

UNIVERSITY OF OKLAHOMA GRADUATE COLLEGE

GROSS INJUSTICES: A QUALIFIED DEFENSE OF MORAL DISGUST IN CRIMINAL
LAW

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

By LUKE TUCKER

Norman, Oklahoma

2025

GROSS INJUSTICES: A QUALIFIED DEFENSE OF MORAL DISGUST IN CRIMINAL
LAW

A DISSERTATION APPROVED FOR THE
SHYAM DEV PATWARDHAN DEPARTMENT OF PHILOSOPHY

BY THE COMMITTEE CONSISTING OF

Dr. Zev Trachtenberg, Chair

Dr. Wayne Riggs

Dr. Stephen Ellis

Dr. Kyle Harper

© Copyright by Luke Tucker 2025

All Rights Reserved.

Acknowledgments

Above all, I wish to thank my advisor, Zev Trachtenberg, for his wisdom and guidance over the years. I am also grateful to my additional committee members, Wayne Riggs, Steve Ellis, and Kyle Harper, for all of their helpful input and advice.

Next, I wish to thank my wife, Shelley Tucker, for two things: first, for her help in creating the original illustrations and photographs found in this dissertation, and second, for her constant love and support in general.

Finally, I thank my parents, Donnie and Sandra Tucker, for always encouraging me.

Table of Contents

<i>Acknowledgments</i>	<i>iv</i>
<i>Abstract</i>	<i>vii</i>
Introduction.....	1
1. Moral Disgust and Moral Dumbfounding	1
2. When Disgust Becomes Law.....	3
3. Debating Disgust in Law	8
4. Defending Disgust in Law	16
Chapter One	21
1. Introduction	21
2. The Presumption of Liberty.....	25
3. Legal Moralism.....	32
4. Moral Dumbfounding	37
5. The Ethics of Incest	41
6. Conclusion	52
Chapter Two.....	56
1. Introduction	56
2. What is Disgust?.....	57
3. The Origins of Disgust	69
4. Beyond Biology	78
5. Disgust Acquisition	94
6. The Role of Culture in Disgust Acquisition	100
7. Conclusion	109
Chapter Three.....	113
1. Introduction	113
2. Ontological Skepticism.....	117
3. The Classic/New Advocacy Distinction.....	124
4. New Advocacy and the Metaphor Challenge	141
5. Classic Advocacy and the Coincidence Challenge: Defending Kekes.....	156
6. Classic Advocacy and the Coincidence Challenge: Harmless Purity Violations	168
7. Conclusion	181
Chapter Four	185
1. Introduction	185

2.	Does Disgust Moralize?.....	189
3.	Does Disgust Dehumanize?	198
4.	Kelly’s Co-opt Thesis	202
5.	The Appeal to Disgust	216
6.	Why Dumbfounding?	234
7.	Vindicating the Inaccessible Premise	244
8.	Vindicating the Wisdom Premise	256
9.	Conclusion	268
References		272

Abstract

This dissertation offers a limited defense of appeals to moral disgust in criminal law. Disgust already influences law in a number of ways, including serving as the primary justification for prohibitions on certain conduct. Whether it should have this influence, however, is contested. Some authors believe feelings of disgust in normative discourse are distracting, or downright misleading. Consequently, the emotion should not influence law, or any public institution. I disagree. The three main skeptical challenges raised against disgust—that it has troubling evolutionary origins, is an “unreasoned” emotion, and its status as a moral emotion, strictly speaking, is discredited by empirical research—don’t withstand scrutiny. More broadly, I argue that moral disgust’s influence on law is both positive and indispensable. In many cases, no other justificatory resource is available to ground laws most citizens consider common sense.

Chapter One sets up discussion by noting that many criminal laws are already based in disgust, and repealing them would run counter to deep moral intuitions. I focus specifically on bans on disgusting, albeit arguably harmless, taboos, such as consensual cannibalism. It’s generally assumed the state needs a good reason to ban something. With harmless taboos, providing a reason is surprisingly hard to do. They don’t cause harm or violate anyone’s rights. It’s even unclear how they threaten “public morals.” This generates a dilemma: either (i) disgust, by itself, can sometimes justify laws, or (ii) bans on harmless taboos are unjustified. Because the latter option is unpalatable, subsequent chapters pursue a case for the former.

Chapter Two sketches a descriptive account of disgust. I cover evidence that disgust is a “basic” emotion with universally recognized expressions and several cross-culturally consistent triggers. It originally evolved as a biological defense against unsafe foods and transmissible sicknesses but, later, was recruited by cultures to serve socio-moral functions. I describe how disgust develops in childhood, how individuals acquire specific disgust triggers, and how culture influences these processes. Moral Foundations Theory is used to explain how cross-cultural

variations in what is considered “morally disgusting” track broader, cultural differences in moral beliefs and practices.

Chapter Three defends the view that “moral disgust” is a genuine emotional state against recent psychological and philosophical challenges. One challenge holds that the common extension of disgust language to immoral behavior is metaphorical and doesn’t literally convey the speaker’s inner state. Another challenge holds that disgust and moral disapproval often co-occur, but this is merely coincidental, not evidence of a distinct, unified emotional response. I argue that neither challenge succeeds. First, I draw on empirical evidence suggesting that, when people report feeling disgusted by immorality, their physiological and neural responses often support a literal interpretation. Second, I argue that the claim “moral disgust is merely a co-occurrence of two unrelated states” relies on controversial metaethical assumptions we should reject on philosophical grounds.

Chapter Four defends the so-called “appeal to disgust” in ethical and policy debates. The appeal cites the fact that many people feel deep disgust toward a practice as evidence that it must be seriously wrong. Critics respond by arguing that disgust, rather than offering guidance, is inherently unreliable or problematic. I rebut these types of arguments. I then argue moral disgust can be “wise.” When widespread, it provides defeasible justification for laws, even before any substantive discussion. This is because our moral disgust reactions are often underpinned by reasons that we lack introspective access to—but reasons that we *would* endorse, if we *could* access them. For support, I draw on gene-culture co-evolutionary theory, which explains how culturally transmitted norms often become opaque to their inheritors. I also examine the case of polygamy, a practice that Western lawmakers banned—guided partly by disgust—long before modern social science revealed its social harms.

Introduction

1. Moral Disgust and Moral Dumbfounding

In 2016, the Scottish public petitions committee considered a petition to legalize incest between consenting adults. It was rejected in less than a minute. The committee’s reasoning: the petition served “no public interest” and was “abhorrent” (quoted in McHugh et al., 2023: 1043). There was no discussion of the substantive issues raised by the petition—such as arguments that anti-incest laws are discriminatory, violate individual rights, or are incompatible with the European Convention on Human Rights. Evidently, the committee members’ instinctive negative reaction, likely rooted in a sense of deep disgust toward incest¹, was sufficient, in their minds, to justify rejecting the petition. Reflecting on this, moral psychologist Cillian McHugh and his colleagues write:

By not addressing these issues, the politicians seemingly maintained a position without providing justification for this position. If an individual member of the committee was questioned on the substantive issues, it is not clear that they would have been able to provide justified reasons to support their judgment, and in some cases may appeal [sic] to emotion [e.g. abhorrence] as justification... If pressed, it is possible that a committee member may have even demonstrated a phenomenon in moral psychology known as “moral dumbfounding” (2023: 1043).

McHugh has a point. Research in moral psychology has found that many people (i) strongly disapprove of incest (even with consent) yet (ii) become dumbfounded when asked to provide reasons supporting their judgment (McHugh et al., 2023; McHugh et al., 2020; 2017; Haidt, 2001; Haidt and Hersh, 2001; Haidt et al., 2000; Haidt et al., 1993). Interestingly, even when subjects acknowledge they are dumbfounded, they often don’t feel abashed and keep maintaining their initial judgment.

¹ Incest causes many people to experience “oral inhibition”—gagging, nausea, and loss of appetite (Royzman et al. 2008). These are unequivocal signs of disgust, not anger or any other state. More on this below.

Incest isn't the only practice that elicits moral dumbfounding. A suite of other practices widely viewed as deeply morally disgusting (albeit arguably harmless) also triggers the effect, including bestiality (Haidt et al., 2000), necrophilia (Haidt and Hersh, 2001), cannibalism (Haidt et al., 2000), and marketizing human body parts (Tetlock et al., 2000). This recurrent link between dumbfounding and moral disgust has led some researchers to call moral disgust an “unreasoned emotion.” This reputation is not helped by other findings in moral psychology, such as that moral disgust is less sensitive to intentions, agency, and mitigating circumstances (Russell and Giner-Sorolla, 2011a; Young and Saxe, 2011) and can be resistant to reappraisal (i.e. changing your mind) (Feinberg et al., 2014).

Given these “unreasoning” features of moral disgust, one might reasonably wonder: Is this emotion really a suitable candidate to influence, even determine, policymaking in a liberal democracy, such as Scotland—or anywhere else? Surely, liberal democracies have grown beyond such pre-Enlightenment gaffes as the “divine right of kings,” alchemy and astrology, and the majoritarian pretension that any lifestyle can be discriminated against as long as most people happen to find it distasteful. Laws, especially those that infringe upon citizens’ freedoms, must be undergirded by articulate reasons.

Actually, in this dissertation, I want to defend the view that moral disgust is a good candidate to influence law in some contexts. Specifically, I will motivate the idea that moral disgust, despite its tendency to cause dumbfounding, can be a “wise” emotional reaction; when widespread among citizens, it can provide *prima facie* justification for laws in a liberal society, even before any rational, substantive arguments for and against are heard. This is because there are interesting reasons to think that our moral disgust reactions are often underpinned by reasons

that we may not have immediate introspective access to—but reasons that we *would* endorse, if we *could* access them.

In other words, the dissertation ultimately aims to defend the Scottish committee’s swift dismissal of the petition to legalize incest, and similar ways in which disgust influences lawmaking. Making this case will require several subsidiary moves, including giving an account of moral disgust, explaining why dumbfounding and disgust are frequently associated, and arguing that disgust’s current role in law is indispensable. Let me give some broad context for the project below.

2. When Disgust Becomes Law

The focus of this dissertation is whether disgust should influence criminal law. By “disgust” I mean the revulsion one feels in response to offensive or unpleasant stimuli—stuff that gets an involuntary “Yuck!” or “Eww!” out of us. The maggots in the dog food. The fingernail clippings your roommate left on a paper towel in the bathroom. The spray of mist on your neck from the stranger behind you on the subway who sneezed. The graphic depictions of surgery in nearly every episode of *Grey’s Anatomy*.

The *Oxford English Dictionary* defines disgust as “Strong repugnance, aversion, or repulsion excited by that which is loathsome or offensive, as a foul smell, disagreeable person or action”—a definition that traces to 1598. Charles Darwin’s *The Expression of the Emotions in Man and Animals* (1876) defined it as “something revolting, primarily in relation to the sense of taste, as actually perceived or vividly imagined; and secondarily to anything which causes a similar feeling, through the sense of smell, touch, and even of eyesight” (p. 254). It is an emotion we experience almost daily, whose commands—“Don’t eat that! Don’t look at that! Get away from that!”—we regularly and reflexively obey.

Some emotion researchers classify disgust as a “basic” emotion, meaning it is fundamentally distinguishable from other emotional states by its characteristic inputs (things that trigger it) and outputs (ways it is expressed) (see for example Ekman and Cordaro, 2011; Rozin et al., 2008; 1999; Haidt et al., 1994; Rozin and Fallon, 1987). In virtue of being basic, disgust is universal in many respects. Certain items almost universally trigger the emotion, such as feces, bugs, rats, spit, sweat, sexual fluids, corpses, rotting meat, and almost anything that is moist, sticky, or gooey (Curtis and Biran, 2001). People everywhere—from the valley girl of suburban Los Angeles to the !Kung hunter-gatherer of the Kalahari Desert—express disgust with the same “gape face”: furrowed brow, wrinkled nose, retracted upper lip (Ekman, 1972) (see Figure 1).



Figure 1. Man making the gape face. Illustration by Shelley Tucker.

While disgust is often triggered by items that would be physically unpleasant to eat or touch, it is not solely concerned with tactile properties. Interestingly, it is also responsive to *moral* stimuli. People across the globe, if you ask what disgusts them, will name bad actions and bad people (Chapman and Anderson, 2013; Tybur et al., 2013; Rozin, et al., 2008; 1999; Bloom, 2004; Curtis and Biran, 2001; Haidt, McCauley, and Rozin, 1997; Imada, Yamada, and Haidt,

1993). People in India express disgust towards “betrayal”; people in the Netherlands name “politicians” and “insulting behavior”; and people in the United Kingdom name “rude people,” “man beating a woman,” “cruelty to a horse,” and “wounding of an old lady” (Curtis and Biran, 2001). Many people who receive unfair offers in economic games exhibit increased activity in brain regions associated with disgust (namely, the anterior insula) (Sanfey et al., 2003). Many people report that they would be averse to receiving a blood transfusion from a child molester (Rozin et al., 1999).

In ordinary speech, the extension of disgust language to morally problematic practices and people is ubiquitous. Traitors are “rats”; freeloaders are “parasites”; greedy businessmen are “pigs at the trough”; opportunistic lawyers are “vultures”; untrue or misleading talk is “bullshit” and habitual purveyors of it are “full of shit”; racists propagate “bile”; and liars “reek” of dishonesty. Calling immorality disgusting is not a quirk of English; it happens in many other languages, including Japanese, Bengali, Russian, and Hebrew, and may be universal (Rozin et al., 2008: 763). It is also not a quirk of adults; children, starting relatively young, can be observed extending disgust language to bad behavior (Aznar et al., 2023; Stevenson et al., 2010; Danovitch and Bloom, 2009). Nor is calling immorality disgusting a new phenomenon. Oedipus says that accidentally violating taboos against patricide and incest has made him “vile” and “utterly unclean.”² Euthyphro believes his father, by committing murder, has “polluted” their household.³ Jesus said that hypocrites on the inside are “full of the bones of the dead and everything unclean.”⁴ And the Ghost in *Hamlet* said that fratricide is “most foul, strange, and unnatural.”⁵

² From *Oedipus Rex* by Sophocles: <https://classics.mit.edu/Sophocles/oedipus.html>

³ From *Euthyphro* by Plato: <https://classics.mit.edu/Sophocles/oedipus.html>

⁴ Matthew 23:27.

⁵ Act I, Scene V: <https://shakespeare.mit.edu/hamlet/full.html>

Because disgust responds to moral stimuli, it may come as no surprise that the emotion has had a strong influence on legal institutions in many countries. In the United States, for example, disgust pervades law. It can serve as a mitigating factor in murder cases⁶ or as an aggravating one⁷ (Nussbaum, 2004). It is central to legal definitions of obscenity (Kahan, 2000). It is central to many nuisance laws, such as those against public nudity, sex, and urination (Nussbaum, 2004). State courts can refuse to return seized property to criminals if doing so would spark disgust in the general public (*Beldotti v. Commonwealth*, 41 Mass. App. Ct. 185).

The fact that an act is widely considered morally disgusting often serves as the primary justification for prohibitions on that act. For example, several taboo-breaking behaviors are widely prohibited—even when harmless and conducted in private with the full consent of those directly involved—seemingly just for being disgusting. For example, incest is illegal in 49 states and a felony in 46, even with consent, and, oftentimes, even if the participants are not genetically related (i.e. there is no increased risk of genetic disorders in offspring) (Posner and Silbaugh, 1996: 207-212). Other taboos that are widely prohibited, even when consensual and harmless, include cannibalism⁸, bestiality (Posner and Silbaugh, 1996: 217-212), necrophilia (*ibid.*, pp. 213-216), and corpse-desecration⁹.

⁶ I.e. perpetrator responsibility can be diluted if their actions were provoked by a bout of overwhelming disgust.

⁷ I.e. the severity of punishment can be greater if a judge or jury finds the murder particularly disgusting or revolting.

