

A HUMEAN APPROACH TO THE BOUNDARIES OF MORALITY

MARK COLLIER

University of Minnesota, Morris

ABSTRACT

Hume maintains that the boundaries of morality are widely drawn in everyday life. We routinely blame characters for traits that we find disgusting, on this account, as well as those which we perceive as being harmful. Contemporary moral psychology provides further evidence that human beings have a natural tendency to moralize traits that produce feelings of repugnance. But recent work also demonstrates a significant amount of individual variation in our sensitivities to disgust. We have sufficient reason to bracket this emotion, therefore, when adopting the general point of view: if we allow idiosyncratic affective responses to shape our fully considered moral judgments, we could no longer reasonably expect spectators with different sensitivities to agree with us.

Keywords: Hume, sentimentalism, moral disagreement, moral domain, disgust

The act is to the highest degree odious and disgusting, that is, not to the man who does it, for he does it only because it gives him pleasure, but to one who thinks of it. Be it so, but what is that to him? He has the same reason for doing it that I have for avoiding it. A man loves carrion – this is very extraordinary – much good may it do him. But what is this to me so long as I can indulge myself with fresh meat? But such reasoning, however just, few persons have calmness to attend to. This propensity is much stronger than it is to be wished it were to confound physical impurity with moral. (Bentham 1785/1978b: 94)

The Journal of Scottish Philosophy 18.1 (2020): 1–16

DOI: 10.3366/jsp.2020.0252

© Edinburgh University Press

www.eupublishing.com/jsp

I. HUME'S BROAD ACCOUNT OF THE MORAL DOMAIN

Hume maintains that the boundaries of morality are widely drawn in everyday life. We ordinarily blame characters for traits that we regard as harmful, but we also routinely condemn personal features that we perceive to be disagreeable.¹ Consider the vice of *meanness*. Agents who debase themselves in pursuit of their ends, or who flatter and court favor with those who mistreat them, are typically regarded as 'disgustful and contemptible' (EPM 7.10n, SBN 253).² This characteristic is seen as a 'real deformity', as Hume puts it, which 'displeases' in the same manner as 'the want of a nose' (EPM 7.10, SBN 253). Similarly, *uncleanliness* excites an 'uneasy sensation' in onlookers (T 3.3.4.10, SBN 611; compare EPM 8.13, SBN 266–7), *melancholy* 'disfigures' characters by 'giving uneasiness' and producing 'aversion and disgust' (EPM 7.2, SBN 251), *dissoluteness* is 'contemptible' and 'odious' (EPM App 4.3–4, SBN 314–315; compare EPM App 4.8, SBN 317–8), and *folly* is 'disgustful' (EPM 7.22n, SBN 258).

Human beings are apt to experience negative moral emotions such as contempt and disgust toward characteristics that we perceive as low and beneath us, but we are also disposed to feel positive emotions such as awe and esteem in response to qualities which raise us up 'above the rest of mankind' (T 3.3.5.14, SBN 613). There are moral highs, in other words, as well as moral lows. Consider the type of moral approbation that 'shining virtues' such as of *magnanimity* or *greatness of mind* often receive in common life (T 3.3.2.13, SBN 600). When we praise the irresolute *courage* of Medea as she singlehandedly battles Creon's armies, or esteem the vaulting *ambition* of Alexander the Great as he sets out to conquer the world, our moral approval is based on 'sublime' feelings of moral elevation (EPM 7.6–7, SBN 252–3; compare T 3.3.2.14, SBN 600; compare T 3.3.2.14, SBN 600; compare T 2.3.8.4–5, SBN 433–434).³

The natural disposition to moralize sublime traits, such as *tranquility* in the face of great adversity, is not founded on any 'views of utility' or 'future beneficial consequences' (EPM 7.29, SBN 260). We commonly feel approbation and esteem toward those who display *cheerfulness*, moreover, even if we believe this agreeable quality to be 'without any utility, or any tendency to farther good, either of the community or of the possessor' (EPM 7.2, SBN 250–1). Indeed, we ordinarily approve of elevating qualities such as *excessive bravery* and *extravagant ambition* even though we explicitly acknowledge these traits to be 'dangerous' (T 3.3.2.15, SBN 601; compare EPM 7.24, SBN 258).

Moral judgments in everyday life are often based, according to Hume, on the intrinsic characteristics of agents (T 3.3.1.28, SBN 590; compare EPM 7.12, SBN 254). When we admire the vaulting *ambition* of military leaders such as Philip II of Macedon, for example, our thoughts 'carry us not ... beyond the hero himself, nor ever regard the future consequences of his valour' (EPM 7.12,

SBN 254). Similarly, our feelings of approbation toward agreeable qualities such as *wit* and *eloquence* are based on immediate sensations that occur ‘without reflection’ on the ‘tendencies of qualities and characters’ (T 3.3.4.11, SBN 612). Moral evaluation in these cases is grounded on the ‘mere . . . appearance’ of these personal qualities, as Hume puts it, rather than any ‘reflection on their tendency to the happiness of mankind’ (T 3.3.1.27, SBN 590).

Moral judgments are analogous, in this respect, to aesthetic evaluations. It is a ‘universal rule’ of taste and criticism, according to Hume, that the appeal of works of art is ‘chiefly deriv’d from their utility’ (T 2.2.5.17, SBN 364; compare T 2.1.8.2, SBN 299). The beauty of a particular house is principally based, for example, on its utility and convenience (T 3.3.1.8, SBN 576–577; EPM 6.24, SBN 244; compare EPM 2.9, SBN 179). But aesthetic appreciation of objects and artifacts is sometimes based on nothing more than our immediate affective reactions to their formal characteristics. We are pleased with the regularity of our bodily features, for example, even though these qualities are ‘neither useful to ourselves, nor others’ (T 3.3.5.4, SBN 615).

