higher-level function that is not merely about the withdrawal of the self, or others distancing themselves from us, but rather motivates a pattern that potentially enables re-incorporation. Basic forms of disgust in other animals are flexible and sensitive in another way as well, viz., in their interactions with other motivational states and emotions. If one is experiencing intense desires, disgust can be down-regulated (Borg & de Jong 2012). If one is experiencing intense disgust, such appetitive states can be down-regulated. Moral disgust is also in continual interaction with other more complex sociomoral emotions, including positive affects such as compassion, empathy, love, moral elevation, familial attachment, and the sense of spiritual purity, and these emotions can modulate disgust reactions. Processes involved in moral disgust also appear to be sensitive to situational appraisals of blame (Krendl et al. 2012).

#### *Character-based versus act-based judgments*

While disgust is more flexible than alleged in these respects, there is nevertheless evidence that it is less flexible than other emotions in certain respects. In particular, increasing evidence indicates disgust is more closely related to assessments based on stable features of individuals (such as character or social identity) than are emotions such as anger. While anger more frequently involves appraisals of acts, disgust is associated with more rigid categorical assessments that are less likely to take into account consequences, excuses, justifications, or intentions connected to acts, and are therefore more difficult to change in the face of cognitive reappraisals than anger. This results in lower sensitivity to situational variables, and a focus on what violations tell us about more stable features of violators, rather than on isolated acts. The corresponding emotions experienced by the targets of disgust and anger, namely shame and guilt, respectively, appear to follow the same pattern, with shame being directed onto more stable features of an individual than guilt (Deonna et al. 2011).

The connection between disgust and character is supported by recent work by Russell and Giner-Sorolla, who, despite being among the most prominent advocates of a purity- or bodily-norm view, have begun to challenge this interpretation of their own results. In Russell and Giner-Sorolla (2011b), they

manipulated an agent's (a) desire to cause harm, and (b) actual harmful consequences of an action. They found that desire to harm was selectively associated with disgust, while harmful consequences predicted only anger. Measurements of participants' inferences in such judgments indicated that perceptions of bad moral character partially mediated the link between disgust and desire to harm, the idea being that such desires in and of themselves are evidence of an attribution of wrongness to more stable features of the agent's identity. Ciaramelli et al. (2013) provide additional support for the preferential link between disgust and character or identity. They found that patients with mPFC damage (BA 10 and BA 32) had a selective deficit in disgust responses to stigmatized groups, and were more willing to engage in contact with them than were control participants. Specifically, while their capacities for judging particular immoral acts remained intact, they were deficient in their ability to generate disgust-based avoidance and moral condemnation on the grounds of social identity, status, or other more stable categorical features.

We believe that the comparative cognitive inflexibility of disgust is an appropriate response to norm violations that appear to be due to character or identity. Such categorical classification is a highly important capacity for navigating the social world, whatever its flaws or limitations, as indicated by the fact that its absence leads vmPFC patients into countless costly and destructive relationships (Damasio 1994; Damasio et al. 1994; Damasio 2005), and opens them to exploitation (e.g., age-related deterioration of the vmPFC correlates with vulnerability to phone scammers in the elderly Denburg et al. (2007)). While anger and guilt are held to be pro-social emotions that can lead to reparations or reconciliation, whether others accept your expressions of guilt depends in large part on what your past behavior reveals about your character. One can only express guilt about a repeated action so many times before the action is attributed to one's character rather than situational factors, and in such cases, what is required to mend the relationship is not merely reparation and another apology, but a more extended reformation of one's character. Hence, when it is someone's character or identity that is called into question, assessors are often better off adopting more categorical and less flexible responses, and interpreting individual acts against this background, adopting a bias against explanations of their actions in terms of context, mitigating circumstances, excuses, reasons, or expressions of regret.

While we believe that the value of character judgments is clear, it conflicts with the liberal belief that moral and legal judgments should be act-based rather than person-based. However, like the question of deontological or utilitarian reasoning, or the assessment of some value systems as morally superior, the notion that character should not play a role in our moral and legal reasoning is far from a settled issue, and various traditions (e.g., virtue ethics) have situated character at the center of their accounts of what it is to be a moral being. Bucking the trend of condemning character-based moral judgments, Pizarro (2011) argues that character judgments are a primary psychological process that is likely to have been generated and stabilized by evolutionary forces, and he proposes that we adopt a person-centered theory, rather than (or at least in addition to) an act-based account. He suggests that character is essential for assigning blame, rather than determining whether or not a moral norm has been violated, such that agents whose actions are attributed to characterological traits are seen as more worthy of blame. He points out that while many of the decisions made by subjects in moral psychology

experiments are characterized as moral “errors” from act-based approaches, an account rooted in character judgments can make sense of these decisions.