⁸ Cannibalism, in itself and as such, is only explicitly criminalized in one state: Idaho (Jarvis, 2024). However, there are legal obstacles to cannibalism in other states, such as federal laws against organ sales and state laws against corpse-desecration. Aside from explicit criminalization, there are several examples from across the U.S. of cannibalism being treated as an aggravating factor in murder cases (i.e. a reason to punish the murderer more harshly) (Jarvis, 2024).

⁹ See, for example, Nussbaum (2004), who explains that “most states have some laws against desecration of the corpse” (p. 154). See also Patrick and Lieberman (2018a), who explain that “in the vast majority of countries and cultures, for most of recorded history, acts like...consensual incest and the desecration of a corpse...have been socially taboo and likewise illegal, despite the fact that they often do no harm to either the individuals involved in the acts or, for that matter, anyone else” (p. 5).

Another class of behaviors that frequently draw lawmakers' attention involves "repugnant markets"—transactions that elicit moral revulsion in the public, even when the parties directly involved benefit from, and explicitly consent to, the trade (Roth, 2007). For example, selling human organs is a felony under federal law, carrying a maximum penalty of five years in prison and a \$50,000 fine (National Organ Transplant Act, 1984), despite the fact that some live organs are high in demand and can be harmlessly donated. The "main barrier" to legalizing such markets is "the widespread view that kidney markets are repulsive, disgusting, and degrading" (Brennan and Jaworski, 2015b). Additional repugnant markets include prostitution, virginity auctions, surrogacy, adoption, vote buying, bribery, horse and dog meat, indentured servitude, and prediction markets in death, assassination, or government overthrows. These markets are seen by many as

inherently gross and profane and as intruding onto the sacred. To put certain things on the market is much like washing hamburger grease off your fingers with holy water or blowing one's nose with pages from the Bible (Brennan and Jaworski, 2015a: 1077).

A third area of disgust-influenced regulation is biomedicine and biotechnology. In 1997, when the Roslin Institute in Scotland successfully cloned Dolly the sheep via somatic cell nuclear transfer, concerns arose that similar techniques could be used to clone humans. Bioethicist Leon Kass responded by calling for bans on human cloning and all research related to it. The crux of his argument: cloning humans is "repugnant"—much like "father-daughter incest...having sex with animals...mutilating a corpse, or eating human flesh" (1997: 687). He added: repugnance expresses "deep wisdom, beyond reason's power to fully articulate it" (p. 687). Then-U.S. president Bill Clinton proposed federal legislation to comprehensively ban cloning humans, stating that it has potential to "threaten the sacred family bonds" at the core of

society.¹⁰ Numerous other countries followed suit. Other emerging biomedical technologies and practices have also, at one time, been the subject of vigorous, disgust-language-laden policy debates, including gene-editing (e.g. Bostrom, 2005; Sandel, 2004), creating new animal species via xenotransplantation (e.g. Mills, 2014; Younger, 2014; Midgley, 2000), enhancement drugs (drugs that enhance natural human abilities like memory, intelligence, or athletic ability) (e.g. Bostrom and Roache, 2008; Kass, 2003), and using cadavers in automotive crash testing.¹¹

3. Debating Disgust in Law

Undeniably, disgust influences law in many countries and in an array of ways. However, whether this *ought* to be the case is a question. It may feel natural to assume that lawmakers and judges are permitted, in some contexts, to treat citizens' deep feelings of disgust as a potential source of justification for laws and legal decisions. But there are important limits on what lawmakers may legitimately prescribe and prohibit, at least in liberal democracies, where norms and values such as liberty, privacy, and state neutrality reign. For example, the First Amendment of the United States Constitution guarantees that "Congress shall make no law" prohibiting the free exercise of religion. Thus, if most citizens happened to find (e.g.) Islam disgusting, obviously that would not justify Congress in outlawing Islam. Citizens in pluralistic democracies hold different and incompatible comprehensive doctrines¹²—sets of views about God and life, right and wrong, good and bad. Which doctrine is right or best is a question that admits of reasonable disagreement. The mandate of the state, says the liberal, is not to force one doctrine

¹⁰ <https://www.science.org/content/article/president-proposes-human-cloning-ban>

¹¹ The foregoing examples are just the explicit functions of disgust in law. Some scholars suggest it is likely having an underappreciated influence on public discourse surrounding other issues, including abortion, suicide, and euthanasia (Lockhart et al., 2023; Cahill, 2013).

¹² John Rawls introduces this term in his 1993 *Political Liberalism*. He writes that comprehensive doctrines form a conception of "what is of value in life, the ideals of personal character, as well as ideals of friendship and of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole" (p. 13).

or another on its citizens—it is to secure citizens’ freedom to develop their own ways of life, according to the doctrine *they* endorse. Of course, the state can come down against some ways of life, such as that of the Mafioso or Nazi, but only insofar as these lifeways threaten *other* citizens’ free pursuit of their preferred lifeways. If disgust is to have a place in law, it must (on the current view) somehow serve the liberal project. However, some authors doubt that it can do this.

Debate about what role—if any—disgust should play in law, especially criminal law, goes back decades (see for example Bollard, 2022; Lieberman and Patrick, 2018a; 2018b; Gert, 2015; Kelly and Morar, 2014; Nussbaum, 2004; Kahan, 2000; Kass, 1997; Miller, 1997; Kekes, 1992; Hart, 1963; Devlin, 1959). While disgust’s proper role in law is a narrow issue, most of the recent discussion in this area has played out in the context of a larger debate about the emotion’s overall normative status. The participants in this debate are the so-called “disgust advocates”—who attribute one or more forms of normative significance to disgust—and “disgust skeptics”—who deny or downplay its normative significance.¹³

The disgust advocate/skeptic debate first arose in the 1990s, blossomed in the early 2000s, and continues to the present. Two early and influential figures in the debate include already-mentioned bioethicist Leon Kass and legal scholar William Ian Miller. Kass’ contribution to the debate came in 1997 in the form of his now-infamous paper, “The Wisdom of Repugnance,” in which he argued for banning human cloning on the grounds of its being repugnant. That same year, Miller published his book *The Anatomy of Disgust*, which provides a general analysis of the emotion and “the important role it plays in organizing and internalizing many of our attitudes towards the moral, social, and political domains” (1997: 18).

¹³ This terminology originates in Kelly (2011) and got picked up by Kelly and Morar (2014), Fischer (2016), Lieberman and Patrick (2018), Plakias (2018; 2013), and Bollard (2022).

Kass and Miller were working in different contexts and were, as far as I can tell, unaware of each other. Nevertheless, their combined contributions quickly garnered interest across bioethics, philosophy of law, and social and political philosophy. Fellow legal scholar Dan Kahan (2000; 1998) used Miller's account of disgust to argue that the emotion plays an indispensable role in criminal law. Further work in bioethics played up the "wisdom" inherent in the public's instinctive disgust not only towards cloning but also towards other emerging biotechnologies that contravene "natural" boundaries or longstanding customs (see for example Levin, 2007; Sandel, 2004; Midgley, 2000).

This push to see disgust as normatively significant was eventually met with an equal push in the opposite direction. Martha Nussbaum (2004a; 2004b; 1999), responding directly to Kass, Kahan, and Miller, argued against citing disgust as a source of moral insight or wisdom in any context, bioethical or legal, on the grounds that an examination of the emotion's cognitive content reveals it to be inherently problematic. Likewise, Daniel Kelly (2014; 2011) has argued that disgust, in virtue of its evolutionary history, is inherently dehumanizing. As such, "there are no defensible uses for disgust in legal or political institutions" (2014: 154) and "any society free of social disgust would be more just, reasonable and compassionate" (p. 153).

At the same time that these normative debates were unfolding, interest in disgust in the hard and soft sciences was surging. Prior to the early 1990s, little empirical research on disgust existed (Strohming, 2014; Rozin, et al., 2008; Olatunji and Sawchuk, 2005). Indeed, political philosopher John Kekes, in a 1992 paper on moral disgust and taboos, acknowledges this lacuna in the very first paragraph, writing that disgust has "not been much discussed in philosophical literature or, indeed, anywhere else" (p. 431). He also notes that, in 1992, there were no entries on disgust in *The Encyclopedia of Social Sciences*, *The Encyclopedia of Philosophy*, or *The*

Dictionary of the History of Ideas. Because of the lack of availability of empirical research on disgust, early contributions to the disgust advocate/skeptic debate mostly featured philosophical, literary, and historical analyses of the emotion. For example, Kekes analyzed the theme of pollution in the ancient Greek, Jewish, and Christian moral traditions. Similarly, Mary Midgley (2000) explored “hubris” and “traditional mixed monsters” (e.g. minotaurs, chimeras) in literature. Kass (1997) and Miller (1997) followed similar patterns. Later in the 1990s, however, empirical research on disgust exploded, with figures like psychologist Paul Rozin leading the charge. It is now a flourishing field and has enjoyed contributions from an array of scholars, including psychologists, evolutionary biologists, cognitive scientists, neuroscientists, cultural anthropologists, and behavioral economists.

The ready availability of empirical research on disgust has had the welcome effect, in recent years, of allowing philosophers and legal scholars to avail themselves of this research and ensure that philosophical discussion remains empirically informed. For example, Nussbaum’s 2004 book *Hiding from Humanity: Disgust, Shame, and the Law* devotes a 52-page chapter to sketching an account of disgust that resonates with research by Paul Rozin and others. Likewise, the account of the emotion in Kelly’s 2011 book *Yuck! The Nature and Moral Significance of Disgust* was praised by one reviewer for “working in partnership with cognitive science, evolutionary psychology and neurophysiology” (Blackburn, 2011: 52). Recently, disgust advocates have also drawn on scientific research to support their arguments. For example, Alexandra Plakias (2018; 2013) draws on work in evolutionary psychology to support her claim that the behavioral tendency associated with disgust (i.e. avoidance or withdrawal) makes it the proper response to contaminating social behaviors—those that pollute beneficial social

interactions by eroding shared expectations of cooperation and are liable to spread if not contained (cf. Kumar, 2017; Hauskeller, 2006).

In the current literature, the disgust advocate/skeptic debate spans at least four different questions¹⁴, which I shall label in the following way (see Table 1):

<i>Moral question</i>	Is disgust is a fitting or appropriate emotional response to any moral wrongs?
<i>Epistemic question</i>	Does disgust, in and of itself, hold any epistemic (or justificatory, or evidential) value?
<i>Political question</i>	What role, if any, should disgust play in public debate, deliberation, explicit social norms, or formalized institutions (e.g. criminal law)?
<i>Ontological question</i>	Is “moral disgust” a genuine emotional state, distinguishable from (1) other states (e.g. anger) as well as (2) physical disgust (e.g. the sort triggered by feces)?

Table 1. Four questions covered in the disgust advocate/skeptic debate.

(1) *The Moral question.* Is disgust a fitting or appropriate emotional response to any moral wrongs? Some authors believe so. For example, as noted above, Alexandra Plakias (2018; 2013) argues that the motivational tendency associated with disgust (withdrawal) makes it an *appropriate* response to contaminating social behaviors. Meanwhile, skeptics believe that disgust is always an *inappropriate* response to immorality. Daniel Kelly and Nicolae Morar (2014) argue that disgust mentally construes its object as dirt, which in the case of immorality is *another*

¹⁴ Bollard (2022) and Kelly and Morar (2014) also note that the disgust advocate/skeptic debate covers a few different questions, which we should not conflate.

person. But all people, even bad ones, ought to be seen as having rights and deserving of moral regard. So, we should resist being disgusted by bad actors.

(2) *The Epistemic question*. Does disgust, in and of itself, hold any epistemic (or justificatory, or evidential) value? Should feelings of disgust be taken to help justify the judgments that they influence? A spectrum of views on disgust's epistemic significance can be found in the literature. Some believe that, at least in certain cases, disgust holds at least some epistemic value (see Fischer, 2016; Midgley, 2000; Kahan, 2000; Kass, 1997; Kekes, 1992). For example, Leon Kass takes the view that feelings of deep disgust towards a practice (e.g. human cloning), when widespread among the public, provide a strong indication that the practice must be seriously wrong, before any arguments for or against have been heard. Meanwhile, disgust skeptics deny that disgust holds epistemic significance. Some skeptics take the strong position that disgust is inherently *unwise*; like any other cognitive bias, its influence on moral reasoning is distortive (see for example Nussbaum, 2004a). Other skeptics take the more moderate view that disgust is neither wise nor unwise. Perhaps it can be an appropriate emotional response to certain wrongs, but only if we have independent evidence that our disgust at those wrongs is properly directed. We cannot cite the mere fact that the emotion was triggered as evidence that it ought to have been triggered.

(3) *The Political question*. What role (if any) should disgust play in public debate, explicit social norms, or formalized institutions, such as criminal law? Should it be permitted to do any justificatory work in specific legal or policy debates? Some authors, such as Dan Kahan (2000), believe disgust plays an indispensable role in criminal law, such that disgust ceasing to have an influence would risk making the law "morally blind" (p. 65). Similarly, John Kekes (1992) argues that disgust is an important political emotion and the proper response to moral

offenses that threaten society's most basic and important conventions. Meanwhile, skeptic Martha Nussbaum rejects such claims; in her view, disgust has a pernicious influence on the operations of law. She does not oppose letting emotions influence the law *per se*; for instance, it is important for the law to have a doctrine of "reasonable provocation," which allows us to hold a person less responsible (i.e. punish them less severely) for a crime that they committed while in the grips of a *reasonable* bout of passion, anger, fear, or nervousness. The problem with disgust, in her view, is that it is *never reasonable*. An examination of its cognitive content reveals that it is rooted in an unconscious desire to see ourselves as somehow higher than animals—a self-delusion which, in order to maintain, requires us to dehumanize marginalized groups. In other words, the emotion is problematic not in certain cases but *by its very nature*. As such, it "should never be the primary basis for rendering an act criminal" (Nussbaum, 2004a: 14).

(4) *The Ontological question*. Is disgust really responsive to moral stimuli? That is, is "moral disgust" a genuine emotional state at all, distinguishable from other emotions (e.g. anger) as well as merely physical disgust? Many scholars tend to assume so. However, recently, a number of disgust skeptics have begun to challenge this assumption (see Bollard, 2022; Fileva, 2021; Giubilini, 2016; Gert, 2015). One challenge is that the widely observed phenomenon of extending disgust language to moral offenses is merely an instance of metaphorical speech. For example, when someone says "Your hypocrisy disgusts me!" the reality is that they are angry, but are coding their anger as disgust for rhetorical purposes. A second challenge is that disgust and moral disapproval sometimes co-occur, but this is a mere coincidence (e.g. the disgust one feels towards a bloody murder is directed at the blood, not the murder *per se*). A third challenge is that disgust is not a moral emotion but rather a *moralizing* emotion. Physical disgustingness (not itself a moral property) can sometimes induce a moral judgment where one does not belong.

For example, cannibalism is widely thought immoral, and this may be *because* it is widely considered physically gross (Giubilini, 2016).

Disgust advocacy and skepticism are not monoliths. Some disgust advocates, like Kass (1997), attribute all three types of normative significance to the emotion. Meanwhile others, like Miller (1997), argue for moral and political but not necessarily epistemic significance. Two different advocates may attribute the same type of normative significance to disgust but on the basis of different arguments. Also, as I discuss in later chapters, different advocates have slightly different accounts of moral disgust, and disagree about what types of moral transgressions typically trigger the emotion. For example, some advocates suggest that moral disgust is a response to violations of “purity” (acts that violate norms related to sanctity or cleanliness, but don’t cause direct harm to anyone), while other advocates suggest it is a response to “anti-social violations” (acts that damage beneficial social relations) (e.g. Plakias, 2018; 2013; Kumar, 2017). Nevertheless, despite their differences, disgust advocates present a more-or-less unified front in the literature. Advocates who attribute one form of normative significance to disgust tend to attribute others. Philosophically, this makes sense. For example, if disgust is a reliable signal of serious wrongness (epistemic significance), then arguably it can justify criminal prohibitions all on its own (political significance).