The widespread propensity in common life to evaluate character traits independently of harm and welfare considerations, according to Hume, provides crucial support for the doctrine of moral sentimentalism. Moral rationalists would have a difficult time, after all, accommodating the fact that we condemn traits such as *uncleanliness* simply because they are apt to make us feel uneasy or disgusted. Hume explains, ‘[W]e may, in this instance, seemingly so trivial, clearly discover the origin of moral distinctions, about which the learned have involved themselves in such mazes of perplexity and error’ (EPM 8.13, SBN 266–267; compare T 3.3.4.10, SBN 611). Samuel Clarke maintains that we discern moral truths in much the same way that we grasp mathematical axioms and theorems. But this makes it hard to understand our natural tendency to moralize *uncleanliness*: there are presumably no eternal or immutable relations of fitness through which one can demonstrate *a priori* the viciousness of going unwashed.⁴

2. HUME’S BROAD ACCOUNT: A CONTEMPORARY EVALUATION

Hume recognizes that the extent of the moral domain in common life is an empirical matter that requires one to adopt an ‘experimental’ approach (EPM 1.10, SBN 174). But there is a significant worry, however, about Hume’s methodology. Hume’s broad account of the boundaries of morality is based on informal studies of ancient Greek and Roman panegyrics, satires, and funeral orations (T 3.3.1.24, SBN 587). We can discover which personal characteristics are moralized in everyday life, Hume maintains, by examining what is said about people when they die (EPM 2.2, SBN 176–7). But even if one accepts that the classical encomia of writers such as Polybius and Livy accurately capture the local

attitudes of ancient Rome and Greece, this is hardly a global survey: we cannot assume that these particular cultures are representative of the moral commitments of diverse populations around the world.

A similar objection can be raised, it should be noted, against several of Hume's critics. Hume's contemporaries unanimously rejected his broad account of the moral domain on the grounds that it is counter-intuitive (Feiser 1998: 298–303). And recent commentators offer a similar critique. John Mackie is highly suspicious, for example, of Hume's wide delineation of morality: we might *admire* agents who display agreeable traits such as eloquence, he maintains, but we surely do not *blame* those who are deficient in these personal qualities (Mackie 1980: 127–8; compare Foot 1978: 76). But how can these critics be certain that their intuitions are widely shared in their local communities, let alone be confident that these attitudes are cross-culturally valid?

In order to properly settle disputes about the boundaries of morality in common life, therefore, we must carefully examine the evaluations of a range of different societies. Consider the type of ethnographic fieldwork conducted by Richard Shweder and his colleagues (1997), for example, which carefully plumbs the moral intuitions of populations in a culture very distant from ours. They present residents of a Brahman temple town in Bhubaneswar, Orissa, India with a series of thirty-nine vignettes that describe violations of normative codes of conduct. In one incident, a son cuts his hair and eats chicken soon after his father's death; in another, a widow eats fish several times a week. Structured interviews are used to capture their attitudes toward these breaches as well as the justifications that are offered in support of these evaluations; cluster-analysis is subsequently employed to extract the central themes of their moral discourse (Shweder et al. 1997: 138).

Shweder and his colleagues discover that praise and blame in Orissa typically revolves around three main foundations: individual autonomy, communal hierarchy, and spiritual purity (Shweder et al. 1997: 140). Moral disapproval of the son's behavior is saturated, for example, with reference to the importance of social roles and community hierarchy; condemnation of the widow's behavior, moreover, often invokes spiritual purity and the sacred order of things. Thus, it appears that residents in this town have a more expansive account of the moral domain than their counterparts in the United States, who tend to restrict their moral approbation to matters of individual autonomy, rights, and harm (Shweder et al. 1997: 142).

One might object that this cross-cultural variation can be accommodated by a narrow account of the moral domain. Populations in Orissa and the United States display conflicting attitudes about which actions are morally permissible, but these moral disagreements might be explained away in terms of different factual beliefs about the welfare consequences at stake. Consider once again the vignette in which a son eats chicken and cuts his hair shortly after the death of his father. Orissa residents typically regard this act as highly immoral, but this condemnation

must be understood against the background of their theoretical commitments. It is widely accepted in this community that spiritual pollution from the dead can be absorbed into the bodies of the living; sons are forbidden to eat meat or cut their hair in the period following the death of their fathers, therefore, because such behavior is perceived to be harmful (Turiel et al. 1987: 207). The same explanatory strategy can be applied to the case of the widow who is blamed for eating fish. It is a common belief in the Odissa community that eating fish causes women to act promiscuously; the wife's actions could be viewed, then, as reducing utility (Haidt et al. 1993: 615).

Jonathan Haidt and his fellow researchers manage to navigate around this type of objection by examining cross-cultural attitudes about *harmless* offenses. While conducting structured interviews with diverse socio-economic groups in the United States and Brazil, these researchers ask participants to morally evaluate vignettes such as FLAG, in which someone cuts up their national flag in order to clean a bathroom, and CHICKEN, where an individual has sexual intercourse with a chicken before eating it. They employ statistical analyses on this data to determine the extent to which those who believe these normative transgressions to be harmless nevertheless adopt a critical moralizing stance (Haidt et al. 1993: 613).

