

Full Citation:

Kelly, D. (2013). 'Moral Disgust and The Tribal Instincts Hypothesis,' *Cooperation and Its Evolution*, Eds. K. Sterelny, R. Joyce, Calcott, B, & B. Fraser. Cambridge, MA: The MIT Press, pages 503-524.

<http://www.amazon.com/Cooperation-Evolution-Life-Mind-Philosophical/dp/0262018535>

Moral Disgust and Tribal Instincts: A Byproduct Hypothesis¹

By Daniel R. Kelly
Department of Philosophy
Purdue University, West Lafayette, IN, USA

I. Introduction: Some Interesting Facts about Morality

Psychological research has been discovering a number of puzzling features of morality and moral cognition recently.² Zhong & Liljenquist (2006) found that when people are asked to think about an unethical deed or recall one they themselves have committed in the past, issues of *physical cleanliness* become salient. Zhong & Liljenquist cleverly designate this phenomenon the “Macbeth Effect,” and it takes some interesting forms. For instance, reading a story describing an immoral deed increased people’s desire for products related to cleansing, like a shower soap, disinfectants, or antiseptic wipes. Moreover, Zhong & Liljenquist found that cleaning one’s hands after describing a past unethical deed actually reduced moral emotions such as guilt and shame. So much so that those who did “wash away their sins” were less likely than other participants to help out a desperate graduate student. Other researchers report similar findings. Schnall and her

¹ I would like to thank the participants in the 2008 Minds and Societies conference at UQAM and the participants in the 2008 Emotion and Commitment conference at ANU for helpful comments on this material.

² In what follows, I’ll call attention only to those findings that are relevant to my argument. For more encompassing overviews of recent work in empirical moral psychology, see Doris & Stich (2005, 2006) and Doris et al. (2010).

colleagues explored how issues of cleanliness influence judgments of moral *severity*. In one experiment (Schnall et al. 2008b), feelings of disgust were induced in participants in a number of ways, including having them remember a disgusting experience, having them watch a disgusting video, or having them fill out questionnaires at a filthy desk or in the presence of a foul and unpleasant odor. Schnall and colleagues found that many participants' moral judgments about events described in vignettes were more severe in such disgusting conditions, even though what triggers the feelings of disgust (a bad smell, a disgusting memory) is "extraneous" to the events described in the vignettes, about which the moral judgments are putatively being made. In another (Schnall et al. 2008a), one group of participants was made to physically wash their hands between experiencing disgust and making a series of moral judgments, while a second group was not. Those who had just washed themselves down were significantly less severe in their condemnation of the vignettes they were asked to evaluate.

Another series of experiments has sought to demonstrate a link between morality and worries about contamination – or more precisely, it explored the link between often tacit worries about contamination, moral taint and *immorality*. Paul Rozin and his colleagues found that many people are slightly reluctant to put on a sweater that once belonged to and was worn by an undescribed stranger, even if it had subsequently been thoroughly cleaned. They also found that people tend to become increasingly reluctant to put on, and in some cases even *touch*, a similarly laundered sweater if they are told that the previous owner had committed some extreme moral violation such as murder. The link between immorality and contamination was made especially evident by the piece of clothing that people were most reluctant to come into contact with, which was treated as

the most aversive and contaminated of all: a sweater that once belonged to the ultimate moral monster, Adolf Hitler (Rozin et al. 1994, also see Haidt et al. 1997 for discussion).

Coming at the issue of morality from another angle are those exploring prejudicial attitudes that members of one group may harbor about members outside their own group. As is all too familiar, either from anecdotal report or even from common first hand experience, members of one cultural “ingroup” will often consider members of other cultural “outgroups” to be below them in one way or another. In exaggerated cases of this, one group is liable to completely demonize and dehumanize an other, considering them not fully human but merely animal, undeserving of any moral treatment or consideration whatsoever, and somehow tainted and tainting. The Indian caste system is often said be deeply informed by worries about the threat of contamination flowing up from lower to higher castes, or even from those, like Europeans, outside the caste structure itself (see Bouglé 1971, especially 22-23; also see Shweder et al. 1997 and Rozin et al. 1999). Perhaps the most extreme and well-known example was the Nazi attitude towards Jews. Indeed, anti-Jewish Nazi propaganda flagrantly invoked the imagery and language of purity, contamination and dehumanization. Hitler rhetoric portrayed Jews as maggots in a festering abscess, hidden away inside the clean, healthy body of the nation. Experimenters have begun exploring the psychological bases of a variety of different kinds of prejudices, and one of the most interesting findings has been that amongst members of a particular ingroup, different emotions are commonly associated with attitudes towards different outgroups (Cottrell & Neuberg 2005).

Finally, Wheatley & Haidt (2005) have used the emotion of disgust to induce some striking and seemingly irrational moral judgments as well. In one of their most devious experiments, they ask subjects to consider the following vignette:

“Dan is a student council representative at his school. This semester he is in charge of scheduling discussions about academic issues. He often picks topics that appeal to both professors and students in order to stimulate discussion.”

All of those given this vignette were asked to rate how morally wrong Dan’s behavior is. However, some of the subjects were also hypnotically induced to undergo a brief flash of disgust when they saw the word “often,” which occurs near the middle of the vignette. Amazingly, many of those hypnotized subjects judged Dan to be doing something morally wrong! They arrived at this judgment despite the completely innocuous description of Dan’s actions. Moreover, they were remarkably persistent in standing by their initial impulses; they upheld their dim opinion of Dan even after admitting they had little or no justification for the judgment. Indeed, when asked to explain themselves, participants ended up saying things like Dan seems like a “popularity seeking snob,” and that he “just seems like he’s up to something.” Most revealing, perhaps, was one subject who bluntly stated “I don’t know [why it’s wrong], it just is.”