This unity-in-diversity is paralleled on the other side. On the one hand, different disgust skeptics favor different accounts of moral disgust. For example, Nussbaum’s account of disgust is heavily influenced by Paul Rozin’s work, particularly his taxonomy of distinct disgust domains. She accepts, with minor alteration, Rozin’s distinction between “core disgust”—a response to likely disease vectors, like feces and saliva—and “animal-reminder disgust”—a response to things that remind us that we, as human beings, are mortal and deeply vulnerable,

such as images of bloody car wrecks and exposed viscera. Meanwhile, other skeptics, such as Mara Bollard (2022) and Joshua Gert (2015), reject Rozin’s notion of animal-reminder disgust wholesale. They argue instead that alleged “animal reminders,” like exposed viscera, repel simply on the basis of their being likely disease vectors. Another point of dispute within disgust skepticism is over the ontological status of “moral disgust.” Some skeptics, such as Mara Bollard (2022), question its ontological status, while others, such as Daniel Kelly (2011), do not. On the other hand, skeptics present a fairly unified front on the big normative questions. They generally refuse to attribute any sort of normative significance to disgust, arguing (e.g.) that its epistemic or moral dubiousness undermines its political significance.

Throughout this dissertation, I will use the term “disgust advocate” broadly to refer to friends of my thesis—that is, those who agree that disgust is both normatively significant and also plays an important role in criminal law—while “disgust skeptic” refers to my philosophical opponents. I don’t think this will be confusing, even though there will surely be times when not all disgust advocates are likely to agree with me, and not all disgust skeptics are likely to disagree.

4. Defending Disgust in Law

The terrain of the disgust advocate/skeptic debate is broad. However, this dissertation broaches a narrow issue at the intersection of the Epistemic and Political questions: whether disgust holds justificatory value (the Epistemic question), and if it does, what this means for its role in law (the Political question). As noted above, currently, several laws and legal decisions (in the U.S. and elsewhere) are justified primarily on the basis of widespread feelings of disgust among citizens. For instance, this is true of many prohibitions on harmless taboos, repugnant markets, and certain biotechnologies. Of course, as has been noted, disgust is frequently

associated with moral dumbfounding; despite our strong feelings of revulsion toward certain practices, it can be quite hard to explain *why* revulsion is warranted. It follows from this that the legitimacy of disgust-based laws stands or falls depending on the justificatory status of disgust itself (the Epistemic question). If it *can* provide justification, then these laws are justified, at least partly. If *not*, these laws are unjustified (unless we can justify them on some other basis, which seems unlikely).

The view I will ultimately defend is this. Under some conditions, disgust should indeed play a justificatory role in debates about law and public policy. Specifically, when an overlapping consensus of citizens with diverse comprehensive doctrines widely shares a deep sense of moral disgust towards a practice, X, this constitutes defeasible evidence that X is seriously wrong and, at least provisionally, ought to be prohibited. In other words, even before any substantive arguments for or against have been heard, widespread disgust provides an initial, though not a final or dispositive, reason for prohibition. This does not mean that discussion of X is off-limits. On the contrary, vigorous debate ought to be carried out. However, starting with a presumption on the side of the majority's disgust—shifting the burden of argument to those who challenge or question it—will improve the *procedure* of public debate.

The upshot of this argument is that certain ways in which disgust currently influences lawmaking—such as the Scottish public petitions committee's dismissal of the petition to legalize incest, and prohibitions in the U.S. on certain harmless taboos—are *prima facie* justified.

While this is the dissertation's overarching goal, several subsidiary moves need to be made to get there. Here is how the argument of the dissertation unfolds. Chapter One sets up discussion by highlighting a point rarely acknowledged by disgust skeptics: *several criminal laws are already based in disgust, and it would be incredibly counterintuitive to repeal these*

laws. I draw out this point by discussing a tension between basic liberal principles and existing bans on disgusting, albeit harmless, taboos. Liberals generally hold the state needs a good reason to ban something. But with harmless taboos, it is surprisingly hard to give a reason. They don't cause harm or violate anyone's rights. It's even unclear how they threaten "public morals." Psychological research on moral dumbfounding highlights a general trend that people often can't explain why they think harmless taboos are wrong—they just *feel* that they are deeply disgusting. This generates a dilemma: either (i) disgust can sometimes justify criminal laws, or (ii) we must admit that bans on harmless taboos are unjustified. Because the latter option is deeply unpalatable, subsequent chapters build a case for the former.

Chapter Two reviews the scientific literature germane to disgust and sketches a descriptive account of the emotion. Obviously, I'm concerned in this dissertation with normative issues, and in normative discussions, it's important to keep our *is*'s separate from our *ought*'s. Nevertheless, in the disgust advocate/skeptic debate, participants on all sides typically assume one descriptive account of disgust or another. This is because the subject of the debate—disgust—is a real psychological and emotional phenomenon with an objective structure. We should not simply stipulate what we mean by "disgust"; if we're going to talk about the emotion, we should know a thing or two about its descriptive character. To that end, I chase down several important threads in the current science. I cover evidence that disgust is a "basic" emotion, with universally recognized expressions and several cross-culturally consistent triggers. It originally evolved as a biological defense against unsafe foods and transmissible sicknesses, but later was recruited by cultures to serve socio-moral functions. I describe how disgust develops in childhood, how individuals acquire specific disgust triggers, and how culture influences these processes. I also use Moral Foundations Theory to explain how cross-cultural variations in what

is considered “morally disgusting” track broader, cultural differences in moral beliefs and practices. The descriptive account of disgust laid out in this chapter serves to ground the arguments of the two subsequent chapters.

Chapter Three delves into the recent swell of work on the Ontological question. Specifically, I defend the view that “moral disgust is a genuine emotional state” against recent psychological and philosophical challenges. One challenge holds that the common extension of disgust language to immoral behavior is merely metaphorical and does not literally convey the speaker’s inner state. Another challenge holds that, while disgust and moral disapproval often co-occur, this is merely coincidental and not evidence of a distinct, unified emotional response. I argue that neither challenge succeeds. First, I draw on empirical evidence suggesting that, when people *report* feeling disgusted by immorality, their physiological and neural responses often support a literal interpretation. Second, I argue that the claim “moral disgust is merely a co-occurrence of two unrelated states” relies on controversial metaethical assumptions we should reject on philosophical grounds. Refuting ontological skepticism about moral disgust in this chapter heads off a major obstacle to the argument of the next chapter.

Chapter Four—the central argument of the dissertation—defends the “Appeal to Disgust,” an informal argumentative move that often appears in debates across ethics, political philosophy, and other areas. The Appeal cites the fact that many people feel strong moral disgust toward a practice as evidence (to some degree) that it must be seriously wrong. Critics disparage the Appeal as anti-reason and propose several arguments that disgust, rather than offering guidance, is inherently unreliable or misleading in moral matters. I rebut these arguments. I then clarify how the Appeal to Disgust really works. On my interpretation, it rests on two premises: (1) we often lack introspective access to the reasons underlying our moral disgust reactions

(*Inaccessible Premise*), and (2) sometimes, widespread moral disgust is undergirded by good reasons—reasons we *would* endorse if we could access them (*Wisdom Premise*). I support the Inaccessible Premise by drawing on gene-culture co-evolutionary theory, which explains how culturally transmitted norms often become opaque to those who inherit them. I support the Wisdom Premise by examining the case of polygyny, a marriage practice that Western lawmakers banned—guided in part by disgust—long before modern social science revealed the social harms associated with it. Upshot: while disgust is fallible, it sometimes points us in the right direction—and that is enough, I believe, to warrant lawmakers’ taking it seriously.

Chapter One

Liberalism and the Legal Status of Harmless Taboos

1. Introduction

Most people assume that incest, cannibalism, marketizing human organs, bestiality, and necrophilia are bad — *really, really* bad. In American society (and elsewhere), these practices are *taboo*, forbidden acts that mainstream opinion sees as sickening, offensive, and rude to even talk about in polite company. We enforce compliance with these taboos via powerful informal sanctions, such as shame and reputational damage. If you get caught French kissing a cousin, canine, or cadaver, your reputation will be forever ruined.

Informal sanctions aside, repugnance toward taboos is frequently expressed in the law. Incest is illegal in every state but Rhode Island, and in the vast majority of states it is a felony (Posner and Silbaugh, 1996); bestiality is criminalized in most states, sometimes as a misdemeanor but usually as a felony (Posner and Silbaugh, 1996); selling human organs is illegal under federal law (National Organ Transplant Act, 1984); cannibalism and necrophilia, though only explicitly banned in a minority of states, can be prosecuted under either corpse-desecration laws, which all states have, or federal laws against organ sales, which apply nationwide.

Punishments for violating the above taboos are typically harsh, often carrying a maximum sentence of years (or decades) in prison and thousands of dollars in fines. For instance, in a handful of states, the maximum penalty for consensual adult incest is life imprisonment. The penalty for selling human organs, which is illegal under federal law, is similarly harsh, carrying a maximum sentence of five years in prison, a \$50,000 fine, or both.

What interests me about the five aforementioned taboos—incest, cannibalism, bestiality, necrophilia, and human organ sales (call them the “foul five” for short)—is the philosophical unclarity surrounding the question of why they should be criminalized at all. None of these practices appears inherently harmful, based on any plausible interpretation of the term “harm.” On the contrary, when conducted in private by consenting adults, they seem harmless. By “harmless,” I mean no one’s physical well-being, mental health, finances, or property is adversely affected; no one’s interests or autonomy are set back; and there’s no danger to public health or safety. They may be deeply disgusting, but is that fact, in and of itself, a reason to criminalize them? Calf fries, C-sections, colonoscopies, flatulating in public, taxidermy, having pet cockroaches, and chewing tobacco are also often seen as disgusting—but we don’t impose prohibitions on those.

Obviously, I don’t mean the foul five can never be harmful. Incest between parents and children who are minors is harmful, since minors lack the capacity to meaningfully consent, and non-consensual sex is just rape. Pedophilia and rape should clearly be crimes. But what about when, for example, an adult child gives full consent to have sex with an adopted parent? Who gets harmed? Concerns about offspring having birth defects aren’t applicable, given the lack of genetic relation. Nevertheless, incest laws typically provide no exemption for such cases. In most states, non-genetically related adults who engage in consensual incest can be prosecuted as felons (Posner and Silbaugh, 1996). In Alabama, Delaware, Minnesota, Tennessee, Texas, and Virginia, this includes adopted siblings. It’s also worth mentioning that most states explicitly classify consensual, non-penetrative sex between adult family members (including siblings) as a felony offense, despite the behavior’s lack of reproductive potential.

Similar stories unfold with the rest of the foul five. Cannibalism and necrophilia are obviously wrong when a cadaver is procured through murder or grave robbery, or when there are threats to public health. But what if, for example, a cannibal or necrophile obtains prior consent from the deceased (who dies of natural causes) and their family, and health inspectors verify that the cadaver is safe for consumption or sexual use? Here, it becomes less evident who gets harmed; nevertheless, I'm not aware of any jurisdiction in which you could do such actions without getting arrested. Similarly, bestiality is clearly wrong when it involves animal abuse, but it's at least theoretically possible for a zoophile to engage in private and harmless activities with a cooperative animal. It's also possible to imagine a healthy, financially secure adult consenting to (e.g.) sell one of their kidneys to a modern artist for use in an eccentric art piece. Nevertheless, in such cases, criminal laws against bestiality and organ sales provide no exemptions.

The foul five may be deeply disgusting, but it's not clear that something should be illegal just because it is disgusting (otherwise, why are calf fries and C-sections legal?). If we can't articulate a reason why the foul five should be criminalized, then imposing absolute bans on them seems inconsistent with basic liberal principles. After all, a basic principle of liberalism is the Presumption of Liberty: citizens have a *prima facie* right to do what they want (especially in private), and if the state wants to ban conduct, X, the burden of argument falls on the state to demonstrate a reason why banning X is necessary, where the reason typically involves harm or some rights violation.

The rest of this chapter considers a few ways to go about justifying bans on the foul five. I consider justifications drawn from mainstream liberal thought: (1) the Harm Principle, (2) the Offense Principle, or (3) Legal Moralism. Ultimately, I argue, none of these options work. The foul five don't merit criminalization under the Harm Principle because they don't inherently

cause harm in any plausible sense. They don't merit criminalization under the Offense Principle because the Principle doesn't justify intrusions into the private sphere (current bans on the foul five extend to private activity).

The third option is to draw on Legal Moralism, the view that the state may sometimes criminalize immoral behavior if the immorality is sufficiently grave to concern the public. But this won't work either because it's difficult to articulate a reason *why* the foul five should be thought morally wrong. Research in moral psychology shows most people morally disapprove of the foul five while simultaneously struggling to verbalize why they think this (in the literature, this is known as "moral dumbfounding") (see for example Haidt et al., 2000). A review of philosophical literature reveals a similar lack of clarity among philosophers regarding what exactly makes the foul five morally wrong. If both ordinary citizens and philosophers find it dumbfounding to explain why the foul five are wrong, it's reasonable to suspect that legislators would be similarly dumbfounded—in which case the state has not met its justificatory burden imposed by the Presumption of Liberty.

In short, this chapter motivates the idea that there is an apparent tension between basic liberal principles and laws against the foul five. The upshot is as follows. *Decriminalizing the foul five would be crazy*. There would be no popular support for this. A presidential candidate who ran on this issue would become a laughing stock. If liberalism doesn't offer resources to justify prohibiting the foul five, then that is a serious blow to its philosophical credibility. It's one thing to be willing to bite certain bullets that one's views entail, but it's another thing for one's views to overturn deeply held moral or political intuitions. Some bullets shouldn't be bitten. It follows that liberals need an argument that can justify laws against the foul five, without compromising any of their basic commitments. What would such an argument look like? Almost

certainly, it will involve attributing to citizens' widespread feelings of disgust—all by themselves—some ability to justify laws, at least in some cases. (After all, our main objection to the foul five is not that they are rude or impolite or harmful, but that they are deeply disgusting.) Giving an account of how disgust can justify law is beyond the scope of this chapter, however. My sole focus here is to establish the *need* for such an account, which subsequent chapters will go on to build out.

2. The Presumption of Liberty

By “liberalism” I mean the dominant political ideology in the West today. It is distinguished by the importance it places on the civil and political rights of individuals (such as the rights to free speech, occupation, association, conscience, etc.), which the state must not intrude upon, except to protect the rights of others. Some leading exponents include Locke, Kant, Mill, Rawls, and Feinberg, to name a few. On the standard account, liberalism grew out of the Wars of Religion, which occurred throughout Europe in the 16th, 17th, and 18th centuries (Weatherford, 1995). Protestants and Catholics eventually recognized that the state could not impose a single faith, and thus the only stable basis for a political regime was to separate church and state, barring the state from intruding on the religious lives of its citizens.

Liberalism later extended this principle about religion to other spheres of life in which citizens hold conflicting beliefs about the meaning of life or how to live (Weatherford, 1995). The goal of the liberal state is not to resolve these conflicts but to create a neutral framework in which citizens can pursue their diverse conceptions of the good life in peace. Obviously, liberal theorists, especially in the 20th century, have introduced diverse — sometimes conflicting — conceptions of notions central to liberalism, such as “liberty,” “neutrality,” “rights,” “harm,” etc. For example, this has led to a divorce between “classical” liberals, who follow (e.g.) Hayek and

Nozick in championing minimalist government, private property rights, and the free functioning of markets against state interference, and “new” liberals, who place emphasis on egalitarianism, social justice, and the welfare state (Courtland and Schmidtz, 2022).