The results of these studies are striking: low socio-economic status (SES) populations in less Westernized cities in Brazil are significantly more likely than high-SES populations in the United States to judge that harmless offenses fall within the boundaries of the moral domain; a majority (63%) of working class Brazilians judge the behavior described in the FLAG vignette to be morally impermissible, even when regarded as victimless, whereas none of the upper class North Americans in the study make this type of evaluation (Haidt et al. 1993: 619). Low-SES populations in Brazil are especially prone to moralize actions that violate purity norms or social hierarchies; this population operates with a 'broader' morality, therefore, than high-SES groups in the United States (Haidt et al. 1993: 625).⁵

What role do emotions play in this moral disapproval of harmless offenses? Paul Rozin and his colleagues attempt to answer this question by presenting populations in the United States and Japan with vignettes similar to those used in the Shweder and Haidt studies and asking participants to choose the facial expression or emotion word that best captures the feelings of hypothetical spectators who witness these transgressions (Rozin et al. 1999: 576). They discover a one-to-one mapping between the three foundations of Orissa moral discourse and a triad of other-critical emotions (Rozin et al. 1999: 585). According to their 'CAD' Hypothesis, violations of communal hierarchies elicit contempt, violations of autonomy provoke anger, and violations of divinity or purity excite disgust (Rozin et al. 1999: 578). Normative transgressions fall within the boundaries of morality as long as they are apt to produce at least one of the

negative emotions in this ‘disapprobation spectrum’ (Prinz 2007: 90; compare Gill and Nichols 2008: 148).

The proposal that negative emotions such as contempt and disgust are sufficient for moralization in everyday life receives independent support from controlled laboratory experiments. Participants in these studies are primed to experience flashes of repugnance in response to arbitrary words, such as ‘take’ or ‘often’, that are randomly embedded in texts describing moral transgressions. Inducing disgust in this manner tends to increase the severity of moral judgments: participants are more likely to evaluate vignettes that include these trigger words as seriously morally wrong (Wheatley and Haidt 2005: 781).

This effect has been replicated across a number of experimental conditions. It does not matter whether disgust is induced through ammonium sulfide (‘fart spray’), filthy desks covered with garbage, or the notorious scene from *Trainspotting* where Ewan MacGregor dives into the ‘Worst Toilet in Scotland’. In each of these contexts, participants who are primed to feel pangs of disgust make harsher moral judgments than those in control conditions (Schnall et al. 2008a: 1102). Moreover, disgust does not only amplify the severity of moral condemnation; it also causes participants to moralize otherwise neutral vignettes. Those who are prompted to experience repugnance are more likely to assert, for example, that a student council representative who chooses topics for discussion at campus assembly has done something morally wrong (Wheatley and Haidt 2005: 782–3).

Going in the other direction, inducing feelings of cleanliness reduces the harshness of moral judgments. Participants who wash their hands with soap after watching the cringe worthy clip from *Trainspotting*, for example, typically make less severe judgments than those who do not clean themselves (Schnall et al. 2008b: 1220–21). Similar effects can be found in studies of self-directed moral condemnation. Participants who clean themselves with antiseptic wipes after recalling unethical deeds, for example, are much more likely to exhibit diminished levels of shame and guilt. It appears that Shakespeare was picking up on a general propensity of human nature, therefore, in his depiction of Lady Macbeth’s desire to scrub her hands after killing Duncan (‘Out, damned spot!’); we do commonly believe, it appears, that we can wash away our sins (Zhong and Liljenguist 2006: 1452).

3. IS THIS A VERBAL DISPUTE?

Frances Hutcheson attempts to draw a categorical distinction between merely agreeable ‘natural Abilities’ such as *magnanimity* and ‘moral Qualitys’ like *benevolence* (Hutcheson 2008: 128). But Hume argues that ordinary language is too vague to permit any clear lines of demarcation. ‘Whether natural Abilities be Virtues’, he writes to Hutcheson in a letter dated September 17, 1739, ‘I follow

the common Use of Language' (Greig 1969: 33–4). And common usage does not enable one to 'mark the precise boundaries' between talents and virtues, or defects and vices; these moral categories are not 'exactly fixed' or 'precisely define[d]' in English or 'any other modern tongue' (EPM App 4.2; SBM 313).

We cannot clearly demarcate natural abilities and virtues, according to Hume, because there is 'so little distinction' between our standard emotional responses to these qualities (EPM App 4.3, SBN 314). Sublime and elevating qualities such as *magnanimity*, for example, provoke sentiments 'of a kind similar' to those which arise from socially useful traits like *benevolence* (EPM 7.29, SBN 260; compare T 3.3.6.6, SBN 617). Good qualities elicit love; great ones excite esteem. But these positive emotions have similar causes and effects and thus are 'at bottom the same passions' (T 3.3.4.2n, SBN 608; compare T 2.2.2.10, SBN 337; compare T 3.3.4.1, SBN 607). Negative emotions such as contempt and hatred also form a natural resemblance class. According to Hume's taxonomy, contempt is a 'species' of hatred (T 2.2.5.1, SBN 357). Contempt is what one feels when the basic emotion of hatred is mixed together or blended with feelings of personal superiority (T 2.2.10.2, SBN 390).⁶

The similarity of our emotional responses to natural abilities and social virtues renders it artificial to partition these qualities into distinct kinds. One can choose to label these traits as talents or virtues. But this is a verbal distinction without a substantive difference: no matter how we talk, we would feel much the same way.