The explicit mention of disgust in the last example points to a common thread running through all of these otherwise disparate findings. Indeed, these are just a few examples of recent work that has been exploring a link between disgust and morality that is fascinating, puzzling, and often troubling. In the rest of this paper, I will offer an account of moral disgust that illuminates this link, and makes sense of some of its more unsettling features. In the next section, I will briefly sketch the tribal instincts hypothesis,

a component of gene-culture coevolutionary theory, which posits a set of cognitive mechanisms that allow humans to navigate an important part of their social world, and which will provide a theoretic background for the rest of the paper. In the third section, I will advance and defend my account of the basic disgust response. This account, encapsulated by what I call the Entanglement thesis, holds that disgust has a pair of primary functions, one associated with protecting the digestive system from potentially poisonous foods, and the other geared towards preventing disease and parasitic infection in general. Finally, in the fourth section, I put the pieces together and argue for the Co-opt thesis, which holds disgust acquired several auxiliary functions associated with sociality, tribal instincts and morality, without losing the features that allowed it to adequately perform those primary functions associated with poisons and diseases. Finally, I show how my account is able to shed light on some of the puzzling features of morality and moral cognition mentioned above, for instance how concerns with spiritual purity or moral taint can be explained as products of a mismatch between elicitor and response, produced when this emotion was brought to bear on issues that it did not initially evolve to deal with.

II. The Tribal Instincts Hypothesis: Social Norms and Ethnic Boundary Markers

The tribal instincts hypothesis supposes that the human ability to interact and cooperate on the scale of entire tribes and cultures is facilitated by a number of reliably produced impulses, which are at least partly innately structured. Indeed, Richerson and Boyd (1998,1999) argue that human “ultrasociality” is too complex to be fully explained by appeal to only kinship and reciprocity, the standard resources of evolutionary theorists

attempting to explain cooperation, and so these tribal instincts promise to be a crucial part of a complete explanation of distinctively human social structure. The tribal instincts hypothesis itself is derived from a much broader theory concerned with human nature and evolution, namely gene-culture coevolutionary theory, or GCC for short. As its name suggests, both culture and cultural evolution, on the one hand, and genetic and biological evolution, on the other, fall within the scope of GCC. However, it does not treat either of these topics in isolation from the other. Rather, GCC sees cultural and genetic evolution as deeply intertwined in humans, and seeks to understand the ways in which culture can and has influenced genetic evolution, and alternatively how genetic factors have influenced the evolution of cultures (Boyd & Richerson 1985, 2005a; also see Richerson & Boyd 2005 for a more accessible overview).

GCC conceives of culture very broadly, as information in the social environment that can be passed from one member to another by social (therefore non-genetic) means. Culture can be transmitted not just through populations but also across generations. Information accumulates as it is passed from one generation to the next, and as such, the entire body of information can be seen as an inheritance system, sharing some analogous properties with the genetic inheritance system. One factor motivating the tribal instincts hypothesis is the insight that continuous and increasing reliance on the information in this cultural inheritance system imposes new requirements on those who rely on it. As the volume and import of the information in the cultural inheritance system both increase, new selective pressures would be created that favor certain psychological capacities, namely those capacities allowing individuals to easily access and use information stored in the cultural repository.

One important result of humans' extended immersion in culture – and according to GCC there are many – is that among the many types of socially relevant cues to which humans are sensitive (indications of kinship, hints about others' intentions, etc.) are cues about norms and indicators of what kinds of norms others embrace. That is, humans have become innately disposed to see their social world in tribal terms, and to react accordingly.³ The enfolding of cultural and natural selective pressures in humans evolution has produced a set of social, tribal “instincts” that are sensitive to particular types of cultural information, namely information that facilitates living within the context of large, cooperative groups or tribes. Into this category falls information about both *social norms* and *ethnic boundaries*.

It is nearly a platitude that human social interactions are regulated by complex systems of norms. GCC sees enormous importance in this platitude, however, and suggests that it is largely these systems of behavior-guiding social norms that make it possible for humans to cooperate or even smoothly interact in large groups. Pairing this evaluation of the importance of social norms with the tribal instinct hypothesis yields the prediction that humans will have a distinctive capacity to cognize those social norms. Recent research suggests that this is, indeed, the case. While the details are far from

³ A note on terminology: the name “tribal instincts hypothesis” I am taking from gene culture coevolutionary theorists like Boyd and Richerson and their collaborators. However, in what follows, I will focus primarily on only two specific types of “tribal instincts,” namely those psychological capacities associated with social norm and those associated with ethnic boundary markers. I take no stand on whether there are other, distinct, tribal instincts beyond these two. Moreover, when I use the term “tribe” I will be guided by the usage in that literature, and so mean, roughly, groups of several hundred people who may not be bound together by relations of kinship or histories of direct cooperative interaction, but who are culturally homologous in that they share beliefs, values, and most importantly, embrace a common cluster of norms and display common ethnic markings. For some discussion of the troublesome terminology, see Gil-White (2006).

I should note that the literature on gene culture coevolutionary theory is not only complex, but it is large and growing. My discussion here will be of necessity somewhat cursory, but I hope to give the flavor of the outlook and highlight those elements most important to my aim of illuminating the nature of what is often called “moral disgust”.

settled, it appears that the capacity has a number of important features, including the ability to easily acquire and internalize norms from the social environment. Once a norm is internalized, it produces characteristic types of motivation to both comply with the norm and to punish those who violate it (Sripada & Stich 2007, see also Nichols 2004).

Whereas social norms help coordinate social interactions within a tribe, ethnic symbols serve to mark the boundaries between different tribes. Such symbols, or ethnic boundary markers, as they are often called in the GCC literature, allow members of the same tribe to identify and selectively interact with each other (McElreath et al. 2003, Barth 1969). Members of the same tribe, in this sense, share a large set of beliefs and values. More importantly, they also share large clusters of social and moral norms. Such beliefs, values, and norms are not themselves immediately visible to the naked eye, however. More directly and easily detectable symbols of various sorts (displayed colors, styles of clothing, use of different dialects, varieties of cuisine, and so forth) often serve to signal information about which norms, values, and beliefs a person holds, and thus which tribe he or she belongs to.

Moreover, there is a sound rationale for displaying such symbols: coordinated interactions will go much more smoothly and efficiently than non-coordinated interactions, and other things being equal, it is in no one's interest to engage in the later, more difficult sort of exchanges. Social norms often govern social interactions in such a way that actors who share the same relevant norms will have similar and complementary expectations about the "proper" form of the interactions, practices, and customs in which they might mutually engage. Against this backdrop, ethnic markers perform an important signaling function: they serve as visible cues for the behavioral dispositions and

otherwise unobservable values and norms that guide the behavior of others. The perceivable symbolic markers thus allow actors to identify and selectively interact with those who have similar and complementary expectations about the type of interaction in question, and to avoid difficult and inefficient exchanges with those who do not.