Specifically, when people are given minimal, incomplete, or ambiguous information about controllability or intentionality, they are likely to take character information into account when asked to arrive at an estimate of these features. Pizarro notes that “applying information about an individual’s previous acts, his or her known behavioral tendencies, or his or her character traits is a valid (albeit not perfect) way to make an assessment, much as we would apply base rate information when making other kinds of judgments under uncertainty”. He also points out that this can help to explain why people’s reactions to utilitarian decisions can vary depending on whether the person finds the decision to be easy versus hard, or depending on their perceived motivations for making the decision, despite the fact that the outcomes are exactly the same. For example, people are more likely to condemn those who perform difficult decisions easily or immediately, compared to those who show painful deliberation. Echoing the point above concerning the ambivalence of researchers’ interpretations of mPFC versus insula/amygdala activity as reflecting disgust or empathy, Pizarro suggests that, from a character-based perspective, one does not merely want individuals to perform the right act; one wants them to do it in the right way and for the right reasons.

Even if it were possible to exclude character judgments from the legal determination of guilt or innocence, or sentencing, it is difficult to imagine that the effects of criminal convictions could be shielded from subsequent character judgments. In some cases, spending two years in prison on a felony charge, deprived of rights and liberties, may be the least consequential aspect of a criminal conviction, given subsequent denials of employment, travel, housing, financing, etc., based on what a criminal charge says about one’s character. Moreover, it is important not to forget that character assessments can also be positive, and that we grant those with good character benefits on this basis, a practice few would reject. In short, judging character is a critical component of social interaction – indeed, it is difficult to imagine how social life would be possible without both negative and positive assessments of character, be they based on direct observation of an actor’s actions or knowledge of an actor’s reputation. At the most elementary level, extensive cooperation, the hallmark feature of our species, can only have evolved in conjunction with the capacity to assess another’s character and maintain a stable mental representation thereof (Nowak 2006). Given the foundational role that character assessment plays in much of what all observers, be they liberal or conservative, value about human behavior, there are grounds for embracing, rather than rejecting, emotions that serve to incorporate those of good character into the community and exclude those of bad character.

### ***An Alternative General Theory of Disgust's Role in Norms***

As noted repeatedly above, support for the idea that disgust is a morally pernicious emotion comes directly from a number of more general theories of disgust, e.g., Rozin et al.’s “animal reminder” theory, and the CAD triad of types of normative systems. As we have also noted, we agree with the growing number of those who argue that such theories are flawed and that there are more attractive alternatives that both erode the criticisms of disgust, and point the way to a more nuanced account of

the role of disgust in norm enforcement, thus making clear why disgust is a valuable moral capacity. Here we sketch our own theory, one that is based upon, or has affinities with, a number of similar views (Kelly 2011; Tybur et al. 2012). On this view, the structural properties of basic forms of disgust were preadapted to deal with threats that went beyond the mere oral incorporation of pathogens. These include the ways in which the cognitive properties of basic disgust discussed above are well suited for certain normative tasks or contexts, but go beyond them. In particular, a central premise of the view is that disgust functions as a low-cost alternative to anger in punishment and other aspects of norm enforcement. Russell and Giner-Sorolla (2013) argue that one reason why anger's behavior is more flexible is that it is responsive to strategic considerations. They note that a reasoned and reparative approach can be less risky and more productive under some circumstances, and in conflicts with a physically or socially stronger person, avoidance or the recruitment of social support may be the better strategy. They argue that anger possesses such flexibility, while disgust does not. We believe that some of these responses are not (or at least not solely) best attributed to anger. Instead, we argue that (i) disgust is itself highly responsive to strategic considerations, and (ii) represents a strategic alternative to anger that embodies the advantages attributed to anger by Russell and Giner-Sorolla, and others.