[I]f, in short, the sentiments are similar, which arise from the endowments and from the social virtues, is there any reason for being so extremely scrupulous about a word, or disputing whether they be entitled to the denomination of virtues? (EPM App 4.6, SBN 316)

Everyone acknowledges that disagreeable traits such as *uncleanliness* make them feel uneasy and that they dread having such epithets hurled in their direction. The philosophical debate over the extent of the moral domain, then, is relegated to a linguistic dispute of interest only to grammarians (T 3.3.4.1, SBN 606; compare EPM App 4.1, SBN 312–3; EPM App 4.22, SBN 322).

Those in favor of restricting the moral domain would likely object, however, that Hume is blurring a crucial distinction. We might have similar emotional reactions to harmful and disagreeable traits, but there are important differences in our behavior toward them. We are *prima facie* disposed to sanction and discipline agents for harming others, but we do not adopt a punitive stance toward merely disagreeable qualities such as *uncleanliness* or *folly*. This asymmetry serves as a plausible criterion, therefore, for excluding harmless but disagreeable traits from the moral domain (Feiser 1998: 308).

Restrictionists about the scope of the moral domain are once again attempting to draw clear lines, however, where there are really only continuous gradations.

Hatred is naturally attended with anger and a desire for the misery of its target (T 3.3.1.31, SBN 591). But contempt and disgust underwrite similar types of sanctions: they reduce good-will and friendship (EPM App 4.8–10, SBN 317–8). Those who display disagreeable character traits typically meet with penalties such as withdrawal, censure, and outrage. We ‘commonly keep a distance’, as Hume puts it, from those we find contemptible (T 2.2.10.9, SBN 392). Social punishment is a matter of degree, therefore, rather than kind. At the upper extreme lies full criminalization and restriction of liberties; at the lower end, stigmatization and exclusion.⁷

Hume concedes that benevolence and cruelty are paradigmatic virtues and vices that ‘bear the greatest figure’ in our lives (T 3.3.1.11, SBN 578; compare T 3.3.1.27, SBN 590). But the centrality of these traits does not preclude the existence of ‘other kinds’ of qualities at the periphery of our moral space (T 3.3.1.25, SBN 588; compare EPM App 4.2, SBN 313). A harmful trait such as *cruelty* is the ‘most detested of all vices’ and produces the ‘strongest hatred’ (T 3.3.3.8, SBN 605; compare EPM App 4.2, SBN 313). Merely disagreeable traits like *uncleanliness* produce affective responses with ‘*inferior*’ degrees of force and vivacity (T 3.3.4.2, SBN 607). But this does not mean that *uncleanliness* is morally neutral; it is only less wrong and generates weaker prohibitions.

4. THE MORAL DOMAIN AND THE GENERAL POINT OF VIEW

Hume challenges his restrictionist opponents to find a ‘sufficient reason for excluding’ natural abilities such as *courage* or *magnanimity* from the ‘catalogue of virtues’ (T 3.3.4.2, SBN 607). One cannot appeal to the fact that these agreeable qualities produce ‘somewhat different’ types of emotional responses in spectators. This criterion is obviously too strong: it rules out canonical social virtues, such as *justice* and *benevolence*, which also excite a ‘different sentiment or feeling in the spectator’ (T 3.3.4.2, SBN 607).

The only alternative justification for restricting the boundaries of morality, according to Hume, involves considerations of free will. Qualities would be shielded from praise and blame, on this proposal, if they lie beyond our volitional control (T 3.3.4.4, SBN 608). But the problem is that this intuition is not universally shared. It is apparent from ‘common life and conversation’, as Hume puts it, that laypeople ‘do not seem much to regard this distinction’ (T 3.3.4.4, SBN 609; compare EPM App 4.2, SBN 313).⁸

[S]entiments are every day experienced, of blame and praise, which have objects beyond the dominion of the will or choice... (EPM App 4.21, SBN 322)

Ancient philosophers, moreover, share these attitudes. They classify *magnanimity* and *courage* as virtues, and *meanness* and *folly* as vices, even though these personal qualities are believed to be entirely ‘independent of the will’ (EPM App 4.20, SBN 321–2; compare T 3.3.4.3, SBN 608). Hume regards Cicero as a model case: he includes agreeable but involuntary qualities like *magnanimity* and *decency* within the catalogue of virtues (EPM App 4.11n, SBN 316–317).⁹

Hume takes himself to have eliminated every possible justification for excluding harmless but disagreeable qualities from the moral domain. But there is an excluded alternative. This rationale is grounded in Hume’s official commitments, moreover, concerning the general point of view: we have sufficient reason to restrict our fully considered judgments to considerations of harm and welfare because this is a necessary condition of the possibility of moral agreement.

Hume recognizes that everyday moral judgments are vulnerable to systematic biases and distortions. Consider the *proximity bias*: it is difficult for us to sympathize with the affections of those who are remote from us, and as a result, we are prone to mistakenly assert that distant characters are less virtuous, even when they display exactly the same traits as those close at hand (T 3.3.1.14, SBN 580–1). We do not reflectively endorse this natural propensity, however, since irresolvable moral disagreements would be inevitable if we allow ourselves to judge characters from this variable standpoint (Cohon 2008: 127). We must adopt the general point of view, therefore, if we are to have any hope of achieving a stable moral consensus (T 3.3.1.15–16, SBN 581–2; compare T 3.3.1.30, SBN 591; compare T 3.3.3.2, SBN 602–3).