Pairing this idea with the tribal instincts hypothesis also yields a prediction about human psychology: humans will tend to express commitment to their tribe and their tribe's characteristic cluster of norms by displaying ethnic boundary makers, will be predisposed to be sensitive to the ethnic boundary markers displayed by others, and will be motivated and make inferences about whom to socially engage and whom to avoid based on them. Moreover, such capacities can give rise to ethnocentric attitudes in favor of one's own tribe (and perhaps, to a lesser extent, friendly and allied tribes) along with its members, customs, values and norms. The dark flipside of this is that those same tendencies can all too easily give rise to prejudicial attitudes against other, hostile tribes, and their members, customs, values and norms. In an eloquent expression of this line of thought, Boyd and Richerson speculate that such attitudes often involve the emotion of disgust:

“[G]roups of people who share distinctive moral norms, particularly norms that govern social interactions, quite likely become ethnically marked. This suggests that ethnocentric judgments easily arise because “we the people” behave properly, while those “others” behave improperly, doing disgusting, immoral things, and showing no remorse for it, either.” (Boyd & Richerson 2005a, 101)

III. The Basic Disgust Response and the Entanglement Thesis: Poisons & Parasites

For the moment, set aside the tribal instincts hypothesis and consider the emotion of disgust itself. A wide range of empirical work has shown that the basic disgust response is comprised of a diverse but highly coordinated set of elements, including affective, behavioral, and cognitive components (see Ekman 1992, Rozin et al. 2008). Among the most recognizable of the behavioral components is the gape face, the characteristic facial expression associated with disgust. In especially intense episodes, production of the gape face can tip into the retching that it so clearly prefigures. Gapes are also accompanied by a feeling of nausea, an orally based sense of aversion, and a reflex-like withdrawal, the quick physical recoil from the disgusting entity. The basic disgust response includes cognitive components as well, namely a more sustained sense of offensiveness and contamination. When some entity is found disgusting, it is considered offensive in a particular way: the thing is repulsive; one does not want to come into physical contact with it; mere physical proximity to the entity often is off-putting, repugnant, barely tolerable. Moreover, while a disgusting entity often captures the attention, even thinking about it is unpleasant. More striking, an entity that is considered disgusting has the ability to transmit its disgustingness to other things it comes into contact with. Those things thus contaminated are thereby considered disgusting, and elicit the same suite of response elements. Together, the operation of these two cognitive components of disgust can quite naturally lead to concerns about cleanliness and desires purify oneself.

Opposite the basic disgust response are the sorts of things that can trigger it. Here, the data are even more puzzling (see again Rozin et al. 2008). Disgust can be elicited by an extraordinarily diverse set of triggers, ranging from the concrete and physical to the

abstract and social. On the one hand, some of the most universally disgusting things are closely associated with the body, like spit, feces, blood, and organic decay of all sorts. Disgust is also sensitive to bodily boundaries in a particular way; the emotion enforces a “no re-entry” policy. If something was once within or a part of a body, even your own, but then exits or breaks off, it thereby becomes an object of disgust; common examples include saliva, blood, hair, fingernails, and severed limbs. Also disgusting to many is a set of creatures that might be called “creepy crawlies”: slugs, roaches, rats, and the like (Davey et al. 1993, Webb and Davey 1993, Ware et al. 1994). Certain types of perfectly edible (i.e. non-poisonous) food disgust some people as well. Common offenders in this category include cuisine like Brussels sprouts, escargot, caviar, pork rinds, Whoppers and deep-fried Twinkies. Other common elicitors of disgust are non-standard sexual behaviors and practices, including most notably incest – though what counts as non-standard, and, in the case of incest, which kin are off limits, varies from culture to culture. Finally, certain types of social behavior can elicit disgust. Crooked politicians and ambulance chasing lawyers are emblematic of such behaviors in our own culture, but violating certain social norms, especially those that govern how antecedently disgusting entities are to be dealt with (norms regulating burial rituals, the correct way to prepare food or maintain bodily hygiene, etc.) can also often elicit disgust (Shweder et al. 1997, Rozin et al. 1999, see also Haidt et al. 1997).

Another noteworthy feature of the elicitors is that some appear to be universally and perhaps innately disgusting. These include those elicitors closely linked to the body, as well as incest (Lieberman et al. 2003, Fessler & Navarette 2004; cf. Prinz 2008 for a dissenting view). In the case of other elicitors, however, what is considered disgusting

can exhibit patterned variation from culture to culture. For instance, whether it is escargot, caviar, and Brussels sprouts that are typically considered repulsive by the locals, or whether it is pork rinds, Whoppers and deep-fried Twinkies, will depend on whether you are at a state fair in the US's rural Midwest or at a posh bistro on Paris's Left Bank. Likewise, while many cultures consider some types of deviant sexual activities not just wrong but also disgusting, exactly which of those activities are deviant in this way varies from culture to culture. In terms of social behaviors, as the relevant norms vary from one group to another, so will the transgressions that are considered disgusting. In extreme cases, the norms and even ideologies of entire opposing social groups can come to disgust as well, e.g. conservative ideology can be mildly disgusting to liberals; liberal ideology likewise to conservatives.⁴

Taking a step back and surveying all of these data invites some difficult questions. First and foremost is the simplest one: given 1) this puzzling and variable array of elicitors, that 2) evokes a response composed of an equally puzzling cluster of components, *how are all of these things connected?* My aim is to sketch an answer to this question. That answer comes in two parts: the Entanglement Thesis and the Co-opt thesis. The rest of this section will be occupied with the former, while the next section will take up in the later.

⁴ This need not imply that *every* time someone describes themselves as “disgusted” by something they actually are disgusted, that they are correctly reporting the full activation of their disgust system. Surely sometimes claiming to be disgusted by, for instance, a particular type of campaign finance reform can be merely one way to verbally express fervent disagreement or outrage. However, there is reason to believe that at least sometimes, social and moral issues can genuinely activate the disgust system, and thus produce (mild) episodes of disgust, complete with the full suite of components. See Chapman et al. 2009; Rozin et al. 2009; for a skeptical view see Bloom 2004; for discussion see Kelly in press, especially chapter 4).