Despite these in-house divisions, we can still identify some central tenets of liberalism that transcend divisions between different theorists. In particular, at least one extremely basic idea, presumably shared by all liberals despite their diverse conceptions of “liberty,” is that liberalism begins with the premise that political authority and its exercise must be *justified* (as opposed to arbitrary, or unreasoned, or unjustified) (Courtland and Schmidtz, 2022). Any time citizens are legally obliged to do something they would not otherwise prefer to do, or legally disallowed from doing something they would otherwise prefer to do, there must be a reason why. It’s not the citizen’s job to persuade the state to be neutral regarding their preferred conception of the good life; it is the state’s burden to offer a clear, compelling rationale for any non-neutral treatment of citizens.

The name for this basic idea is the Presumption of Liberty. The Presumption finds expression in several paradigmatic liberal theorists. For example, Mill (1963) writes:

the burden of proof is supposed to be with those who are against liberty; who contend for any restriction or prohibition.... The *a priori* assumption is in favour of freedom... (p. 262).

Rawls (2001) agrees:

...there is a general presumption against imposing legal and other restrictions on conduct unless there is a sufficient reason... (p. 44).

Feinberg (1984) also agrees:

...most writers on our subject have endorsed a kind of ‘presumption in favor of liberty’ requiring that whenever a legislator is faced with a choice between imposing a legal duty on citizens or leaving them at liberty, other things being equal, he should leave

individuals free to make their own choices. Liberty should be the norm; coercion always needs some special justification (p. 9).

For the purposes of this dissertation, let's understand the Presumption of Liberty as follows:

Presumption of Liberty: Citizens generally possess a *prima facie* right to act as they please (especially in private), such that if the state wants to restrict their behavior, it bears the “burden of argument” to demonstrate why doing so is necessary.

By “burden of argument,” I mean that in a debate about whether conduct, X, should be illegal, those opposing criminalization have an advantage. They need not demonstrate why citizens have a right to do X; it's the state's responsibility to demonstrate why X should be illegal. With this dialectical advantage, opponents can simply argue, “The state has not presented a convincing case for its pro-criminalization stance,” and that should be a winning strategy. Conversely, a passive approach is not available to the state. It can't just say “Criminalizing X is justified because citizens have not articulated an argument for their right to do X.” This would constitute a failure on the state's part to fulfill its argumentative burden.

There are different kinds of reasons to accept the Presumption of Liberty. One sort of deontological argument might be that it's intrinsically wrong for the state to impose arbitrary, not-rationally-supported restrictions on citizens' liberty, insofar as this violates their natural rights. But I'm inclined towards a more pragmatic, historical argument, which simply says that past experiences make it prudent to maintain healthy skepticism about legislative motivations (Barnett, 2003). History has shown that states and majority factions often have dubious motives when it comes to legally coercive actions such as regulating free speech, drafting people into the military, granting market monopolies, and so on. While there may be legitimate reasons to do these things, we can't blindly entrust the state or the majority with unlimited discretion to draft criminal laws. Merely asserting, without any elaboration, that “Coercing people in manner X is

justified” won’t suffice to fulfill the state’s burden of argument; nor will it suffice to show that a majority would be pleased if a minority were coerced in manner X (Barnett, 2003). Allowing such “justifications” would grant the state a dangerous amount of latitude.

Caveat: The reader may ask, *where* does the state make an argument? The legislative record? Relevant court decisions? The court of public opinion? This is a question I’m not quite sure how to answer. As a liberal (in the very basic sense outlined above), I’m of the opinion that the state has an obligation to show through argument that its criminal laws are non-arbitrary. For a law to be non-arbitrary, it must serve the common good in some way (in the very broad sense of “does not merely serve majority or minority interests”). How do we know whether a given law serves the common good or not? Presumably, legislators will make arguments that appeal to liberal values. If it is *not* possible to construct an argument in support of a law, X, that appeals to liberal values, then X violates the Presumption of Liberty (i.e. citizens are not presumed free to make their own choices but, rather, are being treated paternalistically by legislators). As to how my liberal philosophy should play out in the “real world,” or what the enforcement mechanism should be (e.g. courts? a constitution with a bill of rights?), I’m not taking on that question here. (After all, it is secondary to the *philosophical* question of whether it is possible to justify laws against the foul five, which is my focus.) At the very least, it seems to me, legislators should self-regulate their behavior to be consistent with the Presumption of Liberty, and citizens should punish them by voting them out of office when they don’t self-regulate.

2.1. *The Harm Principle*

If bare assertions and appeals to majority won’t suffice to justify criminal laws, how, then, should the state go about fulfilling its burden to provide justification? One traditional

answer from a liberal standpoint is that the state should adhere to the Harm Principle. In Mill's statement, the Principle holds that

the only purpose for which power can rightfully be exercised over any member of a civilized community against his will is to prevent harm to others (1978: 9).

So, the state could justifiably prohibit the foul five if they were causing "harm to others."

But this is a moot point, as it regards the foul five. As noted, there are numerous situations in which incest, cannibalism, selling human organs, bestiality, and necrophilia can occur without injuring anyone's body, mental health, finances, or property; without setting back anyone's interests or autonomy; and without endangering public health or safety. (Pedophilia and rape are harmful, but who is harmed by incest between consenting adults? Murder and grave robbery are harmful, but who is harmed by a cannibal or necrophile gaining explicit consent to obtain an already-dead body? Etc.) Therefore, the state making a bare assertion such as "A blanket ban on the foul five is necessary to prevent harm" is a losing argument.

Objection: The foul five should be illegal (or closely regulated), despite not being *inherently* harmful, for the same reason that the sale, production, and possession of very dangerous products (e.g. automatic weapons) are regulated. A responsible person could use an automatic weapon for benign recreational purposes (e.g. shooting watermelons for YouTube videos). However, as a matter of policy, more harm is prevented—with an acceptable burden on liberty—by closely regulating these weapons than by having no regulation. In other words, the risk to public safety posed by untrained, irresponsible, or ill-intended actors justifies placing a burden on the liberty of *responsible* actors. Likewise, the risk posed to public safety by the foul five is sufficiently high to justify placing a burden on the liberty of responsible actors.

Reply: First, it's worth noting that the foul five aren't regulated; they're entirely banned (e.g. you can't obtain a license to engage in consensual incest). Second, placing burdens on

liberty in light of risks to public health or safety is, I concede, compatible with the Presumption of Liberty, provided the nature of the risk is clearly articulable—which, in the case of something like automatic weapons, it is. However, bare assertions like “The foul five pose a high risk to public safety” won’t do, given this would grant the state a dangerous amount of latitude. What specific risks of harm do the foul five entail? Pedophilia and rape are clearly harmful, but it’s unclear who is harmed, or is *at risk* of being harmed, by incest between consenting adults. Murder and grave robbery are harmful, but it’s unclear who is *at risk* of being harmed by a cannibal or necrophile gaining explicit consent to obtain an already-dead body. And so on. In short, unless the state can articulate exactly how banning the foul five would secure a corresponding payoff in terms of reducing the likelihood of harm, its burden of justification remains unmet.

2.2. *The Offense Principle*

A second option available to the liberal theorist may be the Offense Principle. In Joel Feinberg’s (1985) statement, this Principle holds that

It is always a good reason in support of a proposed criminal prohibition that it would probably be an effective way of preventing serious offense (as opposed to injury or harm) to persons other than the actor, and that it is probably a necessary means to that end (p. 1).

In essence, according to Feinberg, the Harm Principle doesn’t exhaust all the “serious” wrongs that should attract the attention of the criminal law. For example, if a couple on a public bus starts expressing excessive romantic affection, to the point of nudity, other passengers may feel deeply offended (Feinberg, 1985). However, to describe this offense as a “harm” stretches the term too far. It seems more accurate to say the passengers have been rightly and legitimately offended. Therefore, even in a liberal society where citizens hold diverse moral and religious

doctrines, citizens still have an interest in defining “serious offenses” and establishing reasonable legal boundaries around them.

However, I fail to see how the Offense Principle could be used to justify blanket, exception-less prohibitions on the foul five, akin to the current laws in place. Feinberg defines an “offense” as an affront to one’s sensibilities, causing an unpleasant psychological state (annoyance, embarrassment, shame, agitation, etc.). While any such offense can ground criminalization, two necessary conditions must be satisfied, on Feinberg’s account. First, for criminalization against an offense, X, to be justified, it must be the case that the vast majority of people are greatly offended by X. A slight majority being greatly offended is not sufficient, nor is a great majority being slightly offended (Feinberg, 1985). Current criminal laws against the foul five seem to satisfy this condition, given that the foul five are greatly offensive to the vast majority.

However, current laws fail to satisfy the second condition, which holds that even if conduct, X, is deeply and widely offensive, anti-X legislation must undergo certain tests in order to reserve maximal space for individual liberty (Feinberg, 1985). Two examples:

- (1) Can those offended by X alter their behavior in a relatively easy manner to avoid experiencing offense, thereby preserving the liberty of those committing the offense? If so, then coercion is uncalled for. For example, if most people are deeply offended by the fact that some movie theaters show pornographic movies, coercion is not thereby justified, since those who are offended can simply avoid those theaters.
- (2) Can those committing the offense change the time, manner, or place of their conduct to avoid causing offense, without undertaking a considerable burden? If this is not a practical possibility, then coercion interferes too much with the liberty of those committing the offense. For example, even if most people are deeply offended by interracial couples, asking these couples to avoid ever appearing together in public spaces would be an impossible burden.

In light of the above, *some* regulation of the foul five may be justified in public spaces. For example, it seems reasonable to prohibit cannibals from eating lunch on public buses, to prohibit

organ vendors from advertising their merchandise on graphic billboards along busy highways, and so on. But existing prohibitions on the foul five clearly fail one or both tests. Currently, these practices are illegal even when conducted by consenting adults in a harmless and private manner. The Offense Principle does not permit intrusions into the private sphere. Laws against public nudity, intoxication, sex, urination, defecation, and excessive noise don't extend to private conduct because there is no audience present to be offended. Why should the law treat the foul five differently? When there is no non-consenting audience, there is no offense.

3. Legal Moralism

If criminalizing the foul five can't be justified via the Harm or Offense Principles, it might seem like liberal theorists have exhausted their justificatory resources. But, I say, at least one option remains available: we could endorse Legal Moralism, the legal theory according to which the state is occasionally justified in criminalizing immoral behavior as such, provided that the immorality in question is sufficiently grave to be of public concern. Different theorists interpret "sufficiently grave to be of public concern" in different ways, but the basic idea is that if an immorality is so extreme that it threatens the integrity of "public morality," then criminalization is justified, given all citizens have an equal interest in preserving public morality.

3.1. The Hart/Devlin Debate

Legal Moralism is sometimes construed as a rival to political liberalism, but I don't think this is necessarily the case. The perception that Legal Moralism and liberalism are rivals seems to be primarily due to the influence of Patrick Devlin. In the midst of the so-called Hart-Devlin debate during the mid-20th century, Devlin (1959) put forth the argument that (1) a society's identity is largely constituted by the content of its morality, and (2) therefore, a society has a

prima facie right to employ coercion to forestall changes to its morality, when these changes are significant enough to jeopardize its maintenance of identity across time.

In response, Devlin's primary opponent, H.L.A. Hart (1963), pointed out that the above argument has absurd, illiberal implications. In particular, the argument fails to draw a distinction between "positive" morality—the ethical standards that happen to be widely accepted among citizens at a given time—and "critical" morality—the objective, ethical principles that citizens can employ at any time to critique and improve whatever positive morality happens to be in vogue. With this distinction on hand, it becomes clear that Devlin's argument could be coopted to justify perpetuating virtually any positive morality, even if it is *unjust* from the standpoint of critical morality, in the name of "preserving society's identity." This relativism is indeed incompatible with liberalism. It assigns high priority to preserving society's identity, even if achieving this must come at the expense of basic liberal values such as individual liberty, autonomy, privacy, or state neutrality.

But surely Devlin's relativism is not an essential feature of Legal Moralism itself. Contemporary legal moralists such as John Kekes (1992; 2000), Robert P. George (1993), Steven Wall (2013), R.A. Duff (2014), and Michael Moore (2017) simply incorporate Hare's positive/critical morality distinction into their arguments. That is, they remain *bona fide* legal moralists, believing that the state is occasionally permitted to criminalize immoral actions as such, but they place a stringent standard on what qualifies as falling within public concern. Each of these theorists holds that, for actions to qualify, it is necessary (though not sufficient) for them to violate a "critical" ethical principle. This approach alleviates concerns about illiberal relativism. Therefore, if Legal Moralism can be modified to anchor its claims in "critical" ethical principles, compossibility with liberalism can be restored. In that case, liberals could attempt to

justify criminalizing the foul five within a legal moralist framework. This approach might fail on its own merits, but not for lack of compatibility with liberalism. For convenience, let's refer to versions of Legal Moralism that don't lead to illiberal relativism as "Liberal Legal Moralism" (for an overview of such views, see Petersen, 2010).

Endorsing Liberal Legal Moralism would certainly be a convenient move as it regards retroactively justifying criminalizing the foul five, given the *actual* legislative motive for their criminalization often seems to be simply that most people find them inherently morally repugnant. Posner and Silbaugh (1996) state that "the strongly held community norms against incest are reason enough, in the minds of most people, to criminalize incest" (p. 129). Opponents of legalizing organ markets argue that such markets "would diminish human dignity and our sense of solidarity" (Cohen, 2002: 47). State laws against bestiality frequently employ emotionally charged adjectives to describe the act, such as "abominable," "unnatural," "detestable," "infamous," "perverted," and "a crime against nature" (Posner and Silbaugh, 1996: 207-212). State laws against cannibalism and necrophilia often categorize conduct involving human remains as illegal if it would "outrage reasonable family sensibilities" or would be "offensive to the sensibilities of an ordinary person" (Posner and Silbaugh, 1996: 213-216). Whether taboos are harmful is irrelevant, it seems; the predominant concern is preventing our moral ecology from being polluted with behavior that most people consider morally repugnant.

3.2. *Kekes' Legal Moralism*

It will help to briefly recount a specific account of Liberal Legal Moralism. One prominent account comes from John Kekes (2000; 1992). Kekes (2000) explicitly states that his argument "is one that reason recommends to everyone, liberals included" (p. 26). His basic argument is that while diverse societies, bound to encounter moral disagreements, should exalt

the civic virtue of tolerance, it remains an unavoidable fact that some lifestyles threaten the minimum conditions on which civilization (as opposed to lawless, anarchic barbarism) itself depends. These lifestyles warrant disgust and must be placed beyond the bounds of toleration.

Kekes distinguishes between “primary conventions”—which organize and dispense the basic requirements that all good lives need, like safety, health, and nutrition—and “secondary conventions”—which organize and dispense secondary, culturally variable goods that enrich good lives, like religious faith and specific family customs (p. 24). Citizens’ conceptions of secondary conventions are bound to conflict in diverse societies, so tolerance is demanded. But primary conventions are invariable; everyone’s secondary conventions depend on them. Thus, breaches against them are intolerable and must elicit outrage. Outrage, says Kekes, signals that citizens not only abstractly understand that civilized life is underpinned by a set of minimum, necessary conditions, but also that they are deeply committed to their maintenance.

But not all breaches to primary conventions are of equal seriousness. Some not only breach primary conventions but do so *flamboyantly*. Kekes (1992) calls these “taboos.” They are “the worst evil that people can do” (p. 432) and the “secular equivalent of sacrilege” (p. 443). While all immoralities are wrong, taboos are of a different order; the perpetrator’s inner attitude and motive is so beyond the pale as to be incomprehensible to morally committed persons. For example, consider that a person who is not debased could still agree that murder driven by revenge, albeit condemnable, is within the realm of the comprehensible (think *West Side Story* when Tony vengefully kills Bernardo or Exodus 2 when Moses kills the Egyptian). Meanwhile, the idea of Jeffrey Dahmer killing someone and eating their body—no set of circumstances, no story, could make such behavior understandable. Eating a human being is *unspeakable* (literally, just speaking of taboos at the dinner table is bad manners). Dahmer (or any taboo-violator)

“shows such a blatant disregard of what we regard as civilized life that anyone who so acts is placed beyond the pale” (p. 443). His breach conveys *disdain* for primary conventions and, by extension, civil order itself. His act represents “the ultimate object of fear” for all political societies, namely, “the dreadful possibility of the removal of all limits” (p. 443).