When we express moral praise or blame, according to Hume, we ‘expect’ our audience to ‘concur’ with us (EPM 9.6, SBN 272). And we can only reasonably anticipate such concurrence when we appeal to the welfare consequences of traits; utility considerations are the only concerns that ‘touch a string’, as Hume puts it, ‘to which all mankind have an accord or symphony’ (EPM 9.6, SBN 272). But what about merely disagreeable characteristics such as *uncleanliness*? Can we reasonably expect that other spectators will regard these traits as odious or disgusting?

Hume takes it for granted that emotional aversions to disagreeable qualities are sufficiently uniform across populations (EPM 9.12, SBN 277). But this crucial assumption is undercut by recent work in moral psychology. Researchers have discovered a significant amount of individual differences in terms of our sensitivities to disgust. Disgust sensitivity varies across populations along a continuous scale, with high-disgust sensitivity at one of the spectrum, and low-disgust sensitivity at the other. Some personality types have a predisposition to experience repugnance more frequently and more intensely than others; one and the same property, therefore, might elicit feelings of abhorrence in some spectators but not in others (Haidt et al. 1994: 709).

Moral psychologists have found that variation in disgust sensitivity does in fact generate moral disagreement. Consider the experimental work of Shaun Nichols, for example, on the moralizing role that disgust plays in everyday life. Nichols conducts surveys in which participants are presented with a series of vignettes describing etiquette violations that occur during a dinner party. He discovers that transgressions that trigger disgust responses, such as snorting and spitting into a glass of water before drinking it, are the only norm violations that are judged to be immoral (Nichols 2004: 21–22). Nichols also finds that participants in these studies disagree about the moral status of these transgressions. Their different evaluations are correlated, moreover, with their idiosyncratic levels of disgust sensitivity: a majority of high-disgust sensitive individuals regard this type of infraction as a moral transgression, whereas a majority of low-disgust sensitive participants judge it as merely a conventional violation (Nichols 2004: 23–24).

The connection between disgust sensitivity and moral disagreement has been corroborated across a variety of empirical studies: In the experiments conducted by Wheatley and Haidt, for example, a majority of those who are induced to feel disgust judge sexual relations between second cousins to be morally wrong, whereas a majority of those in the control group regard this type of behavior as morally neutral (Wheatley and Haidt 2005: 781). Chapman and Anderson make a similar discovery: high-disgust sensitive judges are twice as likely as low-disgust sensitive judges to moralize violations of norms such as talking in class without raising your hand (Chapman and Anderson: 2014: 344).

Jones and Fitness get analogous results in their work on moral hypervigilance: a majority of high-disgust sensitive individuals reach a guilty verdict after reading a transcript of a murder trial, whereas a majority of low-disgust sensitive individuals find the defendant to be innocent (Jones and Fitness 2018: 619). High-disgust sensitive judges also tend to recommend significantly harsher and lengthier punishments after reading vignettes about crimes (Jones and Fitness 2018: 622–623).

Individual variation in disgust sensitivity has been linked, moreover, to divergent attitudes about the morality of homosexuality. Yoel Inbar and his colleagues discover that disgust sensitivity predicts moral disapproval of gay people (Inbar et al. 2009a: 438; compare Terrizi et al. 2010: 590). High-disgust sensitive judges are also more likely than low-disgust sensitive judges to be strongly opposed to gay marriage (Inbar et al. 2009b: 720).¹⁰

The cross-cultural studies examined in Part 2 of this paper provide additional support for the claim that spectators cannot expect substantive moral agreement when they base their judgments on feelings of contempt and disgust. One of the main conclusions of the Shweder research, after all, is that the moral attitudes of populations in the Chicago and Orissa diverge on whether violations of hierarchical or purity norms fall within the boundaries of morality (Shweder et al. 1997: 142).¹¹ Haidt and his colleagues also locate substantive moral disagreement

in their comparative studies: one of their crucial findings is that high socio-economic status groups in the USA and low socio-economic status groups in Brazil exhibit conflicting attitudes about the morality of stories that are apt to produce 'disgust and disrespect' (Haidt et al. 1994: 625). Finally, Rozin and his colleagues discover 'considerable disagreement' between Japanese and American populations in their judgments about which normative violations would be met with facial expressions of contempt or disgust (Rozin et al. 1999: 581).

Crucially, moral judgments based on utility considerations do not exhibit the same patterns of individual or cross-cultural variation. Haidt and his fellow researchers discover significant moral disagreement about harmless but offensive vignettes, but in the one harm vignette presented to participants, in which one child pushes another off a swing, there was a nearly unanimous consensus across all socio-economic groups and geographical regions that this type of action is always wrong and should be stopped and punished (Haidt et al. 1993: 620). If we narrowly restrict the moral domain to considerations of harm and welfare, then, we can reasonably expect others to share our affections. Everyone feels hatred, for example, toward wanton *cruelty*. But if we allow idiosyncratic emotions such as disgust and contempt to drive our judgments, we would inevitably find ourselves in intractable moral disagreements. We could not expect consensus on mundane questions about what is appropriate at dinner parties, let alone controversial topics such as gay marriage or flag desecration.