The Entanglement thesis holds that disgust is a uniquely human kludge.⁵ Underlying the basic disgust response are two distinguishable cognitive mechanisms that were once distinct, but have become deeply entangled with each other in modern human beings. Through the evolutionary process of descent with modification, these two mechanisms became more and more functionally integrated with each other, eventually forming the single emotion now recognized as disgust. The character of that human emotion remains informed by the character of those two entangled mechanisms and the adaptive problem each initially (and separately) evolved to solve. However, while homologies with similar features and functions to each distinguishable mechanism can be found in primates and other animals, the Entanglement thesis holds that only in humans have these two mechanisms become so tightly intertwined as to form disgust. Thus, the Entanglement thesis also provides an explanation for why this particular emotion is found in human beings, but not other animals (see Rozin et al 2008, Morris et al 2007).

One of those two entangled mechanisms is directly linked to digestion. This mechanism initially evolved to regulate food intake and to protect the gut and gastrointestinal system against substances that are poisonous, toxic or otherwise harmful when swallowed. It mainly protects against such substances by preventing them from being fully ingested. The mechanism can also produce orally based aversion towards specific types of food, to prevent them being eaten. Indeed, foods that, once fully consumed,

⁵ The term “kludge” is taken from engineering and computer science, where it is usually used to refer to a clumsy, piecemeal, or inelegant solution to a problem, or the clumsy, piecemeal, inelegant device used to solve the problem. Kludges are often gerrymandered, constructed to fix problems that were themselves unanticipated. The clumsiness stems from the fact that kludges are often constructed out of whatever parts are available when the problem unexpectedly arises. Those initially unrelated parts are cobbled together to construct the kludge, and in the process often put to new uses. Given that natural selection is a blind, tinkering process that operates without the benefit of foresight, the term also comes in handy in evolutionary explanations, though obviously in such contexts the metaphorical use of terms like “unexpected” or “unanticipated” needs to be taken with a grain of salt.

induced gut-based distress are often not only expelled, but also generate what have been called acquired taste aversions, so that they are not consumed again (Garcia et al. 1974, Bernstein 1999). For shorthand, I will call this the *poison mechanism*.⁶

Returning to the characterization of disgust offered above, certain features of the basic response are easily traced to the poison mechanism and its proprietary adaptive function. In general, the aversion in episodes of disgust is often produced via physiological systems primarily based in the mouth and gut, giving it a strong oral aspect (what Rozin calls a “sense of oral incorporation.”) The gape face often precedes and uses many of the same muscle groups as retching, which is how the body expels substances from the gut and mouth. The accompanying feeling of nausea is also useful in preventing ingestion of food in the first place. Finally, culturally local patterns of disgust to certain types of cuisine also indicate the operation of a mechanism dedicated to monitoring food intake.

The other of the two entangled mechanisms that shaped human disgust is linked to disease and parasites. This mechanism originally evolved to protect the entire organism from all forms of pathogenic or parasitic infection. It does this by causing the organism to avoid any close physical proximity to infectious agents, or anything that is likely to be infected and contagious. Since many infectious agents are microbes that cannot themselves be detected by the naked eye, protecting against infection involves avoiding not only visible pathogens and parasites, but also avoiding places, substances and other organisms that might be harboring them. Unlike the one protecting against poisons, this mechanism is not specific to ingestion, and so obviously has a much larger

⁶ Previous theorists have emphasized this feature of disgust. For example, Griffiths (1997) follows Darwin (1872) in casting the emotion as a food rejection system, and much of Rozin’s work on disgust stresses its oral character (see Rozin et al. 2008 and citations therein).

domain and range of cues to monitor. I'll often refer to this as the *parasite mechanism* in what follows.⁷

Once again returning to the characterization of disgust offered above, certain features of the basic response are easily traced to the parasite mechanism and its proprietary adaptive function. One of these is the reflex-like withdrawal: quickly recoiling from a disgusting entity instantly decreases its physical proximity. The more cognitive sense of offensiveness can effectively prevent getting close to disgusting entities in the first place, and motivate moving further away from them once they are detected. Finally, the sensitivity to the possibility of contamination, the motivation to cleanse oneself and the concern about physical purity are all clearly fitted to the adaptive problems linked to infection.

As in the case of the poison mechanism, evidence of the parasite mechanism can also be found in the common elicitors of the emotion. While disgust does appear to have a special link to the mouth (the intensity of an episode of disgust can usually be increased by imagining the offending entity coming into contact with one's mouth or tongue), its domain is by no means restricted to the oral; the emotion monitors all of the bodily orifices and boundaries (Rozin et al. 1995; Fessler & Haley 2006). Feces and organic decay are some of the most effective vectors of disease transmission, and are also some of the most potent elicitors of disgust, perhaps universally so (Haidt et al. 1994). Finally, phenotypic abnormalities and other reliable indicators of infection in conspecifics are also possible universal and innate elicitors of disgust (Curtis et al. 2004).

⁷ Psychologists have previously noted this aspect of disgust as well. Steven Pinker (1997), for instance, calls disgust "intuitive microbiology", and Curtis et al. (2004) present an impressive body of evidence in support of the claim that disgust is an "evolved response to threats of infectious disease." Kurzban & Leary (2001) cite "parasite avoidance" behavior to explain a certain type of social stigmatization, though the link to disgust is less explicit in their discussion.

Another form of evidence in support of the Entanglement thesis comes from developmental psychology. Patterns in ontogeny suggest that components of the disgust response are on different developmental schedules, and emerge at different ages. Indications of mere distaste (not liking the taste of certain foods, but not reacting to them with disgust) and the ability to make and react to gape faces are present within the first year of life (Bandura 1992). Other components of the mature disgust response such as contamination sensitivity, however, do not emerge until much later. Some researchers mark the usual appearance of full blown disgust as late as four to eight years (Rozin et al. 1985), while others put it earlier, around two and a half to three years (Siegal and Share 1990). Whichever turns out to be closer to the truth, both estimates place components of the response linked to the parasite mechanism as emerging significantly later than those linked to the poison mechanism.