If safeguarding primary conventions requires that citizens be outraged by breaches against them in proportion as they are flagrant, then taboo-violations must evoke *disgust*—a zero-tolerance emotion. In being disgusted by cockroaches, we recoil, cannot stand their presence, and our bodies viscerally demand they be removed from our vicinity. Likewise, flagrant violations to primary conventions—i.e. the commission of a taboo act—must (and, for morally committed witnesses, *does*) evoke this same qualia. Upon perceiving a taboo act being committed,

we are nauseated, our guts heave, we tremble, our hearts race, we both sweat and shiver, we cannot control our physical responses, or we control them only with great effort. ... We regard these matters as fundamentally unsettling, their memories stay with us, they are the stuff of which nightmares are made (p. 433).

At this point in our search to justify criminal bans on the foul five from a liberal standpoint, the usefulness of Kekes’ Liberal Legal Moralism becomes clear. Behaviors that threaten to undermine society’s primary conventions, even if they do not cause direct or immediate harm to anyone, merit criminalization. This does not open the flood gates to public intrusion into the private sphere (which would be illiberal) because actions that merely violate moral norms do not fall into this class of primary-convention-threatening conduct; this classification is reserved for actions that *flaunt morality itself*. Actions that flaunt morality, left unchecked and unmonitored by the law, run the risk of desensitizing citizens to socially destructive behaviors. Yet, mass desensitization to socially destructive behaviors is incompatible with civilization; the survival of primary conventions depends on citizens possessing a deeply

felt commitment to them. So, if the foul five, albeit not inherently harmful, are gravely immoral—i.e. they pose a threat to primary conventions—then enforcing absolute bans on them (akin to those currently in place) constitutes a proportionate response to their danger to society.

4. Moral Dumbfounding

Before we celebrate that we've found a way to rationalize current prohibitions on the foul five, consider a challenge: we need to establish some reason to believe that the foul five are actually (gravely) immoral. This is important because on Kekes' account, only *actual* threats to primary conventions merit criminalization, not merely perceived threats. If most citizens were convinced that (e.g.) wearing a blue tie on Thursdays threatens the minimum conditions that underpin civilization, this would not justify criminalization. Wearing blue ties on Thursdays clearly does not flout primary conventions, and nothing bad happens when people become desensitized to that which they *mistakenly* consider socially disruptive.

Analogically, it follows that if the foul five are (by some coincidence) mistakenly believed to be evil or immoral (when in fact they aren't), then mass desensitization to them wouldn't cause society's downfall (rather, nothing would happen).

This reasserts the importance of the Presumption of Liberty: the burden of argument must fall on the state to clearly articulate why or how exactly the foul five threaten primary conventions, such that mass desensitization to them would be socially catastrophic. Kekes seems to agree, as evinced by the following prefatory remark inserted early in his 2000 essay:

The countless historical and contemporary instances in which individuals were and are unjustifiably oppressed in the name of arbitrary religious, moral, or political ideologies will be abhorrent to those whose allegiance lies with the part of humanity (p. 23).

In other words, Kekes acknowledges that the state has a history of abusing citizens with arbitrary uses of coercion; thus, he hopes to design his Legal Moralism *to be liberal-compatible*, to avoid

opening the door to arbitrary coercion. In that case, it's reasonable to impose on the state the burden of showing through argument that the foul five are immoral.

Can the state meet this burden? I examine this question below. It is ultimately my view that the state *cannot* meet its burden—or at least, how it might do so is unobvious.

4.1. Moral Dumbfounding about the foul five among the public

The liberal legal moralist needs a compelling argument that the foul five are actually immoral (or somehow threatening to society's primary conventions). Given the Presumption of Liberty, the bare assertion "The foul five are immoral" won't do; nor will an appeal to the majority of the form "Most people believe that the foul five are immoral." What substantive argument might the state supply?

At this juncture, it's time to address the elephant in the room: despite our strong instinctive aversions to the foul five, it is not obvious *precisely why* we consider them morally wrong at all (let alone gravely immoral, to the point of being socially destructive). The easiest evils to articulate are those that are somehow associated with a form of widely acknowledged disvalue, such as harm, inequality, injustice, the violation of consent, exploitation, dishonesty, disloyalty, disrespect, etc. But it's unobvious which form of disvalue is present in the case of two siblings having non-penetrative (i.e. infertile) sex, or a healthy person selling a kidney to a modern artist for use in an eccentric artwork, provided the actors are adults who provide explicit, untainted consent and conduct their business in private (to avoid triggering disgust in spectators).

Empirical research has shown that most ordinary people, when questioned, struggle to explain why they disapprove of the foul five (McHugh, 2017; Haidt and Hersh, 2001; Haidt, 2001; Tetlock, 2000; Haidt et al., 2000; Fiske and Tetlock, 1997; Haidt, Koller, and Diaz,

1993).¹⁵ This is a widely documented phenomenon in the field of moral psychology, where it's known as "moral dumbfounding."

The typical procedure in moral dumbfounding research is to start by giving subjects a story in which a certain taboo is broken, albeit no one gets harmed in any way. For example, a study by Jonathan Haidt and colleagues (2000) famously gives the story of Julie and Mark, two biological siblings who, while vacationing together, have a consensual sexual encounter, but use two forms of protection, keep it a secret from everyone they know, and promise to never do it again. Subjects are then asked to evaluate the actors. Unsurprisingly, the overwhelming majority express strong disapproval. But when pressed by interviewers to explain why they disapprove, subjects often become "dumbfounded"—unable to find the words to explain or support their judgment. The fact that a subject is dumbfounded is sometimes evinced by their own direct admission of being stumped or confused. Other times, researchers infer that a subject is dumbfounded based on various telltale signs, such as (1) repeating unsupported claims¹⁶ ("That's terrible!", "It's just wrong to do that!"); (2) saying "I don't know" over and over again; (3) non-verbal displays of confusion like excessive laughing, face-touching, or saying "um"; and (4) beginning to put forward arguments but abruptly abandoning them upon realizing they lead to a dead end (Haidt et al., 2000).

¹⁵ The term "foul five" is my term, not a term found in moral psychology literature. There are other things that provide dumbfounding, like eating dogs in Western culture (see Haidt, Koller, and Diaz, 1993). I'm focusing on the foul five because (1) they're harmless, (2) they're taboo, and (3) they're illegal.

¹⁶ I use the term "unsupported claims" because that is the one the authors of the study themselves use. However, it may not be the best term. When a subject says "That's terrible!" or "That's just wrong!" they *might* be asserting a proposition, which they then fail to support with evidence (hence "unsupported claim"). Or, they *might* simply be conveying their inner emotional state (e.g. "I am currently feeling very agitated")—sort of like how someone who personally dislikes pistachio ice cream could exclaim "Pistachio ice cream is bad!" without intending to assert that it is *actually worse* than any other flavor, or that there is some objective truth about how to rank ice cream flavors. They're just emoting.

Moral dumbfounding was originally defined by Haidt et al.(2000) as “the stubborn and puzzled maintenance of a judgment without supporting reasons” (p. 1). Different definitions have since been proposed, but one element consistently identified across definitions is “an absence of reasons for, or an inability to justify or defend, a moral judgement” (McHugh, 2017: 1). Much research in this field has focused primarily on the incest taboo (see Haidt et al., 2000; Haidt and Hersh, 2001; Haidt, 2001; Haidt, Koller, and Diaz, 1993), but widespread moral dumbfounding has also been observed surrounding the rest of the foul five. For example, for selling human organs, see Tetlock, 2000. For cannibalism, see Haidt et al., 2000. For bestiality, see Haidt and Hersh, 2001. And for necrophilia, see Haidt 2001. As a matter of empirical observation, many people just can’t explain why the foul five are wrong.

4.2. Moral Dumbfounding about the foul five among philosophers

You might expect that moral dumbfounding, while prevalent among the philosophically unsavvy masses, is observed to a much lesser degree among philosophers, given their exceptional sagacity. On the contrary, philosophers are as stumped as everyone else. For example, a string of recent work by philosophers has examined the standard arguments given in support of the legally-enshrined taboo against consensual adult incest and determined them to be transparently untenable (see for example Green, 2017; Bergelson, 2013; Tralau, 2013; Sebo, 2006).

Similar work by philosophers has called into question the standard justifications for other taboos, including organ sales (see Brennan and Jaworski, 2015a; 2015b), bestiality (see Levy, 2003; Singer, 2001), and cannibalism (see Lu, 2013; Shand, 2010; Wisniewski, 2004; 2007). For example, Matthew Lu (2013), concurring with Shand (2010) and Wisniewski (2004; 2007), declares that “no successful argument against [cannibalism] can be articulated within the

dominant consequentialist or neo-Kantian deontological approaches in normative ethics” (p. 433). Lu personally believes that “a virtue ethics of a broadly Aristotelian form” can show cannibalism to be wrong, but acknowledges “it is surprisingly difficult to explain why” (p. 433). As for bestiality, Neil Levy (2003) reviews and ultimately rebuts several standard objections to the practice (e.g. it’s a perversion, it violates consent, it wrongfully uses animals as mere means). He finally concludes, we could argue from a communitarian standpoint that the practice is (in a sense) irrational; nevertheless, there is “nothing *immoral* [emphasis original] about bestiality” (p. 453).

5. The Ethics of Incest

The previous section claims that philosophers don’t know why the foul five are wrong. To give this claim some concreteness (beyond offering citations), let me briefly review a non-exhaustive, but still representative, sample of arguments against incest. This focus on the ethics of incest is mostly arbitrary; I could have focused on any of the foul five. But the case of incest is illustrative. Below, consider some failed philosophical arguments in support of the claim that “Incest is wrong.”¹⁷

5.1. *The genetic argument*

The “genetic argument” basically holds that incest is wrong because the offspring of closely related parents have an increased risk of genetic disorders.

This argument fails by virtue of being both over- and under-inclusive. It’s under-inclusive in that it ties the wrongness of incest to the risk of inbreeding, yet we ordinarily find any case of sex between family members disgusting, even if the risk of pregnancy is (virtually)

¹⁷ Some of the arguments considered here are ones that I have seen gestured at in philosophical scholarship, others are arguments that I have simply encountered in casual conversations with friends and colleagues.

zero, as in the story of Julie and Mark (or if the participants are same-sex, or if one of the participants is post-menopausal, or if the participants engage in non-penetrative sex only, etc.).

The argument is over-inclusive in that it establishes a general principle (sex is wrong any time there's an increased risk of genetic disorders) that we typically don't—and wouldn't—apply to non-incestuous couples who choose to procreate despite being aware of an increased risk of Autism Spectrum Disorder, Huntington's disease, hemophilia, sickle cell anemia, Down syndrome, Tay Sachs, cystic fibrosis, or other genetic disorders (Green, 2017).

5.2. The family argument

The “family argument” holds that sex within families is wrong insofar as it is likely to incite competition and emotional fallout.

Apart from sounding ridiculous (Julie and Mark shouldn't copulate because mom or dad might get jealous?), the family argument also fails by virtue of being over- and under-inclusive. It's over-inclusive in that many practices and arrangements may cause family dysfunction (helicopter parenting, divorce, infidelity, cohabiting with someone your parents dislike, alcoholism, workaholism) that don't typically evoke moral disgust or merit legal sanction (Green, 2017).

And the argument is under-inclusive in that it's at least possible for intra-family sex to be dysfunction-free, as evinced by sibling marriages among the rulers of the Incan empire, native Hawaiian royalty, ancient Persian rulers, and the Ptolemaic dynasty in Egypt (of which Cleopatra, who married her younger brother, was part) (Benokraitis, 2005).

5.3. The perversion argument

The “perversion argument” is this: (1) Incest is a sexual perversion. (2) If incest is a perversion, then it is immoral. (3) Therefore, incest is immoral.

Both premises are contestable. The first claims that incest is a sexual perversion. To assess this claim, we need an account of “sexual perversion.” One prominent account comes from Thomas Nagel (1969). According to Nagel, a sexual perversion is a desire that deviates from the natural psychology of sexual desire. The natural psychology of sexual desire is defined as “a desire that one’s partner be aroused by the recognition of one’s desire that he or she be aroused” (p. 12). That is, if X and Y sexually desire each other, then not only are X and Y mutually aroused (X is aroused by Y and vice versa), but also X is aroused *by Y’s arousal for X*, and Y is aroused *by X’s arousal for Y*. In short, sexual desire involves “reflexive mutual recognition” (p. 12).

On this view, a perversion occurs when this psychology (reflexive mutual recognition) is “truncated or incomplete” (p. 13). For example, shoe fetishism is a perversion here because *mutual* arousal is necessarily absent (the object of one’s arousal can’t reciprocate arousal). Consensual sadomasochism is also a perversion because with sadomasochism, the arousal is not reflexive; the partners necessarily objectify each other (the passive partner is aroused *by being objectified*, and the active partner is aroused *by objectifying the passive partner*).

What about incest? Is it a perversion on Nagel’s account? No. It’s possible for Julie and Mark to complete the natural psychology of sexual desire. They can be reflexively and mutually aroused.

Even if incest were a perversion, would this necessarily imply that it is wrong? The principle here seems to be that perversions are necessarily immoral. Nagel denies this claim, since affirming it would have counterintuitive implications, like entailing that masturbation is immoral (Levy (2003) and Solomon (1974) deny the claim for the same reason). Also, some

paradigmatically immoral sex acts, e.g. adultery, pass the perversion test, i.e. are non-perverse. Thus, the link between “perverse” and “immoral” seems weak.

5.4. *The Evolution Argument*

The “evolution argument” claims that incest is immoral because humans have been programmed by biological evolution to be incest-averse. Since incest carries the risk of inbreeding, animals that avoid incest, seeking instead to mate with non-relatives, thereby enjoy a competitive advantage in the ability to produce healthy offspring. Formally: (1) If incest-aversion bestows a reproductive advantage, then incest is immoral. (2) Incest-aversion bestows a reproductive advantage. (3) Therefore, incest is immoral.

Both premises of this argument seem contestable, but let’s focus on (1). Premise (1) seems to equate biological fitness with moral goodness, but this is a highly dubious equation. Few people, I would guess, would consider it *morally* valuable to propagate their genes with maximum efficiency. Moreover, we don’t normally morally condemn people who choose not to propagate their genes, such as homosexuals¹⁸, voluntary celibates, couples who choose to exclusively adopt, couples who are voluntarily childless, and so on.

Additionally, the argument doesn’t seem to recognize the distinction between an individual’s incest-aversion and a culture’s incest taboo. Most societies have an incest taboo, and it almost always applies to some of your family members to whom you’re not genetically related. For example, in the book of Mark, chapter 6, John the Baptist criticizes Herod for marrying the divorced wife of his half-brother, on the basis that such a relationship is forbidden in the Old Testament. Obviously, Herod isn’t genetically related to his half-brother’s ex-wife. Likewise, in

¹⁸ Of course, many people do believe homosexuality is wrong, but typically not for the reason that it hinders homosexuals’ ability to propagate their genes.

contemporary America, many states legally forbid sex with (a) adopted family members (e.g. step parents/children, adopted siblings), and (b) family members to whom you're only related by law (e.g. your siblings' spouses), even though having sex with adopted family members or in-laws carries no risk of inbreeding. Therefore, it seems that the incest taboo, as it manifests culturally, is not directly concerned with maximizing individuals' biological fitness.