In Clark and Fessler (in prep.) we propose an evolutionary account of how disgust's role has expanded in humans, coming to serve functions beyond those for which it originally evolved, viz. protection from the ingestion of pathogens. A full overview of the evolutionary pressures and affordances that have produced moral disgust is beyond the scope of this paper. Nevertheless, some evolutionary background will be illuminating. One stunning fact about disgust is that in other animals, its role appears to be restricted solely to avoidance of orally consumed pathogens, in the form of conditioned taste aversion – among nonhuman species, disgust plays no role in the avoidance of nonoral pathogens (such as insects, bodily fluids, etc.), mate-choice, and other uniquely human forms of disgust. Most relevant for the present purposes, unlike anger and aggression, whose roles in (proto)norm enforcement run deep in ancestral species, it is only in humans that disgust plays a role in normative dynamics.

We believe that disgust's role in uniquely human patterns of norm enforcement is due in significant part to the fact that disgust expressions and behaviors have no known signaling functions in other animals – even in processes as simple as learning food aversions, signs of disgust in conspecifics do not influence the observers' tendency to avoid the associated foods (Galef and Osborne 1990). Signaling functions are essential for emotions to play a role in norm enforcement. Signaling anger towards a norm violator has the advantage that one's disapproval of the violation, one's personal commitment to enforce the norm, and the punishment administered, are clearly conveyed both to the violator and to third parties. However, despite such advantages, signaling anger can be a highly costly move, as it increases the chances of aggressive confrontation and retaliation. In such situations, avoidance may be preferable. However, fear-based avoidance appears to enhance the power of the violator, and lower that of those who flee rather than punish. While mere avoidance (moving away) may impose some cost in the form of ostracism, simple avoidance of this type sends no clear, specific and active signal of disapproval. Such was the condition of our nonhuman primate ancestors.

We propose that disgust filled a niche in normative psychology by providing a means of actively signaling disapproval to norm violators, backed by the threat of exile or extermination, but without the potential costs of direct aggressive confrontation. One finding that supports this idea comes from Pond et al. (2012), who found that disgust sensitivity is negatively correlated with aggression. More specifically, those high in disgust sensitivity were found to have decreased (a) trait physical and verbal aggression, (b) behavioral aggression, (c) physical aggression, and (d) approval of romantic partner violence. They speculate that these negative correlations may stem from a desire to avoid the disease threats involved in aggression, and suggest that disgust may lead to ostracism rather than attack as a protective tactic. Fessler et al. (2004) found that disgust and anger led to differential changes in risk-taking among men and women, a pattern to be expected based on strategic considerations concerning differences between men and women in the risks and rewards of direct aggression.

Emotion researchers generally focus exclusively on either the individual or the dyad, and they rarely attend to the game-theoretic consequences of signaling across groups; however, disgust's proposed advantages play a role not merely at the level of dyadic signaling, but also at the level of groups. Signaling disgust not only affects the target, but also third-party assessments of the violator, and can serve to recruit support. Disgust is also especially preadapted for ecologies of higher-order punishment (which are unique to humans) wherein the social costs of norm violation are in essence "contagious". Hence, it will often be cheaper to treat norm violators as if they carried transmissible disease than to attack them, and active disgust-based signals of disapproval to third parties augment avoidance by protecting the actor from being punished for (implicitly) supporting the norm violator.

The relationship between the costs of various forms of actions taken against norm violators and the signaling component of reactions to norm violators comes into stark relief once it is recognized that there will almost always be more individuals who are unaffected by a given actor's behavior than there are who are affected by it. This simple dynamic has profound implications for the creation of a moral order. As has been redundantly illustrated throughout history, the exploitation or extermination of others does not depend on united action by all possible actors, but rather hinges on action by a few, and inaction by the rest. The fact that it is costly, sometimes impossible, to aggress against a much larger coalition means that those who would exploit others rely on a lack of coordination among third parties who might object to their actions. If, however, observers can signal to one another that they disapprove of the given action, pluralistic ignorance is eliminated, and coordination is facilitated.

While active intervention by third parties is sometimes the only way to stop exploitation, because of the power of large coalitions, mere signaling of disapproval by many observers will often suffice in this regard, as actors will recognize that, should they persist, a large coalition is likely to coalesce against them. Signaling disapproval can thus preclude both the exploitation itself and the costs to observers of intervening. The obstacle in such situations is often the risk that a disapproving observer will become embroiled in the event before sufficiently broad signaling has occurred such that the power of large numbers will reduce the costs to that observer. From the perspective of the observer, moral outrage is thus a costly reaction, as anger often leads the individual to confront the offending party directly. In contrast, moral disgust motivates distancing the self from the offending party while signaling one's

disapproval to that actor and any other observers. Moral disgust is thus an inexpensive means of overcoming pluralistic ignorance, amassing a large set of potential punishers, and thereby truncating exploitation at little cost to third parties. This affordance for low-cost maintenance of the moral order means that observers will often play their part, in contrast to the high demands that active individual intervention places on them.