While the Entanglement thesis makes sense of much of the extant data and proximate structure of disgust, it also raises questions about the ultimate explanation of the emotion, and the sorts of selective pressures that drove the poison and the parasite mechanisms together. Though evolutionary accounts of mental structure are notoriously difficult to confirm or falsify, there are several factors that would have been instrumental in creating what we now recognize as disgust in modern humans. First, there must have been a non-trivial degree of functional overlap between the two systems; food is major vector of disease transmission, and so the parasite mechanism would have antecedently been sensitive to food related cues, making the two mechanisms easily poised to become entangled by the right kind of tweak here or there. A major, related tweak, however, might have come when ancestral humans began to consume more and

more meat, procured either by advances in hunting or active scavenging. The increase in meat consumption marks an important difference between humans and their primate cousins, and has been linked to increased brain size (see Aiello and Wheeler 1995, Sterelny 2003). This is relevant to the evolution of disgust because ancestral humans did not have a long history of scavenging, and so would not have been able to simply rely on a scavenger's powerful gastro-intestinal system to neutralize the dangers in the larger quantities of meat they were developing the wherewithal to acquire. This situation is just the sort that would have pressured the parasite mechanism to become even more sensitive to potential foods, thus creating the kind of selection that would have pushed that mechanism to simply merge with the poison mechanism.

This line of thought also suggests why researchers fail to find anything fitting the description of disgust in other mammals like dogs, who *are* long time scavengers, and have evolved the type of iron-clad gut to accommodate it, or in other primate species, who's evolutionary past *did not* include the transition to a meat intensive diet the way human's did. As such, an ultimate explanation along these lines is compatible with the claim that disgust is a uniquely human emotion, while at the same time being compatible with there being unentangled homologies of its component parts present in a range of other species.⁸

⁸ Another factor that could have helped push the poison and parasite mechanisms towards entanglement has to do with the increased need for a perspicuous signal that early humans could have used to socially transmit (perhaps parochial) information about disease and parasite related issues in the local environment. The gape face, associated with retching and the poison mechanism, provided a good candidate. Selection for an effective signaling system would have generated pressure to link the gape with the parasite system, which could have also contributed to the integration of the two systems. Moreover, there would be little incentive for deception given the relevant adaptive problems, especially those having to do with contagious disease, and so no clear selective pressure militating *against* the evolution of a clear, trustworthy signaling system. See the third chapter of Kelly (in press) for more detailed discussion.

Finally, it is worth noting that despite the differences in ontogenetic and evolutionary history, the Entanglement thesis holds that the poison and parasite mechanisms have merged in mature, modern humans. Once the emotion of disgust is fully developed in an individual, the many components of the response come as a package; they are thereafter produced together with law-like regularity, forming what philosophers of science sometimes call a “nomological cluster” (Boyd 1991). In the case of disgust, this means that any elicitor will reliably produce those clustered components, both those linked to the evolutionary problem of food regulation (reflexive production of gape face, nausea, sense oral incorporation) and those linked to pathogen and parasite avoidance (quick withdrawal, sense of offensiveness, sensitivity to contamination, and the desire to cleanse or purify). While the intensity or vividness of different episodes of disgust will obviously vary from the mild to the extreme, elicitors of all sorts trigger this full nomological cluster, from creepy crawlies, to bodily fluids, to the relevant types of moral transgressions.

IV. Moral Disgust and the Co-opt Thesis

Recall the question posed earlier. Given, on the one hand, the puzzling cluster of affective, behavioral, and cognitive components that make up the disgust response, and on the other hand, the equally puzzling array of elicitors that trigger the response, how are all of these things connected? The first part of my answer was largely (though not exclusively) focused on the character of the response itself. The second part of my answer is the Co-opt thesis, which assumes the Entanglement thesis is correct, and

embeds it within the context of GGC and the tribal instincts hypothesis.⁹ It takes the set of elicitors as its point of departure.

The Co-opt thesis holds that as humans became more reliant on social groups and the cultural information they provided, basic disgust was co-opted by the emerging tribal instincts to help perform a variety of novel functions that arose in conjunction with this increased sociality. In doing so, disgust's most characteristic features, features that initially evolved to solve adaptive problems linked to poisons and parasites, were brought to bear on those new functions in the social domain. Moreover, it is exactly this *imperfect fit* between the basic disgust response and many of those social functions it was later co-opted to perform that gives rise to the sorts of puzzling results turning up in the recent research on moral cognition. In short, some of the more troubling features of moral judgments discussed in the first section can be understood as cognitive byproducts, generated by the mismatch between “unanticipated” problems and the kludgy solution disgust helps provide.

As an example of co-optation, consider the gape face discussed above. Gaping utilizes most of the same facial muscular movements as retching, the physical act that it sometimes precedes and accompanies. According to the Entanglement thesis, as the poison and parasite mechanisms fused, however, that facial expression was acquired a new purpose: it was recruited to send signals to conspecifics. Generally speaking, the importance of communication increased in tandem with human sociality and the propensity to live in larger and larger groups. As the significance of communication rose,

⁹ While my presentation of the Co-opt thesis assumes the Entanglement thesis is correct, the two are logically distinct. Others may consistently subscribe to something similar in spirit to the Co-opt thesis, i.e. the idea that disgust initially evolved to do something other than the functions it now performs in human moral psychology, while rejecting the Entanglement thesis, and replacing it with an alternative account of what, exactly, disgust initially evolved to do.

so too did the need for perspicuous signals that could cleanly transmit important information. Faces and different facial expressions in general could already convey a rich assortment of information, and on my view, the gape face was co-opted to perform a similar signaling role. It can warn others, including small children, against eating something known to be toxic or poisonous. The gape also expresses a sort of “Warning! Biohazard!” message, useful for cautioning others to avoid nearby pathogens or contaminated areas. In being co-opted, the gape went from merely preceding the actual expulsion of substances from the mouth to acting as a warning sign. Moreover, once this broader signaling function was acquired, gapes were able to transmit other socially relevant information as well, including information related to the other functions disgust acquired in the social and moral arenas.¹⁰ For instance, expressions of disgust, together with the types of things that elicit them – disgust towards specific types of food, like deep-fried Twinkies or escargot, or towards particular behaviors and social practices, like driving a gas guzzling SUV or smoking cigarettes – can themselves act as ethnic boundary markers, signaling information about group membership, and hence which values and norms a person is committed to.¹¹