If biological fitness isn't the concern, what about cultural fitness? It could be argued that incest is wrong because it decreases the fitness of one's culture. It turns out that, historically, "endogamous" cultures (those with a norm *against* mating with outsiders) generally fare worse than "exogamous" cultures (those with a norm in favor of mating with outsiders) (Henrich, 2016). This is because, when (e.g.) you marry your sister, you acquire no brothers-, sisters-, or parents-in-law. In contrast, if you marry an outsider, you acquire many in-laws. In-laws are useful insofar as they have a reason to support you and be loyal to you, and life is better when you have loyal supporters. Therefore, exogamy is socially valuable insofar as it builds alliances, and alliances make life better. Therefore, the incest taboo is morally valuable insofar as it supports exogamy.¹⁹

Again, this argument is dubious. Put in formal terms, it states: (1) Exogamy (incest-aversion on a cultural scale) is a culturally fit trait. (2) If a trait increases a culture's fitness, then it is morally good. (3) Therefore, exogamy is morally good.

Premise (1) seems true. But (2) faces many problems. First, again, it's an open question whether cultural fitness is a *morally* valuable thing. If the Nazis are exogamous, thereby

¹⁹ To be clear, Henrich (2016) details the cultural logic of incest taboos, in his capacity as an anthropologist. He *does not* claim that incest taboos are morally valuable, in light of their culturally adaptive benefits. *I* am making this extension, to have more fodder for philosophical discussion.

increasing their cultural fitness, this information is irrelevant from a moral standpoint. We place no moral value on the future propagation and advancement of Nazi culture.

Second, (2) seems to be *over*-inclusive. We don't typically condemn voluntary celibates, or individuals who marry someone with no or very little living family, for committing the sin of not building alliances via marriage. Yet, if premise (2) is true, perhaps we should, to be consistent. Of course, that would be highly counterintuitive.

Third, the argument is *under*-inclusive. Imagine Luke and Leia are biological siblings, but Leia was adopted at birth by another family, such that if the two were to marry, they would each thereby gain in-laws. Wouldn't it still be disgusting for Luke and Leia to marry, despite the fact that their marriage would technically be exogamous? If so, then it doesn't seem like the wrong-maker for incest is simply that it decreases cultural fitness.

5.5. *The Bible argument*

The "Bible argument" is this: (1) If the Bible forbids incest, then incest is immoral. (2) The Bible forbids incest. (3) Therefore, incest is immoral.

Premise (1) is ambiguous. It could be read as an endorsement of divine command theory (the wrongness of incest lies in God's forbidding it), or it could be read as ascribing epistemic credibility to moral claims found in the Bible, such that the Bible's condemning incest is sufficient evidence that incest is wrong.

Either way, this argument obviously has limitations insofar as both readings of (1) will be controversial in a diverse liberal democracy. Not everyone believes in the Bible, so if the Bible is the only reason to believe incest is wrong, then people who don't believe in the Bible have no reason to condemn incest. We can't make laws based on controversial religious doctrines, since that would violate basic liberal values.

Premise (2) also seems dubious. Does the Bible forbid incest? In the New Testament, incest is only mentioned twice, once by John the Baptist in Mark 6 and once by St. Paul in 1 Corinthians 5. In both cases, the speaker simply assumes incest is bad based on the fact that it is forbidden in the Old Testament, which the speaker considers authoritative. So, really, we should focus on the relevant Old Testament verses.

Indeed, incest is condemned in the Old Testament. Leviticus 18-20 contain approximately 60 proscriptions against incest (Carmichael, 1995). However, those proscriptions don't resemble the incest taboo as it manifests in contemporary America. For example, Leviticus doesn't explicitly condemn relationships between cousins, whereas many U.S. states treat such relationships as a felony.

Setting this worry aside, there's a further problem. Suppose divine command theory is false, such that the Bible just *reveals* moral facts to us (it doesn't *determine* what moral facts there are). This raises the question: Why, then, does God forbid incest in the Old Testament? What are his reasons?

According to biblical scholar Calum Carmichael (the only scholar I know of who addresses this question), the justification for forbidding incest is

to ensure that family life is as sexually unimpassioned as possible. . . In a polygamous set-up, when a father takes a new wife, a son by a previous wife may be around the same age as the new wife... In ancient Mediterranean society, where there was little social mixing of the sexes, it cannot surprise that a son may conceive a desire for his father's new wife. To permit the two men to compete for her sexual favors is a recipe for violence (1995: 126).

Thus, Carmichael suggests that God, in the Bible, forbids incest on the basis of the family argument. But, as we already established in section 5.2, the family argument isn't particularly persuasive, being both under-inclusive and over-inclusive.

5.6. *The intuitionist argument*

The foregoing arguments try to justify the belief that incest is wrong *inferentially* (i.e. based on an inference from prior premises). But maybe that's the wrong approach, like trying to inferentially justify the perceptual belief "My shirt is blue." I know my shirt is blue because that's what my eyes relay. Trust in my eyes is foundational, not based on some further belief. Likewise, maybe the wrongness of incest is knowable non-inferentially. I often know when things are bad immediately upon contemplating them, without thinking for longer than a few milliseconds (e.g. torturing babies for fun). And that's how I know incest is bad—it just *immediately seems* bad. Absent defeaters, no further argumentation is needed. If anyone doesn't "see" the wrongness of incest, their moral sense must be defective. (It's sort of like how, even if a color-blind person can't see that my shirt is blue, that doesn't make my belief in my shirt's blueness unjustified or irrational. Their eyes just don't work.)

Obviously, ethical intuitionism has critics, but I have no interest in attacking intuitionism *per se*. Instead, suppose *arguendo* intuitionism is true. Does that suffice to render endorsing the incest taboo—with no further evidence—rational? In other words, presupposing intuitionism, is non-inferentially-supported belief in the wrongness of incest rational?

To proceed, we need an account of "ethical intuition." Without exhaustively covering the possibilities, consider two possible accounts: (1) W.D. Ross' self-evidence theory, and (2) Michael Huemer's seeming-state theory.

Ross' self-evidence theory

According to Ross, moral truths are self-evidently true, analogous to how mathematical axioms are self-evidently true (here, I'm relying on the overview of Ross found in Bedke, 2010). With simple mathematical propositions like "two and two make four," we know they're true

immediately upon grasping the terms involved, without further reflection. If you grasp the concept “two” and also the concept “addition,” you should immediately see that the claim “two and two makes four” is true. In fact, it’s impossible to imagine anything other than two and two making four, since that would be a contradiction.

Similarly, you might think certain basic moral truths are knowable *a priori*, simply upon grasping the terms involved. For example, say a moral practice is intrinsically bad if the practice itself violates a value we hold dear, like respect for individual autonomy (human beings being left alone to guide their lives as they see fit). In that case, chattel slavery (property ownership over another human being) is intrinsically bad. After all, chattel slavery is logically incompatible with respect for autonomy. It’s logically impossible to imagine an “autonomous slave” (the very phrase is an oxymoron). Slavery here isn’t bad because it has undesirable consequences, like making the slave unhappy. It would still be bad even if the slave were (somehow) made happy by the arrangement.

Does this account of “ethical intuition” help us justify the belief that incest is wrong? No. It’s not the case that simply upon grasping the concept “incest,” we see that incest violates a moral value we hold dear. Consensual adult incest is defined as “consensual sex between two adults who happen to be close family members.” Does this concept necessarily imply a moral contradiction, whereby a value we hold dear is necessarily violated? Which value is in violation? Respect for autonomy? The principle of utility? The doctrine of the golden mean?

Huemer’s seeming-state theory

Next, consider Michael Huemer’s (2007; 2005) seeming-state account of “ethical intuitions.” Huemer’s account has two components. First, intuitions have a non-epistemic component. If S has an “intuition” that p, this just means it seems to S that p is true. Second,

intuitions have an *epistemic* component. For Huemer, if it seems to S that p, then, absent defeaters, S has some degree of justification for believing p. (This is basically Phenomenal Conservatism applied to ethics (Huemer, 2005).)

An important feature of Huemer's account is the "absent defeaters" clause in the epistemic component. For Huemer, ascribing some epistemic credibility to ethical intuitions doesn't have any outrageous implications because epistemic credibility is only ascribed when defeaters are absent. If any defeaters are *present*, then we should doubt our intuitions and seek independent, confirming evidence.

Huemer (2007) suggests several specific defeaters that we should look out for. You should doubt the reliability of your intuition that p if this intuition:

- (1) conflicts with another one of your intuitions,
- (2) has been influenced by your culture,
- (3) has been influenced by evolutionary pressures,
- (4) happens to support doing something that would promote your self-interests, or
- (5) is accompanied by strong emotion.

For any ethical intuition you might have, if any of (1-5) obtain, you are not epistemically justified in believing p on the mere basis that p seems true to you. Instead, you should doubt the reliability of your intuition and seek out confirming evidence.

In the case of the intuition "Incest seems wrong," defeaters 1, 2, 3, and 5 seem to apply. Regarding (1), my intuition that "Incest is wrong" conflicts with my other intuition that "An action is not wrong unless it somehow entails, intrinsically or in its consequences, some form of disvalue that I acknowledge, such as harm, inequality, injustice, the violation of consent, exploitation, dishonesty, disloyalty, disrespect, etc.—but incest doesn't necessarily involve any of these."

Regarding (2), my intuition that “Incest is wrong” seems influenced by my culture. Incest is a common taboo throughout many human cultures, but there is diversity regarding which particular relationships are forbidden. In contemporary America, for example, it’s socially forbidden (and often illegal) to marry a first cousin. But in other cultures, cousin marriages are permitted (e.g. Old Testament-era Jews). And in still other cultures, only “parallel” cousin marriages (marrying a cousin from a parent’s same-gender sibling) are forbidden, while “cross-cousin” marriages (marrying a cousin from a parent’s opposite-gender sibling) are permitted or even encouraged (e.g. this is the norm for the Dravidian peoples of South India) (Henrich, 2016). These cross-cultural variations in the incest taboo have no biological explanation. The best explanation is that cultures themselves influence whether their individual members develop an aversion to incest and, if so, what shape it takes.

Two defeaters being present is sufficient, I think, to show that Huemer’s ethical intuitionism won’t help justify the belief that incest is wrong, so for the sake of brevity, I won’t elaborate on any additional defeaters. However, as additional support, note that Huemer (2007) himself explicitly states that, on his theory, the intuition “Incest is wrong” is a probably an “unjustifiable prejudice,” given the presence of defeaters (especially defeaters 2 and 3 above), and thus we should doubt the intuition and examine “specific arguments” proposed by advocates of the taboo (p. 388). Of course, throughout this section (Section 5), we *have* examined specific arguments, and failed to find a good one.

5.7. Conclusion

In summary, this section considered a sampling of philosophical arguments for the claim “Incest is wrong.” Each is unpersuasive, I think. And while there are other arguments I did not consider, I’m not aware of any arguments that are *better* or *stronger* than those considered (and

refuted) above (though I'm open to hearing suggestions). Therefore, while I wouldn't say philosophers *cannot* establish the immorality of incest, I am comfortable saying that they as yet *have not*. Therefore, like ordinary citizens, philosophers are morally dumbfounded by incest.

As noted earlier, I have focused here on arguments against incest. But this choice was arbitrary, such that a different version of this chapter might have focused on a different member of the foul five. Therefore, I believe that philosophers would be morally dumbfounded with regard to each member of the foul five.

The upshot is this. If both ordinary citizens and philosophers are morally dumbfounded as to why the foul five are immoral, then it seems there is no argument floating around in the ether that legislators can utilize to retroactively justify current criminal bans on the foul five, from a standpoint of Liberal Legal Moralism. Liberal Legal Moralism holds that the state may outlaw immoralities as such, if they are sufficiently grave, but the Presumption of Liberty still obtains; the state cannot simply assert "Incest is gravely immoral" without elaborating, nor can it rely on appeals to majority assumptions. It must give an argument.

Of course, it has *not* given an argument as yet, and I predict that, if it tries, it will find doing so to be a daunting task.

6. Conclusion

This chapter has sought to motivate the following thesis: *imposing blanket criminal prohibitions on the foul five seems inconsistent with liberal principles*. This thesis is interesting, I think, because the foul five are, in fact, illegal in most U.S. states, often carrying harsh criminal penalties, such as thousands of dollars in fines and years in prison.

The chapter's main argument was this:

Main Argument

1. A policy is illiberal if it violates a basic principle of liberalism.
2. A basic principle of liberalism is the Presumption of Liberty (PoL).
3. Criminalizing the foul five violates the PoL.
4. Therefore, criminalizing the foul five is illiberal.

The argument is valid (the premises logically entail the conclusion). Are the premises true? It would seem so; premises (1) and (2) are obviously much more plausible than their denials. The premise most liable to be seen as controversial is premise (3). Hence, I have focused my attention on supporting (3).

The Presumption of Liberty (PoL) holds that citizens have a *prima facie* right to do what they want (especially in private), such that, if the state wants to ban conduct X, the burden of argument falls on the state to demonstrate a reason why banning conduct X is necessary. It seems to me, criminalizing the foul five violates the PoL if the state can't articulate a reason why criminalizing the foul five is necessary. This follows from the definition of the PoL (the burden of argument falls on the state to demonstrate a reason why banning conduct X is necessary).

As it turns out, it doesn't seem like there's any such reason that would be open to the state. After all, there are really only three different approaches we could take:

- 3a. We could say the foul five are *harmful*, thus meriting criminalization under the Harm Principle.
- 3b. We could say the foul five are *offensive*, thus meriting criminalization under the Offense Principle.
- 3c. We could say the foul five are *gravely immoral*, thus meriting criminalization under Legal Moralism.

But these approaches won't work. (3a) won't work because the foul five aren't inherently harmful, on any plausible definition of "harm." (3b) won't work because the Offense Principle doesn't permit state intrusions into the private sphere. And (3c) won't work because only a *liberal-compatible* form of Legal Moralism is plausible, but Liberal Legal Moralism still imposes on the state the burden of showing through argument that conduct X is actually

(gravely) immoral—a burden it will find difficult to meet, given that moral dumbfounding about the foul five is widespread among both the public and philosophers.

So, we have established premise (3) of the Main Argument—the last crucial piece I needed. Hence, the Main Argument goes through. Hence, banning the foul five is illiberal. And liberals should be opposed to illiberal policies. Hence, liberals should be opposed to banning the foul five—or so it seems at this stage.

6.1. *The upshot*

What is the relevance of this argument? The answer is that we can now see more clearly that we face a problem: *we have two conflicting ideas*. On the one hand, we cannot give a liberal argument for criminalizing the foul five (first idea). On the other hand, *that* the foul five ought to be criminalized is deeply intuitive (second). The tension between these ideas shows that there is philosophical work to do. Advocating for the decriminalization of the foul five sounds absolutely crazy, such that if liberalism provides no resources to justify laws against the foul five, that is a philosophical shortcoming. Thus, we need to find a way to provide normative justification for laws against the foul five, ideally from *within* a broadly liberal outlook.

I propose that such a justification may lie in an investigation of disgust, particularly as a moral or political emotion. As referenced at various points throughout this chapter, the foul five aren't just widely viewed as wrong; they're widely viewed as *sick*. In the psychological literature on moral dumbfounding, subjects struggle to articulate justification for their moral condemnation of the foul five, but they know one thing at least: *the foul five disgust them*. Indeed, at least for some of the foul five, criminal bans on them are explicitly rooted in the public's disgust. For example, state laws against bestiality explicitly characterize the act using disgust language (Posner and Silbaugh, 1996: 207-212).

Is it possible to give some sort of account on which disgust—particularly moral disgust—can justify laws, at least in some cases? The remainder of this dissertation pursues this question. The next chapter begins this project by surveying the scientific literature on disgust. This is a natural starting place because disgust is a real psychological state with an objective structure. If we're going to explore its potential normative significance, we should probably know a thing or two about what it is, how it works, and what it responds to. What is its characteristic psychology? What are its evolutionary origins? What types of moral offenses does it respond to? Do these moral disgust triggers vary culturally, and if so, why? However we assess disgust normatively should be informed by the empirical realities regarding these questions.