Hence, the core conceptual features of disgust present extensive affordances for leveraging the signaling power of moral disgust in maintaining the social order. Were we to suggest an overarching function for all the forms of disgust, it would be the protection of the boundaries and integrity of the “self” against pathogenic foreign elements, including substances, organisms, agents, or cultures. The “extended self” is not limited to the body, or even the individual, but rather radiates out through a series of more abstract categories with which we identify, including possessions, personal spaces, and in-groups of various distance from the self (e.g., family, friends, organizations, cultures). Thus disgust responses are also broadly suitable for rejecting norm violators, at least when seen at a higher, more symbolic level of description, where the boundaries of the self are extended to social groups (a ‘social perimeter’), and rejection, expulsion, etc. are seen as removing the offending actor from the ‘body politic’ rather than the literal body. In these domains, disgust serves to mark the normative boundaries surrounding our social and moral identity, leading to avoidance or expulsion of those who violate such norms.

### ***Conclusion***

We have sought to establish the following points:

(1) Disgust has been assessed almost exclusively in terms of its role in facilitating (often extreme) violations of dearly held liberal values. It does indeed play a role in such violations, but it is also applied in the service of core liberal values.

(2) The idea that disgust is intrinsically dehumanizing is plagued by various problems. This notion is supported by general theories that are themselves flawed; there is very little empirical evidence for a disgust-dehumanization connection; the interpretation of that evidence as indicating dehumanization is questionable, and may instead indicate more general psychological tendencies related to shifts in in-group/out-group boundaries; comparisons between humans and nonhumans are too heterogeneous to be encompassed by a single concept, both in terms of the entities with which humans may be compared, and whether such comparisons are favorable or unfavorable; and finally, disgust is not uniquely involved in producing dehumanization.

(3) The notions that utilitarianism is superior to deontology, or that some types of value systems are superior to others, are highly contentious and rooted in an ethnocentric perspective. We have not attempted to settle the meta-ethical debates we have discussed. Instead, we hope to have pointed out that questions about (a) whether utilitarianism is superior to deontology, (b) whether some types of value systems are superior to others, (c) whether moral flexibility is preferable to moral inflexibility, or

(d) whether moral principles and evaluations should be act-centered rather than character-based, are more contentious than has been alleged, and that the importation of simplified, unsettled meta-ethical conclusions into empirical research has the potential to make a mess not only of ethics, but also of empirical disgust research itself, especially when the authors and discussions are not rooted in a solid understanding of the complexity of debates about ethical theories, or the cultural variation in which kinds of general principles are adopted.

(4) For two reasons, the wholesale condemnation of particular emotions is rarely likely to be a successful strategy. First, the value or liabilities of a given emotion are not inherent to the emotion per se, but rather depend on the context in which the emotion is deployed and the nature of the problem being addressed. Overall, while all emotions can go wrong under some circumstances, they generally go right, which is why they have been developed and retained throughout the course of human evolution. Second, given that there are many types of norms, both within and beyond morality, and that there are comparatively few emotions that have to regulate such norms, it makes prima facie sense that the functions of emotions in norms attach to more general structural features of moral violations rather than to particular normative domains. Discussions of the effectiveness of shame and pride, for example, tend to emphasize such broad features (e.g., whether the emotion is directed onto an isolated action (guilt) or whether it was directed onto more stable and global features of the violator (shame)), and we believe that such an approach is preferable.

(5) Finally, the reader may wonder whether we need to reject the purity view, i.e., whether disgust could be selectively associated with purity norms while retaining the advantages that we attribute to it. We think this is unlikely. In addition to the fact that there are problems with the purity view itself (including substantial evidence that disgust occurs to non-purity violations) disgust appears to be linked to character, and there is no prima facie reason to think that character is more important for purity violations than other types of violation. Furthermore, Russell and Giner-Sorolla's results suggest that disgust is associated with character *as opposed to* purity norms

Moral disgust has much to recommend it as a tool of liberal change. Indeed, harkening back to the title, liberals' condemnations of moral disgust in themselves exhibit the features of moral disgust – exclusion, avoidance, expulsion, extermination, a rigid wholesale condemnation of the emotion in virtue of its "character" itself, the idea that it is somehow more crude or coarse (and connected to the sticky body), a cleansing of moral psychology to counteract our tendencies towards it, first-order punishment of those who promote its merits, and second-order punishment of those who refuse to sufficiently distance themselves from its promoters. While we encourage their use of this mode of humans' moral psychology, we think such sentiments are misdirected in this case, and that it would be unfortunate indeed were liberal scholars to conclude that it is time we wash our hands of moral disgust.