¹⁰ It will be useful to make terminological caveat, in order to forestall a number of (interesting, open) philosophic questions about the definition of “morality”, and thus the proper domain of the moral – which judgments, mechanisms, or roles are “really” about moral issues and which are not. (For discussion of these issues, see Nado et al. 2009, Kelly & Stich 2007, Nichols 2004). In the text, “moral disgust” will be used in the same way it is used in the empirical psychological literature, namely to capture those roles that disgust systematically plays in social life. These include its role in guiding social coordination, motivating behavior in potentially cooperative situations, influencing interactions between individuals and between groups of people, and influencing judgments about similar matters.

¹¹ It is worth pointing out that on the explanation offered here, expressions of disgust have acquired a role in signaling commitment, but it is commitment to a group and the set of norms that bind it together. This is an importantly different type of commitment than that often associated with emotional expression, according to which social emotions like guilt and anger are fundamentally commitment devices that evolved to help navigate the Machiavellian vicissitudes of deception, credibility, and defection (for the classical statement of this view, see Frank 1988; see also Pinker 1997). For more discussion of the similarities and differences between these two types of commitment and the different explanations of cooperation they are associated with, see Richerson & Boyd (2001) and Boyd & Richerson (2005b).

For more insight into the character of the auxiliary functions that has disgust acquired, recall GCC and the tribal instinct hypothesis. GCC maintains that one factor which greatly contributes to the human ability to cooperate and coordinate on such a large scale (compared, e.g. to other primates and most other animals) is that human social interactions are governed by a complex set of norms. Common sense and anecdotal evidence is supported by recent research showing that disgust is indeed operative in a number of different types of these social norms. In these cases, the emotion provides the types of intrinsic motivation mentioned above, including motivation to comply with the norm in question, to avoid the actions they prohibit, and to punish or direct punitive attitudes at transgressors of the norm. Indeed, disgust has been shown to play such roles in a number of different types of norms, including the rules of table etiquette (Nichols 2002a, 2002b, 2004), taboos restricting the consumption of meat (Fessler & Navarrete 2003), and taboos against incest (Lieberman et al. 2003, Fessler & Navarrete 2004).

More generally, the anthropologist Richard Shweder and his colleagues have called attention to an entire class of norms that follow the logic of disgust, which they calls purity norms (Shweder et al. 1997, Haidt et al. 1997, Rozin et al. 1999). As their name suggests, purity norms are often understood as regulating issues of purity, not only guarding the sanctity of the physical body, but also protecting the soul from contamination and spiritual defilement. Indeed, purity norms are often distinguished from other classes of norms, such as harm norms or fairness norms, in that transgressions of purity norms usually do not result in direct physical harm or the inequitable treatment of any person.¹² More traditional or religious cultures often see transgressors of a purity

¹² I do not take these categories – harm norms, purity norms, fairness norms – to delineate categorically distinct and disjoint sets, but rather poles on a continuum. One complicated and contentious issue is the

norm as defiling themselves by disrespecting the sacredness of God (or the gods), or by violating the divine order. Purity norms are not completely absent from largely secular cultures, however; their presence is just not as central to the social structure or prevailing moral code. They are often given a different justification in secular cultures, as well: transgressions of purity norms are usually conceived of as “crimes against nature” or violations of the natural order. According to Shweder, norms fitting this description regulate a range of issues, such as the proper foods to eat, when it is admissible to eat them, and often the proper way to prepare them; the details of sexual activities and even sleeping arrangements amongst family members; proper attire in a variety of settings, especially ritual and religious settings; the proper way to deal with organic materials, like corpses, blood, feces, and so forth; and how to interact with members of other social groups, particularly how to avoid being polluted by members of lower classes or castes. In addition to the obvious themes of purification and contamination, preliminary research supports the idea that the character of purity norms is heavily influenced by the emotion of disgust (see especially Rozin et al 1999).

Also confirming commonsense suspicions are recent neuroimaging experiments that link the disgust response to prejudices and ethnic membership. This research shows disgust to be operative in sustaining a class of biases and prejudicial attitudes towards those in particular outgroups or tribes.¹³ As was mentioned above, distinct emotions are

role of perceived harm in various norms and norm transgressions, including so-called purity norms; see Turiel et al. 1987, Haidt et al. 1993, Kelly et al. 2007, Sousa et al 2009, and Stich et al. 2009.

¹³ In using the term “tribes” here, I do not mean to suggest that today’s social networks are still structured along anything approximating tribal lines. However, this is completely compatible with the possibility that important components of the human psychological system devoted to social cognition still *see* social interactions in largely tribal terms, and conceptualize and attempt to navigate those interactions accordingly. Indeed, this is one way to understand the tribal instinct hypothesis: many of the mechanisms underlying social cognition, even in contemporary human beings, originally evolved to allow living in tribal sized groups, and as such they are sensitive to the types of observable cues that are likely to convey

often associated with the different types of attitudes directed at different outgroups and their members (Cottrell & Neuberg 2005). Particularly interesting (if not completely surprising) is the demonstration that disgust is often the emotion linked to the most extreme prejudices, directed at members of the lowliest, most vilified and dehumanized ethnicities (Harris & Fiske 2006).