We must bracket our emotional aversions and dislikes, then, if we are to have any reasonable hope of achieving a stable moral consensus. This is not to say that it is psychologically possible to suspend these affective responses, even in situations where we reflectively acknowledge a reason to do so. Hume recommends that we should 'correct' any 'variable' feelings of praise and blame that arise from fluctuating relations of self-interest or proximity, but he acknowledges that these biases and propensities often remain 'stubborn and inalterable' (T 3.3.1.16, SBN 582). Even when we are incapable of 'correcting our sentiments', however, we retain the ability of 'correcting our language' (T 3.3.1.16, SBN 582). We might not be able to regulate how we feel about particular traits, in other words, but we can at least control how we talk about them.

Recent work in social psychology provides further evidence for this proposal. Consider once again the research of Yoel Inbar and his colleagues on implicit attitudes toward homosexuality. They discover that participants in their studies often manage to overlook their affective reactions when making explicit moral judgments. Many participants in their studies harbor negative evaluations of gay kissing in public, for example, but they are able to 'consciously override' these feelings when they conflict with their fully considered egalitarian principles (Inbar et al. 2009a: 438). Similarly, in the cross-cultural studies conducted by Haidt and his colleagues, many high socio-economic status participants from the

United States report experiencing intense pangs of disgust when considering the CHICKEN incident, but they are able to ‘decouple’ these gut feelings from their avowed moral judgments (Haidt et al. 1993: 626).

5. CONCLUSION

Recent work in moral psychology provides a good deal of support for Hume’s descriptive claim about the role that negative emotions such as disgust play in moralizing harmless traits. But empirical research also undercuts Hume’s normative claim that there is no principled basis for excluding merely disagreeable qualities from the moral domain. There is significant variation between individuals, and across cultures, in terms of our sensitivities to disgust. As a result, one and the same action might strike high-disgust sensitive judges as morally impermissible but low-disgust sensitive judges as permissible. This provides a sufficient reason for restricting the boundaries of morality to harm and welfare considerations in our reflective moral judgments: if we allow our evaluations to be based on our contingent predilections, or what we happen to find odious or repugnant, we cannot reasonably expect fully informed judges with different emotional sensitivities to go along with us.

The rationale for bracketing our feelings of disgust when issuing moral judgments grows even stronger once we acknowledge the implicit nature of these immediate affective responses. Hume maintains that moral evaluations often occur automatically and ‘without reflection’; when queried about our justification for condemning immediately disagreeable qualities such as *uncleanliness*, we can only respond with an inchoate sense of ‘*je-ne-scai-quoi*’ (T 3.3.4.11, SBN 612). These personal qualities are apt to produce feelings of disapproval, as he puts it, ‘but how, or why, or for what reason, [we] cannot pretend to determine.’ (EPM 8.14, SBN 267).

Contemporary research provides additional evidence that our affective responses can operate beneath the level of conscious awareness. Consider once again the experimental work of Wheatley and Haidt on induced disgust. Participants in the disgust condition typically experience *aporetic* puzzlement or dumbfounding when asked why they morally condemn the Student Representative who merely raises question for discussion; one of them simply responds: ‘I don’t know [why it’s wrong], it just is.’ (Wheatley and Haidt 2005: 783).

The problem is not merely that one cannot provide reasons, then, that will convince those with different emotional sensitivities. When these moral evaluations are based on implicit attitudes, the problem is that one cannot give any reasons at all. This might explain why Hume restricts himself to welfare considerations when trying to change other minds about controversial issues of his day. When advocating for monogamy over polygamy, for example, Hume focuses

on the beneficial consequences of this type of marital arrangement (Miller 1987: 183–4). Disputes about the ‘boundaries of moral good and evil’ cannot be decided with ‘greater certainty’, he writes, than by putting our natural emotional responses to the side and carefully weighing evidence about the impact of traits and actions on ‘public utility’ (EPM 2.17, SBN 180).

Jeremy Bentham further develops the insight that utility is ‘uniquely suitable’ for moral and political debate in liberal societies (Darwall 1994: 75). If homosexuality is condemned on the grounds that it is ‘liable to disturb marriage’, for example, one might hope to settle the matter by appeal to the available evidence (Bentham 1978a: 404–405). But there are no factual standards for resolving debates about what is morally impure or offensive to one’s taste (Bentham 1978b: 94). Hume’s ethical theory can be regarded as a halfway house, therefore, in the development of philosophical utilitarianism. Hume maintains that ‘the foundations of all *virtue* are laid in *utility* . . . after a few exceptions made’; Bentham merely fails to understand ‘what need there was for the exceptions’ (Bentham 1962, 142n).