## References

- Alderman, J., Dollar, K. M., & Kozlowski, L. T. (2010). Commentary: Understanding the origins of anger, contempt, and disgust in public health policy disputes: Applying moral psychology to harm reduction debates. *Journal of public health policy, 31*(1), 1-16.
- Baron, J., & Ritov, I. (2009). Protected values and omission bias as deontological judgments. *Psychology of learning and motivation, 50*, 133-167.
- Baron-Cohen, S. (2012). *The science of evil: On empathy and the origins of cruelty*. Basic books.
- Barrett, H.C. (2005). "Adaptations to predators and prey", in D. Buss (ed.) *The handbook of evolutionary psychology* (pp. 200-223), John Wiley and Sons, Inc.
- Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition, 121*(1), 154-161.
- Bastian, B., & Haslam, N. (2011). Experiencing dehumanization: Cognitive and emotional effects of everyday dehumanization. *Basic and Applied Social Psychology, 33*(4), 295-303.
- Bastian, B., & Haslam, N. (2010). Excluded from humanity: The dehumanizing effects of social ostracism. *Journal of Experimental Social Psychology, 46*(1), 107-113.
- Bloom, P. (2013). The baby in the well: the case against empathy. *The New Yorker, 20*.
- Bloom, P. (2013). *Just Babies: The Origins of Good and Evil*. Random House LLC.
- Borg, C., & de Jong, P. J. (2012). Feelings of disgust and disgust-induced avoidance weaken following induced sexual arousal in women. *PLoS one, 7*(9), e44111.
- Borg, J. S., Lieberman, D., & Kiehl, K. A. (2008). Infection, incest, and iniquity: Investigating the neural correlates of disgust and morality. *Journal of Cognitive Neuroscience, 20*(9), 1529-1546.
- Buckels, E. E., & Trapnell, P. D. (2013). Disgust facilitates out-group dehumanization. *Group Processes & Intergroup Relations 16*(6):771-780.
- Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science, 323*(5918), 1222-1226.
- Chapman, H. A., & Anderson, A. K. (2013). Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological bulletin, 139*(2), 300.

Chua, A. (2011). *Battle hymn of the tiger mother*. Bloomsbury Publishing.

Ciaramelli, E., Braghittoni, D., di Pellegrino, G. (2012). It is the outcome that counts! Damage to the ventromedial prefrontal cortex disrupts the integration of outcome and belief information for moral judgment. *Journal of the International Neuropsychological Society*, 18, 962–71.

Clark, J.A. (forthcoming). "Digging disgust out of the dumpster: A neuropsychological defense of moral disgust as an adaptive emotion", Invited chapter for *The Revolting Self: Perspectives on the Psychological and Clinical Implications of Self-Directed Disgust*, eds. Overton, P., Powell, P. & J. Simpson

Clark, J.A. & Fessler, D.M. T., (in prep.) "The messy evolution of disgust".

Curtis, V., de Barra, M., & Aunger, R. (2011). Disgust as an adaptive system for disease avoidance behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1563), 389-401.

Damasio, H. (2005). Disorders of social conduct following damage to prefrontal cortices. In *Neurobiology of human values* (pp. 37-46). Springer Berlin Heidelberg.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264(5162), 1102-1105.

Damasio, A. (1994). Descartes's Error: Emotion. *Reason and the Human Brain*, Grosset and Putnam, New York.

Danovitch, J., & Bloom, P. (2009). Children's extension of disgust to physical and moral events. *Emotion*, 9(1), 107.

De Block, A., & Cuypers, S. E. (2012). Why Darwinians Should Not Be Afraid of Mary Douglas—And Vice Versa The Case of Disgust. *Philosophy of the Social Sciences*, 42(4), 459-488.