Finally, the pieces are ready to be put together. On the one hand, GCC provides details about a number of relatively novel adaptive problems that arise in the wake of increased human sociality and reliance on cultural information, and posits a set of tribal instincts that evolved to help deal with them. On the other hand, we have basic disgust, an emotion that appears to have been cobbled together from parts that originally and separately evolved to deal poisons and parasites, but which also appears to be acting as an important component of certain tribal instincts. The Co-opt thesis offers an explanation: at some point, the poison and parasite mechanisms that comprise basic disgust were co-opted to perform a variety of novel *auxiliary* functions unrelated to either poisons or parasites. Furthermore, the Co-opt thesis maintains that in performing those novel functions linked to social norms and monitoring ethnic boundaries, the full nomological cluster of components that make up the basic disgust response is brought to bear on those social functions, from the more cognitively complex sensitivity to contamination, to the gape face and physical recoil, to the more visceral feelings of nausea and repulsion. Moreover, those social behaviors and social attitudes driven by disgust will be also informed, perhaps implicitly, by the components of the emotion, for instance worries

information relevant to which tribe someone is a member of, and which cluster of norms they embrace. Moreover, those mechanisms are wont to process that information, make inferences, and form intentions in ways that are well suited to a tribal existence, even if the inferences and intentions they produce are not always optimal or even efficient in the context of our modern social institutions.

about contamination and feelings of revulsion, that initially evolved in response to adaptive problems associated with toxic foods and diseases avoidance.

As such, the Co-opt hypothesis opens up the possibility of explaining persistent concerns about contamination and purity in moral affairs to be understandable but misplaced. Such concerns are revealed as byproducts of the imperfect fit between the character of the disgust response and the new function it has been co-opted to perform in conjunction with human tribal instincts. The social norms that recruit disgust appear to require, most basically, some kind of avoidance and aversion motivation. In co-opting disgust in particular, the activities proscribed by those norms, as well as those actors who transgress them, are not *simply* avoided and found aversive. Rather, they are also subliminally infused with a very specific kind of offensiveness, are often considered tainted and contaminating, so much so that they can induce a desire to cleanse or purify oneself. The same type of explanation applies to tribal instincts that monitor ethnic boundaries and their symbolic markers. According to GCC, what is needed is motivation to avoid members of other tribes who have internalized different social norms, in order to avoid uncoordinated (and perhaps hostile) exchanges. When disgust is the emotion co-opted to provide that motivation however, along with it come attendant components like contamination sensitivity, offensiveness, visceral aversion – the full nomological cluster of the disgust response.

Also troubling can be the way in which feelings of disgust can induce judgments that are remarkably persistent. Remember Dan the “popularity seeking snob”:

“Dan is a student council representative at his school. This semester he is in charge of scheduling discussions about academic issues. He often picks topics that appeal to both professors and students in order to stimulate discussion.”

Those hypnotized to feel disgust at the word “often” maintained their initial judgment that Dan was doing something morally wrong even when they were unable to provide any supporting reasons. The vividness and visceral power of the emotion could lead people to remain doggedly committed to other attitudes and norms that involve disgust, even if those attitudes and norms can be shown to be unjustified or rationally unfounded (cf. Haidt et al. ms).

V. Conclusion

While the view of disgust advanced here is primarily descriptive and explanatory, it is tempting to think that it has normative implications, and that the account provides the materials to construct an argument concerning what sort of moral authority ought to be accorded the emotion, and what role feelings of disgust deserve in moral deliberation and debate. Though I believe this suspicion is correct, here is not the place to spell out and defend such an argument (though see Kelly in press, chapter 5). Rather, I will conclude by pointing out that whatever its conclusion, an argument of this sort will begin from facts about the nature of the emotion itself, and be much informed by the rarified perspective of evolutionary theory. This may be surprising to certain philosophical sensibilities; such facts and evolutionary considerations are often thought to be morally neutral, morally irrelevant, or even corrosive to a moral outlook. I doubt such a view is sustainable, however, especially in this case. At the very least, in calling attention the imperfect fit between this cognitive system that initially evolved to deal with poisons and

parasites, on the one hand, and the social dynamics it was later co-opted to help navigate, on the other, the Entanglement and Co-opt theses is able to expose concerns with things like moral taint and spiritual purity as baseless projections even as it explains their source and prevalence.

References

- Atran, Scott. (2002). *In god we trust: The evolutionary landscape of religion*. New York: Oxford University Press.
- Bandura, A. (1992). Social cognitive theory of social referencing. In S. Feinman (Ed.), *Social referencing and the social construction of reality in infancy* (pp. 175-208). New York: Plenum.
- Barth, F. (Ed.) (1969). *Ethnic groups and boundaries: The social organization of cultural differences*. Boston: Little Brown & Co.
- Bernstein, I. (1999). Taste aversion learning: A contemporary perspective. *Nutrition*, 15(3), 229-234.
- Boyd, Richard. (1991). Realism, anti-foundationalism and the enthusiasm for natural dinds. *Philosophical Studies*, 61, 127-148.
- Boyd, R. & P. J. Richerson. (1985). *Culture and the evolutionary process*. Chicago: University of Chicago Press.
- Boyd R., & P. Richerson. (2005a). *The Origin and Evolution of Cultures*. New York: Oxford University Press.
- R. Boyd & P. Richerson. (2005b). Solving the Puzzle of Human Cooperation, In S. Levinson (ed.), *Evolution and Culture*. Cambridge MA: The MIT Press, pp 105–132.
- Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.
- Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to ‘prejudice.’ *Journal of Personality and Social Psychology*, 88, 770-789.
- Curtis, V., Aunger, R. & Rabie T. (2004). Evidence that disgust evolved to protect from risk of disease. *Proceedings of the Royal Society: Biological Science Series B*, 271(4), S131-S133.
- Daly, M. & Wilson, M. (1988). *Homicide*. Hawthorne, NY: Aldine.
- Darwin, C. (1872). *The expressions of emotions in man & animals* (1st ed.). New York: Philosophical Library.
- Davey, G., Forster, L, & G. Mayhew. (1993). Familial resemblances in disgust sensitivity and animal phobias. *Behavior Research and Therapy*, 31(1), 41-50.
- Doris, J. & Stich, S. (2005). As a matter of fact: Empirical perspectives on ethics. In F. Jackson and M. Smith (Eds.), *The Oxford handbook of contemporary philosophy*. Oxford: Oxford University Press.
- Doris, J. & Stich, S. (2006). Moral psychology: Empirical approaches. *The Stanford Encyclopedia of Philosophy (Summer 2006 Edition)*, Edward N. Zalta (Ed.), URL = <http://plato.stanford.edu/archives/sum2006/entries/moral-psych-emp/>.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169-200.
- Fessler, D. & Haley, K. (2006). Guarding the perimeter: The outside-inside dichotomy in disgust and bodily experience. *Cognition and Emotion*, 20(1), 3-19.
- Fessler, D. & Navarrete, C. (2003). Meat is good to taboo: Dietary proscriptions as a product of the interaction of psychological mechanisms and social processes. *The Journal of Cognition and Culture*, 3(1),1-40.