Chapter Two

To “Eww” Is Human: The Nature, Origins, and Moral Significance of Disgust

1. Introduction

This dissertation is about what role, if any, the emotion of disgust should play in criminal law. I am especially interested in whether disgust can justify criminal prohibitions against actions or practices that are widely viewed as disgusting by the general populace, but are otherwise harmless (arguably). This question becomes particularly salient when one realizes that disgust appears to be the *only* possible justification for certain criminal prohibitions, as is the case with the foul five (as per the argument of the previous chapter). If prohibitions on the foul five are justified—which seems intuitively obvious—and if disgust is the only possible justification for these prohibitions, it follows that disgust must be capable of doing justificatory work all on its own. As to exactly *how* this could be so, I aim to give a defensible account by the end of the dissertation.

The aim of the present chapter is to present a picture of the basic empirical character of disgust. I will do this by surveying scientific research on the emotion. Obviously, this literature is vast. My general strategy for navigating it will be to cover several major landmarks in classic disgust research, and also to chase down some important threads in contemporary research. The picture of disgust that will arise from this discussion will serve as an anchor for discussion of the normative significance of the emotion in the next two chapters.

Here is how the chapter unfolds. Section 2 discusses the basic character of disgust, including its characteristic inputs (items that trigger the emotion) and outputs (the emotion’s facial, behavioral, and neurophysiological signature). Section 3 discusses the evolutionary origins of disgust, with a focus on three prominent theories: Valerie Curtis’ infection-avoidance

theory, Paul Rozin's food-rejection theory, and Daniel Kelly's entanglement theory. Section 4 delves into the various "elaborated" domains of disgust. Researchers note that some disgust, known as "core disgust," serves a biological function, such as helping humans avoid infection or reject unsafe food. However, most disgust triggers are not things that can make you sick. Beyond core disgust, three additional domains are examined: animal-reminder disgust (disgust at objects or behaviors evocative of animality and mortality), interpersonal disgust (disgust at intimate contact with unfamiliar, unknown, or undesirable others), and moral disgust (disgust at moral transgressions). Section 5 covers research on disgust development in childhood. Specifically, I discuss the trajectory of development, as well as the mechanisms by which specific disgust triggers are enculturated. Finally, Section 6 discusses the influence of culture on disgust, particularly moral disgust. This section draws on insights from Jonathan Haidt's "Moral Foundations Theory" to explain differences in moral disgust judgments between Americans and Japanese people, and also between American liberals and conservatives. The chapter concludes with a summary of the preceding sections, which serves to set up a basic model of disgust.

2. What is Disgust?

In the movie *Indiana Jones and the Temple of Doom*, there is a scene where Indiana and his companions, Willie Scott and Short Round, have dinner at a palace somewhere in India. They all sit at a long table and enjoy (er, endure) a meal of exotic dishes, including beetles, snakes, and eyeball soup. Most sickening of all is dessert: "chilled monkey brains," served in hollowed-out monkey heads. When the lids are lifted, a red, gelatinous substance (the brains) is revealed, which the locals begin scooping out with spoons. This causes Willie to faint and fall out of her chair—overcome with disgust.

Simply put, disgust is a feeling of revulsion triggered by offensive or unpleasant stimuli. The *Oxford English Dictionary* contains a definition of disgust that dates back to 1598 and has remained steady over time. That definition is as follows:

Strong repugnance, aversion, or repulsion excited by that which is loathsome or offensive, as a foul smell, disagreeable person or action, disappointed ambition, etc.; profound instinctive dislike or dissatisfaction.

Scientific definitions of disgust align with the OED definition but expand on it to capture additional, more subtle features. The first modern scientific definition of disgust comes from Darwin's (1876) *The Expression of the Emotions in Man and Animals*, where it is defined as

something revolting, primarily in relation to the sense of taste, as actually perceived or vividly imagined; and secondarily to anything which causes a similar feeling, through the sense of smell, touch, and even of eyesight (p. 254).

Later, Angyal (1941) characterized disgust as follows:

The threat of disgusting objects to the individual is a very specific one. There is no clear notion present of a damage or harm which the disgusting object would inflict upon one's person. Disgust is directed rather against close contact with certain objects, implying mainly the fear of becoming soiled (p. 394).

More recently, Rozin and Fallon (1987), drawing on Angyal's definition, have defined disgust as "Revulsion at the prospect of (oral) incorporation of an offensive object," where "the offensive objects are contaminants; that is, if they even briefly contact an acceptable food, they tend to render that food unacceptable" (p. 23).

Disgust is often considered a "basic emotion" (I will define this term in a moment) (see Darwin, 1876; Ekman and Cordaro, 2011; Haidt, Rozin, and McCauley, 1994; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999, 2008). Darwin (1876) proposed disgust as one of six basic emotions (the others being anger, sadness, happiness, surprise, and fear). Paul Ekman has also identified disgust as a basic emotion. Initially, Ekman listed six basic emotions (the same six

as Darwin) (see Ekman, 1972). Later, he expanded his list to include several others (e.g. contempt, amusement) (see Ekman and Cordaro, 2011). Despite changes and additions to the list of basic emotions over the last century, disgust's status as a basic emotion has not been questioned (Rozin, Lowery, and Ebert, 1994).

What does it mean for an emotion to be “basic”? According to Ekman and Cordaro (2011), an emotion is basic if it meets two desiderata. First, it is “discrete,” meaning it can be fundamentally distinguished from other emotions by its facial, vocal, and physiological expressions, as well as by the events that characteristically precede it. Second, it has evolved to allow humans to react to fundamental life tasks, such as fighting, mating, escaping danger, mourning the death of family, and so on. Disgust meets both desiderata: it has fundamentally distinguishable expressions (its own facial expression, behavioral pattern, physiology, and phenomenology) and has evolved to allow humans to react quickly and automatically to everyday life tasks (the precise task for which disgust first evolved is disputed, but it probably involved avoiding infection and unsafe foods; see Section 3).

Disgust, qua a basic emotion, shows signs of universality (Angyal, 1941; Bloom, 2004; Curtis, de Barra, and Aunger, 2011; Darwin, 1876; Ekman, 1972; Rozin, Haidt, and McCauley, 1999, 2008). For example, Ekman (1972) provides evidence that the same “disgust face” (lowered eyebrows, wrinkling on the side and bridge of the nose, upper lip raised, lower lip raised and slightly protruding) is recognized by both Americans and Japanese people (see also Ekman and Friesen, 1971). Feces are a universal object of repulsion (Angyal, 1941; Bloom, 2004; Curtis and Biran, 2001; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999, 2008). Children as early as kindergarten are capable of matching the disgust face to disgust

stimuli, even before they have learned the word “disgust” and added it to their vocabulary (Aznar et al., 2023; Danovitch and Bloom, 2009).

2.1. The inputs of disgust

Disgust is triggered by a characteristic set of inputs. Common triggers include bodily wastes, rodents, bugs, rotting food, death, and poor hygiene. Various properties can also trigger disgust (stickiness, sliminess, moistness, squishiness, dampness) (Curtis and Biran, 2001). There is some variation among individuals and cultures regarding what is viewed as disgusting (for specific differences, see Curtis and Biran, 2001; Imada, Yamada, and Haidt, 1994). Rozin, Haidt, and McCauley (2008) write that “cultures differ about whether dogs are best friends or dirty scavengers, about whether or not corpses should be touched during mourning, and about whether mouth-to-mouth kissing is erotic or disgusting” (p. 766). Nevertheless, there is a striking amount of cross-cultural convergence on specific elicitors (Curtis and Biran, 2001; Nussbaum, 2004). Curtis and Biran (2001) interviewed subjects from India, Burkina Faso, the Netherlands, and the United Kingdom and found that feces, vomit, sweat, spit, blood, pus, open wounds, sexual fluids, certain animals (maggots, flies, cockroaches, rats, snakes), dead bodies, spoiled food, poor health, and immoral behavior were frequently named as sources of disgust by subjects in each country.

Disgust triggers are viewed as both offensive and contaminating (Angyal, 1941; Curtis and Biran, 2001; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley 1999, 2008). The nature or contact history of the item itself is the reason for rejection, not merely its properties (such as taste or texture) (Angyal, 1941; Haidt, McCauley, Rozin, 1993; Rozin and Fallon, 1987; Rozin, Lowery, and Ebert, 1994). When you ask people why they won’t eat a disgusting item (e.g. a cockroach), the answer typically references the nature or contact history of the item (“I will not

eat this cockroach because it is a cockroach”) (Rozin and Fallon, 1987). The nature or contact history of disgusting items makes them unclean, dirty, or debased (Angyal, 1941; Curtis and Biran, 2001; Rozin and Fallon, 1987). As such, the prospect of incorporating disgusting items (especially through the mouth) induces distress; the concern is that one will become debased oneself via incorporation (Rozin, Haidt, and McCauley, 2008). Disgusting objects are also contaminating, meaning if an innocuous item makes contact with a disgusting item, the innocuous item can be compromised. For example, most people refuse to drink from a glass of juice with which a (sterilized) cockroach has made brief and traceless contact (Rozin, Millman, and Nemeroff, 1986).

Disgust, being rejection based on offensiveness, differs from other types of rejection. This can be seen most readily in the context of food rejection. Rozin and Fallon (1984, 1987), in their work on food rejection, identify four types of reasons humans reject food items: (1) distaste, (2) danger, (3) disgust, and (4) inappropriateness. Distasteful items are rejected because of their properties (“This item tastes bad”). Dangerous items are rejected because of their consequences (“If I ingest this item, I will get sick”). Inappropriate items are rejected for lacking nutritional value (“I don’t eat paper because paper is not food”). In contrast, disgusting items are rejected because of their nature or contact history (“I will not eat this cockroach because it is a cockroach” or “I will not eat this cookie because a cockroach touched it”).

In eating, there are three “ingestion stages” (Rozin and Fallon, 1987): (Stage 1) in the environment, (Stage 2) in the mouth, and (Stage 3) inside the body. Distasteful items are rejected at Stage 2, but not Stages 1 or 3 (those who dislike broccoli are not distressed by broccoli being on the plate or inside the body). Dangerous items are rejected at Stages 2 and 3 (the prospect of poisonous mushrooms entering the mouth or body is bothersome), but not Stage 1 (poisonous

mushrooms in one's immediate vicinity are not bothersome). In contrast, disgusting items are rejected at all three ingestion stages. The presence of disgusting items in one's immediate vicinity is distressing, as is the prospect of the item entering the mouth or body.

Another point of contrast between distaste and danger, on one side, and disgust, on the other, is contamination/contact history. Distasteful and dangerous items are not viewed as contaminating. If a mushroom was once poisonous but has been detoxified, it is acceptable (Bloom, 2004). Likewise, if a piece of broccoli makes contact with a desirable soup and is then removed, the soup remains acceptable (it is not tainted by contact with the broccoli) (Rozin and Fallon, 1987). In contrast, disgusting items retain their offensive nature even when their distasteful or dangerous properties have been removed. For example, most people refuse to eat candied cockroaches (distasteful properties removed) and also refuse to eat cockroaches contained in indigestible capsules (dangerous properties removed). Additionally, disgusting items transfer offensiveness to otherwise acceptable items with which they have made contact, such as the glass of juice that has made brief and traceless contact with a sterilized cockroach.

2.2. The outputs of disgust

Disgust, when triggered, manifests in a set of characteristic outputs, including a distinct facial expression, physiological signature, behavioral pattern, and phenomenological feel (Angyal, 1941; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999, 2008). These outputs manifest quickly and automatically, without voluntary or conscious invitation, and continue until the expression process is complete.

The facial expression associated with disgust is often referred to as the “gape face” (Bloom, 2004; Rozin, Haidt, and McCauley, 1999; 2008; Rozin, Lowery, and Ebert, 1994). This name is a reference to the characteristic gaping of the mouth, as if preparing to vomit or expel

contents from the mouth. Darwin (1876) describes the disgust face as “the lower lip being turned down, the upper lip slightly raised, with a sudden expiration, something like incipient vomiting, or like something spit out of the mouth” (p. 16). Ekman (1972) notes that the facial expression of disgust involves activation of muscles around the brows, forehead, eyelids, nose, and mouth. Specifically, the brows are typically drawn down, the bridge of the nose is wrinkled, the lower eyelids are raised, the cheeks are raised, the mouth is opened, the upper lip is retracted, and the outer corners of the mouth are pulled down. Some authors emphasize that most of the facial muscles activated in the expression of disgust are concentrated around the mouth (Darwin, 1876; Rozin, Lowery, and Ebert, 1994).

Rozin, Lowery, and Ebert (1994) suggest that the disgust face comes in slightly different varieties, depending on which senses the disgust elicitor stimulates (taste, smell, touch) and what type of offense the elicitor represents (e.g. external threat to the body versus internal threat to the body). The nose wrinkle is most associated with bad smells; the gaping mouth is most associated with oral irritation; and the retracted upper lip is most associated with threats external to the body (e.g. objects of moral disgust like incest, hypocrisy, and sleazy politicians).

In terms of physiology, the most notable physiological response associated with disgust is nausea (Angyal, 1941; Oaten, Stevenson, and Case, 2009; Rozin, Haidt, and McCauley, 1999; Rozin, Haidt, and McCauley, 2008). In extreme cases, vomiting can occur. More generally, disgust is associated with arousal of the parasympathetic nervous system (Chapman and Anderson, 2013; Rozin, Haidt, and McCauley, 1999; Rozin, Haidt, and McCauley, 2008). While the sympathetic side of the autonomic nervous system (associated with fear and anger) prepares the body for immediate action (escaping predators, fighting), the parasympathetic side prepares the body for relaxation and recovery. Thus, subjects in the grips of disgust experience lowered

heart rate, lowered blood pressure, lowered breathing, lowered skin conductivity, and increased salivation (throat-tightening and fainting are also possible). These responses are not under voluntary control, but instead are regulated by the autonomic nervous system.

Disgust has also been associated with increased neural activity in a region of the brain called the anterior insula (Rozin, Haidt, and McCauley, 2008). This is supported by studies that use fMRI machines to observe the brain activity of subjects looking at disgusting images (see Schnaich, Lieberman, and Kiehl, 2008). Additional support comes from studies showing that subjects with damage to the anterior insula and basal ganglia have impaired ability to experience disgust or recognize the disgust face in others (Hayes, Stevenson, and Coltheart, 2007).

The behavioral response associated with disgust is known as “distancing” (Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999; 2008). Distancing involves creating space between the source of disgust and oneself, to alleviate distress. Distancing takes different forms, including (1) removing oneself from the disgust trigger (e.g. by walking away), (2) redirecting one’s attention (e.g. by changing the subject of conversation or closing one’s eyes), and (3) taking action to get the trigger away from oneself (e.g. by washing the hands or expelling offensive contents from the mouth) (Haidt et al., 1997; Rozin, Haidt, and McCauley, 1999). It is worth noting that three common synonyms for disgust are “repulsion,” “revulsion,” and “abhorrence,” each of which has etymological roots related to distancing behavior: “repulse” traces to the Latin *repellere*, meaning “to drive back”; “revulsion” traces to the Latin *revellere*, meaning “to pull away”; and “abhor” traces to the Latin *abhorreere*, meaning “to shudder away from.”

Finally, disgust is distinguished by a characteristic conscious experience or phenomenological feel. It is generally a negative emotion that is unpleasant to feel, and it can

range from mild to intense.²⁰ John Kekes (1992) provides examples of objects and situations that he predicts will evoke intense disgust in his readers (“eating feces,” “being splattered with someone’s brain,” “the sight and smell of a putrefying corpse”) and proceeds to describe the experience as follows:

To say that we are sickened ... is not a metaphor; we are nauseated, our guts heave, we tremble, our hearts race, we both sweat and shiver, we cannot control our physical responses, or we control them only with great effort. It is unthinkable that we could combine such experiences with the enjoyment of a good meal or flirtation, or that we would respond to them merely as annoying interruptions of an interesting conversation or good book. We regard these matters as fundamentally unsettling, their memories stay with us, they are the stuff of which nightmares are made, and we find it usually hard and often therapeutic to talk about them (1992: 433).