Denburg, N. L., Cole, C. A., Hernandez, M., Yamada, T. H., Tranel, D., Bechara, A., & Wallace, R. B. (2007). The Orbitofrontal Cortex, Real-World Decision Making, and Normal Aging. *Annals of the New York Academy of Sciences*, 1121(1), 480-498.

Deonna, J. A., Rodogno, R., & Teroni, F. (2011). *In defense of shame: The faces of an emotion*. Oxford University Press.

Ehrlich, P. R., & Ornstein, R. E. (2010). *Humanity on a tightrope: thoughts on empathy, family, and big changes for a viable future*. Rowman & Littlefield Publishers.

Ember, M., Ember, C. R., & Low, B. S. (2007). Comparing explanations of polygyny. *Cross-Cultural Research*, 41(4), 428-440.

Eng, S. J. (2008). *Experimental Exposure-willingness (EEW): Testing a Behavioral Measure of Attraction Vs. Aversion to Disgust Stimuli* (Master's thesis, University of California, Los Angeles).

Fessler, D. M., Pillsworth, E. G., & Flamson, T. J. (2004). Angry men and disgusted women: An evolutionary approach to the influence of emotions on risk taking. *Organizational Behavior and Human Decision Processes*, 95(1), 107-123.

Galef, B.G., McQuoid, L. & Whiskin, E. (1990) Further evidence that Norway rats do not socially transmit learned aversions to toxic baits. *Learning & Behavior* 18(2): 199-205. doi: 10.3758/BF03205259

Giner-Sorolla, R. & Russell, P.S. (in prep.). " Not just disgust: Fear and anger attitudes also predict intergroup dehumanization"

Giner-Sorolla, R., & Espinosa, P. (2011). Social cuing of guilt by anger and of shame by disgust. *Psychological Science*, 22(1), 49-53.

Goldberg, J.H., Lerner, J.S. and Tetlock, P.E. (1999) 'Rage and reason: the psychology of the intuitive prosecutor', *European Journal of Social Psychology* 29, 781–795

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5), 1029.

Greene, J. D., Cushman, F. A., Stewart, L. E., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111(3), 364–371.

Harris, L. T., & Fiske, S. T. (2011). Dehumanized perception: A psychological means to facilitate atrocities, torture, and genocide?. *Zeitschrift für Psychologie/Journal of Psychology*, 219(3), 175.

Harris, L. T., & Fiske, S. T. (2009). Social neuroscience evidence for dehumanised perception. *European review of social psychology*, 20(1), 192-231.

Harris, L. T., & Fiske, S. T. (2007). Social groups that elicit disgust are differentially processed in mPFC. *Social cognitive and affective neuroscience*, 2(1), 45-51.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low neuroimaging responses to extreme out-groups. *Psychological Science*, 17(10), 847-853.

Helzer, E. G., & Pizarro, D. A. (2011). Dirty liberals! Reminders of physical cleanliness influence moral and political attitudes. *Psychological science*, 22(4), 517-522.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., ... & Ziker, J. (2010). Markets, religion, community size, and the evolution of fairness and punishment. *science*, 327(5972), 1480-1484.

Herz RS. 2003. The effect of verbal context on olfactory perception. *J Exp Psychol Gen.* 132(4):595–606.

Holbrook, C., Piazza, J., Fessler, & Daniel M. T. (in press[a]) Conceptual and Empirical Challenges to the “Authentic” Versus “Hubristic” Model of Pride. *Emotion*.

Holbrook, C., Piazza, J., Fessler, & Daniel M. T. (in press[b]). "Further Challenges to the Authentic”/“Hubristic” Model of Pride: Conceptual Clarifications and New Evidence". *Emotion*

Inbar, Y., Pizarro, D. A., & Bloom, P. (2012). Disgusting smells cause decreased liking of gay men. *Emotion*, 12(1), 23.

Inbar, Y., Pizarro, D. A., & Bloom, P. (2009). Conservatives are more easily disgusted than liberals. *Cognition and Emotion*, 23(4), 714-725.

Kahan, D. (1998). "The anatomy of disgust in criminal law", *Michigan Law Review*, vol.96:1621.

Kass, L. R. (1997). Wisdom of repugnance: Why we should ban the cloning of humans, the. *Val. UL Rev.*, 32, 679.

Kelly, D. R. (2011). *Yuck!: the nature and moral significance of disgust*. The MIT Press.