- Fessler, D. & Navarrete, C. (2004). Third-party attitudes toward sibling incest: Evidence for Westermarck's hypothesis. *Evolution and Human Behavior*, 25, 277-294.
- Frank, R. (1988). *Passions within reason: The strategic role of the emotions*. New York: W.W. Norton & Company.
- Garcia, J., Hankins, W. & K. Rusiniak. (1974). Behavioral regulation of the milieu interne in man and rat. *Science*, 185, 824-831.
- Gil-White, F. (2001). Are ethnic groups biological 'species' to the human brain? *Current Anthropology*, 42 (4), 515-554.
- Griffiths, P. (1997). *What the emotions really are*. Chicago: University of Chicago Press.
- Haidt, J., C. McCauley, and P. Rozin (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences*, 16, 701-713.
- Haidt, J., Rozin, P., McCauley, C. & Imada, S. (1997). Body, psyche, and culture: The relationship between disgust and morality. *Psychology and Developing Societies*, 9, 107-131.
- Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuro-imaging responses to extreme outgroups. *Psychological Science*, 14(10), 847-853.
- Kelly, D. (Forthcoming). *The Nature of Disgust*. Cambridge, MA: The MIT Press.
- Kelly, D. & Stich, S. (2007). Two theories about the cognitive architecture underlying morality. In P. Carruthers, S. Laurence & S. Stich (Eds.), *Innateness and the structure of the mind: Foundations and the future*. (New York: Oxford University Press).
- Kurzban, R. & Leary, M. (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin*. 127(2), 187-208.
- McElreath, R., Boyd, R. & Richerson, P. (2003). Shared norms can lead to the evolution of ethnic markers. *Current Anthropology*, 44(1), 123-29.
- Morris, P., Doe, C. & Godsell, E. (2007). Secondary emotions in non-primate species? Behavioral reports and subjective claims by animal owners. *Cognition and emotion*.
- Nado, J., Kelly, D. & Stich, S. (2009). Moral judgment. In J. Symons & P. Calvo (Eds.) *The Routledge Companion to the Philosophy of Psychology*. New York: Routledge, pp. 621-633.
- Nichols, S. (2002a). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, 84, 221-236.
- Nichols, S. (2002b). On the genealogy of norms: a case for the role of emotion in cultural evolution. *Philosophy of Science*, 69, 234-255.
- Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment*. New York: Oxford University Press.
- Pinker, S. (1997). *How the mind works*. New York: W.W. Norton & Co.
- Richerson, P. & Boyd, R. (1998). The evolution of human ultra-sociality. In I. Eibl-Eibesfeldt and F. Salter (Eds.), *Ideology, warfare, and indoctrinability*. Oxford: Oxford University Press.
- Richerson, P. & Boyd, R. (1999). Complex societies: The evolutionary origins of a crude superorganism. *Human Nature*.

- Richerson, P. & R Boyd. (2001). The Evolution of Subjective Commitment to Groups: A Tribal Instincts Hypothesis. In R. Nesse (ed.), *The Evolution and the Capacity for Commitment*, Russell Sage, pp. 186-220.
- Richerson, P. & Boyd, R. (2005). *Not by genes alone*. Chicago: University of Chicago Press.
- Rozin, P., Fallon, A. & Augustoni-Ziskind, M. (1985). The child's conception of food: The development of contamination sensitivity to "disgusting" substances. *Developmental Psychology*, 21, 1075-79.
- Rozin, P., Haidt, J., & McCauley, C. (2000). Disgust. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions, 2nd Edition*, New York: Guilford Press.
- Rozin, P., Haidt, J., McCauley, C, and Imada, S. (1997). Disgust: Preadaptation and the cultural evolution of a food-based emotion. In H.M. Macbeth (Ed.), *Food preferences and taste: Continuity and change*. Oxford: Berghahn.
- Rozin, P., Hammer, L., Oster, H., Horowitz, T., & Marmora, V. (1986). The child's conception of food: Differentiation of categories of rejected substances in the 16 months to 5 year age range. *Appetite*, 7, 141-151.
- Rozin, P., Lowery, L., & Ebert, R. (1994). Varieties of disgust faces and the structure of disgust. *Journal of Personality and Social Psychology*, 66(5), 870-881.
- 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), 574-586.
- Rozin, P., Millman, L. & Nemeroff C. (1986). Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology*, 50, 703-712.
- Rozin, P., Nemeroff, C., Horowitz, M., Gordon, B. & Voet, W. (1995). The borders of the self: Contamination sensitivity and potency of the body apertures and other body parts. *Journal of Research in Personality*, 29, 318-340.
- Schnall, S., Haidt, J. & Clore, G. (Manuscript). Disgust as embodied moral judgment.
- Shweder, R., Much, N., Mahapatra, M. and Park, L. (1997). The "big three" of morality (autonomy, community, and divinity), and the "big three" explanations of suffering. In Brandt, A. & Rozin, P. (Eds.), *Morality + Health*. New York: Routledge.
- Siegal, M. & Share, D. (1990). Contamination sensitivity in young children. *Developmental Psychology*, 26(3), 455-458.
- Sripada, C. & S. Stich. (2006). A framework for the psychology of norms. In P. Carruthers, S. Laurence & S. Stich (Eds.), *The innate mind: Culture and cognition*. New York: Oxford University Press.
- Ware, J., Jain, K., Burgess, I., & Davey, G. (1994). Disease-avoidance model: Factor analysis of common animal fears. *Behavior Research and Therapy*, 32(1), 57-63.
- Webb, K. & Davey, G. (1993). Disgust sensitivity and fear of animals: Effect of exposure to violent and revulsive material. *Anxiety, Coping and Stress*, 5, 329-335.
- Wheatley, T. & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, 16(10), 780-784.
- Zhong, C. & K. Liljenquist. (2006). Washing away your sins: Threatened morality and physical cleansing. *Science*, 313 (5792), 1451-2.