REFERENCES

- Ardal, Pall S. (1966) *Passion and Value in Hume’s Treatise*. Edinburgh: Edinburgh University Press.
- Bentham, Jeremy (1962) *A Fragment on Government, in The Works of Jeremy Bentham: Volume 1*, John Bowring (ed.), New York: Russell and Russell.
- [1785] (1978a) ‘Offenses Against One’s Self: Paederesty (Part 1)’, ed. Louis Compton, *Journal of Homosexuality* 3(4): 389–405.
- [1785] (1978b) ‘Jeremy Bentham’s Essay on “Paederesty” Part 2’, ed. Louis Compton, *Journal of Homosexuality* 4(1): 91–107.
- Chapman, H.A. and Anderson, A.K. (2014) ‘Trait physical disgust is related to moral judgments outside of the purity domain’, *Emotion* 14(2): 341–348.
- Cohon, Rachel (2008) *Hume’s Morality: Feeling and Fabrication*, Oxford: Oxford University Press.
- Darwall, Stephen (1994) ‘Hume and the invention and utilitarianism’, in M. A. Stewart and J. P. Wright (eds), *Hume and Hume’s Connexions*, Edinburgh: Edinburgh University Press, pp. 58–82.
- Fieser, James (1998) ‘Hume’s wide view of the virtues: An analysis of his early critics’, *Hume Studies* 24(2): 295–311.
- Foot, Philippa (1978) *Virtues and Vices*, Berkeley: University of California Press.
- Gill, Michael B. and Nichols, Shaun (2008) ‘Sentimentalist pluralism: moral psychology and philosophical ethics’, *Philosophical Issues* 18: 143–163.
- Greig, J.Y.T. (ed.) (1969) *The Letters of David Hume: Volume 1*, Oxford: Clarendon Press.
- Haidt, J., Koller, S.H. and Dias, M.G. (1993) ‘Affect, culture, and morality, or is it wrong to eat your dog?’, *Journal of Personality and Social Psychology* 65(4): 613–628.
- Haidt, J., McCauley, C. and Rozin, P. (1994) ‘Individual differences in sensitivity to disgust: a scale sampling seven domains of disgust elicitors’, *Personality and Individual Differences* 16: 701–713.
- Henrich, J., Heine, S. J. and Norenzayan, A. (2010) ‘The weirdest people in the world?’ *Behavioral and Brain Sciences* 33(2–3): 61–83.

- Hutcheson, Francis [1725] (2008) *An Inquiry into the Original of Our Ideas of Beauty and Virtue*, Indianapolis: Liberty Fund.
- Hume, David (1975) *Enquiries Concerning the Human Understanding and concerning the Principles of Morals*, ed. L.A. Selby-Bigge, 3d edn, rev. by P.H. Nidditch, Oxford: Clarendon Press.
- (1978) *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, rev. by P. H. Nidditch, 2d ed., Oxford: Clarendon Press.
- (1987) *David Hume: Essays Moral, Political, and Literary*, ed. E.F. Miller, Indianapolis: Liberty Fund.
- (1998) *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp, Oxford: Clarendon Press.
- (2000) *Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton, New York: Oxford University Press.
- Inbar, Y., Pizarro, D.A., Knobe, J. and Bloom, P. (2009a) 'Disgust sensitivity predicts intuitive disapproval of Gays', *Emotion* 9(3): 435–439.
- Inbar, Y., Pizarro, D.A. and Bloom, P. (2009b) 'Conservatives are more easily disgusted than liberals', *Cognition and Emotion* 23(4): 714–725.
- (2011) 'Disgusting smells cause decreased liking of Gay men', *Emotion* 12(1): 23–27.
- Jones, A. and Fitness, I. (2008) 'Moral hypervigilance: the influence of disgust sensitivity in the moral domain', *Emotion* 23: 719–737.
- Lai, C.K., Haidt, J. and Nosek, B.A. (2014) 'Moral elevation reduces prejudice against gay men', *Cognition and Emotion* 28(5): 781–794.
- Mackie, John L. (1980) *Hume's Moral Theory*, London: Routledge.
- Nichols, Shaun (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment*, New York: Oxford University Press.
- Prinz, Jesse J. (2007) *The Emotional Construction of Morals*, Oxford: Oxford University Press.
- Rozin, P., Lowery, L., Imada, S. and 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): 574–586.
- Schnall, S., Haidt, J., Clore, G.L. and Jordan, A.H. (2008a) 'Disgust as embodied moral judgment', *Personality and Social Psychology Bulletin* 34(8): 1096–1109.
- Schnall, S., Benton, J. and Harvey, S. (2008b) 'Cleanliness reduces the severity of moral judgments', *Psychological Science* 19(12): 1219–1222.
- Shweder, R.A., Much, N.C., Mahapatra, M. and Park, L. (1997) 'The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering', in A. Brandt and P. Rozin (eds), *Morality and Health*, New York: Routledge, pp. 119–169.
- Terrizi, J.A., Shook, N.J. and Ventis, W.L. (2010) 'Disgust: A predictor of social conservatism and prejudicial attitudes toward homosexuals', *Personality and Individual Differences* 49(6): 587–592.
- Turiel, E., Killen M. and Helwig, C. (1987) 'Morality: its structure, functions, and vagaries', in J. Kagan and S. Lamb (eds), *The Emergence of Morality in Young Children*, Chicago: University of Chicago Press, pp. 155–243.
- Wheatley, T. and Haidt, J. (2005) 'Hypnotic disgust makes moral judgments more severe', *Psychological Science* 16(10): 780–784.
- Zhong, C.B. and Liljenquist, K. (2006) 'Washing away your sins: Threatened morality and physical cleansing', *Science* 313(5792): 1451–1452.