Krendl, A. C., Moran, J. M., & Ambady, N. (2013). Does context matter in evaluations of stigmatized individuals? An fMRI study. *Social Cognitive and Affective Neuroscience*, 8(5), 602-608.

Mencimer, S. (2010). “Rick Santorum’s anal sex problem”, *Mother Jones*, September/October issue. "Excerpt from Santorum interview". *USA Today*. April 23, 2003. ISSN 0734-7456.

Navarrete, C. D., & Fessler, D. M. (2006). Disease avoidance and ethnocentrism: The effects of disease vulnerability and disgust sensitivity on intergroup attitudes. *Evolution and Human Behavior*, 27(4), 270-282.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science*,314(5805), 1560-1563.

Nussbaum, M. C. (2009). *Hiding from humanity: Disgust, shame, and the law*. Princeton University Press.

Nussbaum, M. C. (2010). *From disgust to humanity: Sexual orientation and constitutional law*. Oxford University Press.

- Olatunji, B. O., Adams, T., Ciesielski, B., David, B., Sarawgi, S., & Broman-Fulks, J. (2012). The Three Domains of Disgust Scale Factor Structure, Psychometric Properties, and Conceptual Limitations. *Assessment, 19*(2), 205-225.
- Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. *The social psychology of morality: Exploring the causes of good and evil, 91-108.*
- Pond Jr, R. S., DeWall, C. N., Lambert, N. M., Deckman, T., Bonser, I. M., & Fincham, F. D. (2012). Repulsed by violence: Disgust sensitivity buffers trait, behavioral, and daily aggression. *Journal of personality and social psychology, 102*(1), 175.
- Reilly, S. & Schachtman, T.R. (2008) *Conditioned taste aversion: neural and behavioral processes*: Oxford University Press, USA.
- Rifkin, J. (2009). *The empathic civilization: The race to global consciousness in a world in crisis*. Penguin.
- Rozin, P. et al. (2008). Disgust. In M. Lewis, J. Haviland (Eds.), *Handbook of Emotions* (3rd edn), Guilford (2008), pp. 757–776
- Rozin, P., Lowery, L., Imada, S., & Haidt, J. (1999). The CAD triad hypothesis: a mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of personality and social psychology, 76*(4), 574.
- Rozin, P., Haidt, J., & McCauley, C. R. (1993). Disgust. In M. Lewis & J. Haviland (Eds.), *Handbook of emotions* (pp. 575–594). New York: Guilford Press.
- Russell, P. S., & Giner-Sorolla, R. (2011a). Social justifications for moral emotions: When reasons for disgust are less elaborated than for anger. *Emotion, 11*(3), 637.
- Russell, P. S., & Giner-Sorolla, R. (2011b). Moral anger, but not moral disgust, responds to intentionality. *Emotion, 11*(2), 233.
- Russell, P. S., Piazza, J., & Giner-Sorolla, R. (2013). CAD Revisited Effects of the Word Moral on the Moral Relevance of Disgust (and Other Emotions). *Social Psychological and Personality Science, 4*(1), 62-68.
- Russell, P. S., & Giner-Sorolla, R. (2013). Bodily moral disgust: What it is, how it is different from anger, and why it is an unreasoned emotion. *Psychological bulletin, 139*(2), 328.
- Sherman, G., Haidt, J., & Coan, J. (2007). Nazis really are disgusting. *Unpublished manuscript, University of Virginia.*

Stürmer, S., Snyder, M., Kropp, A., & Siem, B. (2006). Empathy-motivated helping: The moderating role of group membership. *Personality and Social Psychology Bulletin*, 32(7), 943-956.

Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, 28, 531-542

Tangney, J. P., & Dearing, R. L. (2003). *Shame and guilt*. Guilford Press.

Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual review of psychology*, 58, 345.

Tracy, J. L., & Robins, R. W. (2007). The psychological structure of pride: a tale of two facets. *Journal of personality and social psychology*, 92(3), 506.

Tracy, J. L., Cheng, J. T., Robins, R. W., & Trzesniewski, K. H. (2009). Authentic and hubristic pride: The affective core of self-esteem and narcissism. *Self and Identity*, 8(2-3), 196-213.

Tybur, J. M., Lieberman, D., Kurzban, R., & Descioli, P. (2013). Disgust: Evolved function and structure. *Psychological review*, 120(1), 65-84.

Vartanian, L. R. (2010). Disgust and perceived control in attitudes toward obese people. *International Journal of Obesity*, 34(8), 1302-1307.