NOTES

- ¹ This paper focuses on other-regarding judgments about character traits, especially negative ones, since these types of critical evaluations are more extensively studied in cognitive science. But it should be pointed out that Hume's official account includes the self-regarding emotions of pride and humility as well.
- ² EPM = David Hume, *An Enquiry concerning the Principles of Morals*, ed. Tom L. Beauchamp; T = David Hume, *Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton. References to EPM and T are followed by corresponding references to the Selby-Bigge editions of the same works by Hume, which are abbreviated as 'SBN'.
- ³ It sometimes appears as if Hume endorses a narrow account of the moral domain, according to which character traits are virtues or vices if and only if they impact utility and are apt to produce the indirect passions of love or hatred in spectators (T 3.1.2.5, SBN 473). Hume writes, 'These two particulars are to be consider'd as equivalent, with regard to our mental qualities, virtue and the power of producing love ... vice and the power of producing ... hatred.' (T 3.3.1.3, SBN 575; compare T 3.3.5.1, SBN 614).
On this restrictive view, moral judgments are made solely on the basis of harm and welfare considerations. We praise *benevolent* agents because we sympathize with the joy and happiness of their recipients (T 3.3.3.2, SBN 602); similarly, we condemn displays of *cruelty* because we vicariously share the pain of their 'miserable sufferers' (T 3.3.3.8, SBN 605). But we can now see that this is not Hume's official position. We have a natural propensity to admire the *great* as well as the *good*.
- ⁴ These eternal and necessary moral truths are supposed to impose practical obligations, moreover, on all rational creatures, including 'the deity himself' (T 3.1.1.4, SBN 456). One might read Hume as pointing out, then, an absurd consequence of Clarkean rationalism: namely that God has a strict moral duty toward cleanliness.
- ⁵ It appears that Hume's early modern and contemporary critics are too quick to reject his broad account of the boundaries of morality as counter-intuitive; these intuitions are not always shared by those who living outside of 'WEIRD' (Westernized, educated, industrialized, rich, and democratic) societies (Henrich et al. 2010: 73).
- ⁶ This helps resolve an important tension in Hume's moral theory. As we have seen, Hume appears to sometimes endorse a narrow account of the moral domain, according to which character traits are virtuous or vicious if and only if they are perceived to impact utility and produce love or hatred in spectators (T 3.1.2.5, SBN 473). But we are now in a position to recognize that Hume is using love and hatred in these passages as 'generic' names. (Ardal 1966: 114). Love and hatred are basic emotions that have derivative forms. Hume often focuses attention on these 'principal passions' without mentioning their derivative forms, since his official view is that they are, for all practical purposes, 'the same affections' (T 2.3.10.31, SBN 448). Hume's official position, in any case, is that the full set of moral emotions is not limited to the four social emotions of love, hatred, pride and humility. In order for character traits to be subjected to praise or blame in common life, it is sufficient that they elicit negative emotions such as contempt and disgust, or positive affections such as sublime awe and elevation.
- ⁷ The same can be said about positive emotions: both love and esteem produce pro-social motivations, but they differ in terms of their relative strengths; love toward the *good* produces a desire to promote their welfare in a 'more eminent degree' than esteem toward the *great* (EPM Appendix 4.6n, SBN 317; compare T 3.3.4.2n, SBN 608).
- ⁸ Hume offers an independent reason to reject the voluntariness condition: it is too strong and would exclude every personal quality from the moral domain. His official position, after all, is that we lack voluntary control over any of our traits (T 3.3.4.3, SBN 609).

- ⁹ Indeed, Hume informs Hutcheson in the September 1739 letter that he had Cicero's *On Duties* 'in my Eye in all his reasonings' about the 'catalogue of Virtues' (Greig 1969: 33–4).
- ¹⁰ Interestingly, positive emotions such as awe and elevation have the opposite effect: negative attitudes toward gay men tended to decrease among participants who watch video clips about heroic or uplifting characters (Lai et al. 2014: 788).
- ¹¹ The fact that these researchers survey attitudes in an orthodox Hindu temple town, moreover, confounds cultural influences with religious ones (Turiel, Killen and Helwig 1987: 201). Hume maintains that theological concerns cause modern moralists to artificially restrict the limits of the moral domain; in this case, it seems, ontological commitments about the spiritual world appear to have moved these boundaries in the opposite direction.

Your short guide to the EUP Journals
Blog <http://eupublishingblog.com/>

*A forum for discussions relating to
[Edinburgh University Press Journals](#)*



EDINBURGH
University Press

1. The primary goal of the EUP Journals Blog

To aid discovery of authors, articles, research, multimedia and reviews published in Journals, and as a consequence contribute to increasing traffic, usage and citations of journal content.

2. Audience

Blog posts are written for an educated, popular and academic audience within EUP Journals' publishing fields.

3. Content criteria - your ideas for posts

We prioritize posts that will feature highly in search rankings, that are shareable and that will drive readers to your article on the EUP site.

4. Word count, style, and formatting

- Flexible length, however typical posts range 70-600 words.
- Related images and media files are encouraged.
- No heavy restrictions to the style or format of the post, but it should best reflect the content and topic discussed.

5. Linking policy

- Links to external blogs and websites that are related to the author, subject matter and to EUP publishing fields are encouraged, e.g. to related blog posts

6. Submit your post

Submit to ruth.allison@eup.ed.ac.uk

If you'd like to be a regular contributor, then we can set you up as an author so you can create, edit, publish, and delete your *own* posts, as well as upload files and images.

7. Republishing/repurposing

Posts may be re-used and re-purposed on other websites and blogs, but a minimum 2 week waiting period is suggested, and an acknowledgement and link to the original post on the EUP blog is requested.

8. Items to accompany post

- A short biography (ideally 25 words or less, but up to 40 words)
- A photo/headshot image of the author(s) if possible.
- Any relevant, thematic images or accompanying media (podcasts, video, graphics and photographs), provided copyright and permission to republish has been obtained.
- Files should be high resolution and a maximum of 1GB
- Permitted file types: *jpg, jpeg, png, gif, pdf, doc, ppt, odt, pptx, docx, pps, ppsx, xls, xlsx, key, mp3, m4a, wav, ogg, zip, ogv, mp4, m4v, mov, wmv, avi, mpg, 3gp, 3g2.*