Thus, in Kekes’ description, disgust is negative, visceral (felt in the body), and can be intense.

2.3. *Disgust is meaning-laden*

It is important to stress that disgust is laden with meaning (Angyal, 1941; Curtis and Biran, 2001; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999, 2008). Defining disgust solely by its expressive, physiological, and behavioral features would be inadequate. After all, oral rejection is observed with dangerous food items, parasympathetic arousal (nausea, lowered heart rate) is associated with emotions other than disgust (e.g. grief), and neural activity in the anterior insula can happen in response to several negative stimuli (e.g. pain, fear, sadness). A central component of the experience of disgust that must be acknowledged is the interpretation of the trigger as *offensive* (dirty, soiled, etc.). Contact with disgust triggers is avoided specifically *because* they are perceived as dirty. Angyal (1941) draws attention to circumstantial evidence

²⁰ Well, disgust is *typically* experienced as negative. Still, disgust induced by artistic experiences *can* be pleasant. For example, there are people (like myself) who love horror movies, and horror movies characteristically contain monsters with disgusting properties (they’re slimy, gooey, sticky, decaying, etc.). Likewise, lots of people, and most children, enjoy “potty humor” (humor dealing with poop, pee, barf, and farts). This puzzling phenomenon (disgust induced by art can be pleasant) belongs to a broader family of puzzles regarding negative emotions in art being sometimes pleasant (e.g. the paradox of tragedy, the paradox of horror). See Carroll, 1990.

that disgust triggers are viewed as dirty: the nastiest words, insults, and curses that one can fling at others frequently invoke “excrements” and “those parts of the body that have some relation to excretion” (p. 397).²¹

Not only do disgusting items elicit a sense of offense, but the intensity of this sense increases with more intimate contact (Angyal, 1941; Rozin and Fallon, 1987; Tybur et al., 2013). Angyal (1941) notes that

There is already some degree of unpleasantness in having disgusting objects in one’s immediate surroundings, and more so if they soil one’s clothes. It is even more disgusting to touch them with one’s bare skin, and very much more so to take them into the mouth, not to mention ingesting them (p. 394).

Thus, the intensity of disgust increases with the level of intimacy: vicinity (less intense), skin (more intense), mouth/ingestion (extremely intense). Many authors note that the most intense sensation of disgust is associated with oral incorporation (Angyal, 1941; Darwin, 1876; Rozin and Fallon, 1987). For example, Kelly (2011) writes that “Feces are disgusting enough; imagining eating them is downright vile.”

Why exactly does contact with (or oral incorporation of) a disgusting object cause distress, and why does the level of intimacy affect intensity? Most people refuse to eat candied cockroaches, despite the fact that the item has no distasteful or dangerous properties. What would happen if you *did* eat the cockroach?

Some authors suggest that a partial answer may be captured in the phrase “you are what you eat” (Nemeroff and Rozin, 1989; Rozin, Haidt, and McCauley, 1999, 2008). Paul Rozin (among others) has argued that disgust is an emotion that serves as guardian of both the body and “the soul” (he often refers to disgust as “the body and soul emotion”) (Rozin, Haidt, and

²¹ Examples I can think of: “shithead,” “piece of shit,” “bullshit,” “eat shit,” “pisshead,” “asshole,” “buttface,” “dickhead,” “cunt,” “pussy,” “asswipe,” and “douchebag.”

McCauley, 1999, 2008; see also Bloom, 2004; Haidt et al., 1997). The mouth—which is central to many definitions of disgust and is the area around which most of the muscle activation involved in the disgust face is concentrated—is viewed as the gateway to the body (Rozin and Fallon, 1987). Literally, the mouth is the critical transition point from one’s “outsides” to one’s “insides”; the face is made of ectodermic skin, while the lips are made of endodermic skin (Rozin and Fallon, 1987). But more than this, Rozin theorizes that people conceive of themselves as embodied beings, with the self (or personality, or moral character) and the body somehow sharing an important link. As such, physical substances incorporated into the body can transfer properties to the non-physical self.

The principle “you are what you eat” is overtly accepted in many traditional cultures, such as the Tairora of Papua New Guinea and Jewish thought on the kosher dietary system (Nemeroff and Rozin, 1989). There is also evidence that Western people accept the principle, albeit in an unacknowledged way. For example, Nemeroff and Rozin (1989) found that American college students associate more aggression and hairiness with people who eat wild boar than with people who eat wild turtles. This finding is a bit hard to believe, but it gets the idea of the “you are what you eat” principle across: people who eat cockroaches are cockroachy, people who seek sexual pleasure from animals are bestial, etc.

Of course, in literal terms, the “you are what you eat” principle is obviously false. Eating a cockroach will not make you grow antennae or mandibles. Does this mean disgust is irrational? Some authors go there (for example, see Oaten, Stevenson, and Case, 2009). Rozin sometimes describes disgust, and especially the perceived contamination of neutral items by contact with disgusting items, as a form of “magical thinking” (see §4.1 below). However, it may not be that people in the grips of disgust literally believe in magic. The sense of offensiveness involved in some varieties of disgust (interpersonal, moral) may be underpinned by a kind of symbolic

thinking (see Nemeroff and Rozin, 1994). Subjects in some studies report that they would refuse to put on a sweater previously worn by Hitler because doing so would somehow symbolically affiliate them with Hitler (Nemeroff and Rozin, 1994; 1990). The claim that disgust may engage a deep, perhaps unacknowledged, kind of symbolic thinking is further supported by the observation that in the course of child development, disgust does not appear until after the ability to engage in abstract thought develops (Rozin and Fallon, 1987).

The offensiveness that humans associate with disgust triggers adds a layer of meaning that is distinctively human, not shared by animals. On the one hand, human disgust has a number of corollaries in animals. For example, many animals experience distaste (rejection of unpleasant tastes and smells), and the face they make in response to distasteful stimuli resembles the nose wrinkle and mouth gape seen in the human disgust face (see Figure 2 for an example). Mice have been found to avoid contact with sick conspecifics (Boillat et al., 2015). And rats avoid ingesting foods that have made them sick or nauseous in the past (Garcia et al., 1955). However, despite these resemblances, human disgust is unique in a number of ways—perhaps most obviously in that disgust objects are mentally interpreted as offensive. Animals don’t find distasteful objects *offensive*. (The uniqueness of human disgust is further discussed in Section 3.3.)

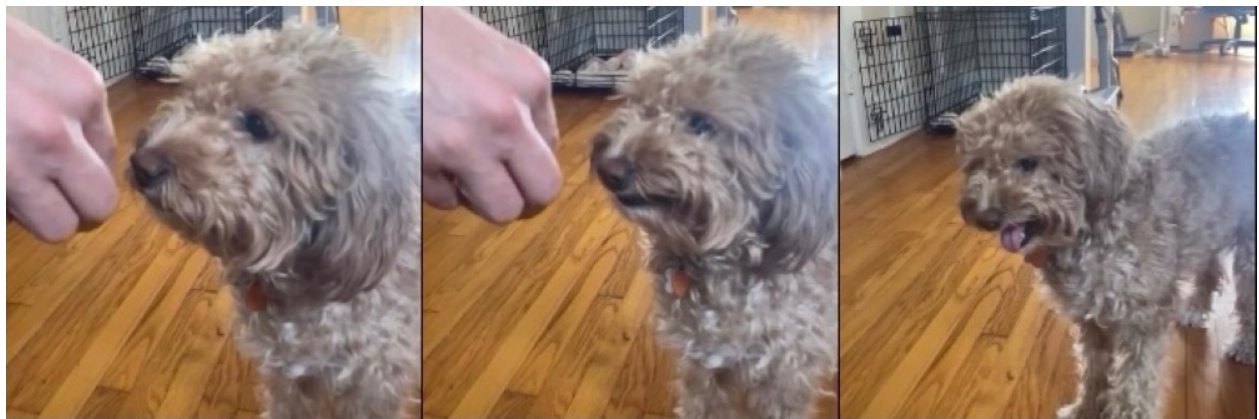


Figure 2. My dog, Kona, reacting to the smell of a stink bug crushed in a paper towel.

Photograph by Luke Tucker.

3. The Origins of Disgust

So, disgust has a highly recognizable shape. It is a basic emotion with signs of universality; it is a negative emotion associated with rejection and avoidance; and while the inputs that trigger it vary slightly among individuals and cultures, the outputs by which it is expressed are strikingly consistent. This consistency and universality demand explanation. Why does disgust have the shape that it has? What function does it serve in human life?

3.1. *Infection-avoidance*

Here, much insight can be gained from adopting the perspective of evolutionary theory. It seems likely that disgust acquired its current shape through a selection process (i.e. it was selected for because it served an adaptive function). The strongest evidence for this is twofold. First, the universality of disgust suggests that it is part of humans' evolutionary lineage—a feature of our psychology present in ancestors before differentiation across local circumstances appeared. People facing dramatically different environmental and cultural conditions respond to strikingly similar disgust inputs and exhibit suspiciously similar outputs. Second, the preconscious, involuntary nature of disgust suggests it has deep roots in our biology. As noted, the disgust response encompasses involuntary physiological arousal, specifically of the parasympathetic side of the autonomic nervous system (Rozin, Haidt, and McCauley, 2008). Thus, a major component of the response is, like with the respiratory and circulatory systems, unchosen and automatic. Moreover, there is evidence that, when it comes to disgust, the tail wags the dog; people feel disgust first and come up with reasons for their feelings second (if they can at all), as illustrated by the dumbfounding phenomenon (Haidt, 2001). And disgust can be stubborn, often persisting even after one has become convinced that an item is not toxic or dangerous. For example, Bloom (2004) notes that “during World War II, American pilots in the

Pacific went hungry because they refused to eat insects and toads, even though they had been explicitly taught that these foods were safe” (p. 68).

One theory is that disgust originally evolved as a sort of “behavioral immune system” to guard against infection (Curtis and Biran, 2001; Curtis, de Barra, and Aunger, 2011; Oaten, Stevenson, and Case, 2009; Tybur et al., 2013). Humans are, of course, equipped with *internal* immune systems. Specifically, the gut contains a complex microbiome inhabited by a delicately calibrated and balanced population of harmless (and some not-so-harmless) microorganisms. These microorganisms, to which the gut is acclimated, keep harmful intruders out by preserving ecological space for only themselves. In other words, they fight not for our health but for their own territorial supremacy, protecting us from more harmful would-be inhabitants by accident. This gut-based immune system is known to have been shaped by a selection process; just as pressure from macro predators (lions) drives adaptation in prey species (gazelles), pressure from microscopic threats (parasites and pathogens) has shaped the human internal immune system (Curtis and Biran, 2001). But this internal system is not bulletproof. Humans do lots of extra stuff to augment our immunity behaviorally (Strohming, 2014). The use of spices in cooking to combat microbes is an example of this; cultural affinity for chili peppers (and other spicy foods with anti-microbial properties) increases with proximity to the equator, where warmer temperatures means more microbes (Henrich, 2016). Disgust also falls into this category. By internalizing anxiety about likely sources of infection, we improve our survival odds (Tybur et al., 2013).

Valerie Curtis has argued that once the peculiar list of inputs and outputs of disgust are viewed through the lens of infection-avoidance, “the pattern falls into place” (Curtis and Biran, 2001: 28) (see also Curtis, 2013; Curtis, Aunger, and Rabie, 2004; Curtis, de Barra, and Aunger,

2011). On the inputs side, the most common disgust elicitors are notorious disease sources. For example, Curtis and Biran (2001) note that feces, a universal disgust elicitor, is the source of ascariasis, hepatitis A and E, polio, amoebiasis, hookworm, pinworm, whipworm, cholera, tetanus, and over 20 known bacterial, viral, and protozoan causes of intestinal tract infection. Other common disgust triggers—blood, vomit, snot, saliva, flies, rats, hair, spoiled food—are also rich with infection potential. Relatedly, Curtis, Aunger, and Rabie (2004) found that subjects systemically rated images of “infection-relevant stimuli” as more disgusting than similar images of “infection-irrelevant stimuli.” For example, a white towel with a reddish-yellow stain (infection relevant) was rated twice as disgusting as a white towel with a blue stain (infection irrelevant) (see Figure 3 for a replica); a louse (infection relevant) was more disgusting than a wasp (infection irrelevant); and an inflamed, pus-emitting skin lesion (infection relevant) was more disgusting than a clean skin burn (infection irrelevant).



Figure 3. Replica of the towel stimuli used by Curtis, Aunger, and Rabie (2004). Photograph by Luke Tucker.

The infection-avoidance theory also has explanatory appeal when it comes to the outputs side of disgust. Curtis and Biran (2001) note that parasites and pathogens target four entry points

of the body—the airways, the gut, the skin, and the genitals—and these are the areas with regard to which contact with disgusting stimuli arouses the most anxiety. As previously noted, the mouth is frequently reported as the site of the most intense disgust experiences, more so than (e.g.) the hands. This could be because the mouth is doubly vulnerable, being closely connected to two infection-vulnerable locations: the airways and the gut.

The infection-avoidance theory can also help explain certain puzzles observed about disgust. For example, several authors have noted that rotting meat is a stronger source of disgust than rotting vegetables (Angyal, 1941; Rozin and Fallon, 1987). Curtis explains this disparity by reference to the fact that rotting meat has a higher concentration of pathogens than do rotting vegetables (Curtis and Biran, 2001). It has also been observed that interpersonal affinity reduces disgust. Individuals are less disgusted by their own poop than others' poop; mothers are less disgusted by changing their own infant's diapers than those of other infants; and sharing a toothbrush with one's own spouse or best friend is less disgusting than sharing with a stranger (Curtis, Aunger, and Rabie, 2004). This too can be readily explained by the infection-avoidance theory: the further one goes outside of one's inner circle, the risk of encountering novel pathogens increases. Whatever pathogens are in one's own poop, or one's infant's poop, or one's spouse's or best friend's saliva, one has likely already been exposed to.

Its virtues notwithstanding, the infection-avoidance theory has limitations. For example, Curtis herself notes that certain disease vectors are not widely viewed as disgusting, such as the mosquito (see Curtis and Biran, 2001). According to the U.S. Centers for Disease Control and Prevention, the mosquito is the “world's deadliest animal,” killing more people annually, through transmission of malaria and other diseases, than any other creature on earth. Despite this,

in areas where malaria-infected mosquitoes are common (e.g. India and Africa), mosquitos, albeit deeply annoying, are not widely viewed as disgusting (Curtis and Biran, 2001).

3.2. *Food-rejection*

Paul Rozin agrees that the evolutionary history of disgust was partly shaped by pressure from infection threats. However, he stresses that its earliest origins trace to the rejection of unsafe foods. It is no coincidence, says Rozin, that the etymology of “disgust” traces to the Latin *disgustus*, meaning “bad taste” (Rozin, Haidt, and McCauley, 1999, 2008). Nor is it a coincidence that, in humans, the distaste and disgust faces are the same (both involve activation of the same muscles around the mouth and nose) (Rozin, Haidt, and McCauley, 1999). Disgust, says Rozin, is likely an example of the evolutionary concept of “preadaptation.”

Preadaptation happens when an organism has a system that was initially meant to serve one function, and evolution later co-opts it to serve an additional function. A classic example of preadaptation in humans is the mouth and lungs being used for speech (Rozin, Haidt, and McCauley, 2008). In speech, the mouth and lungs are teammates: the lungs provide air to produce sound, which the mouth and tongue manipulate to enunciate words. This system might seem expertly designed for speech, but the mouth and lungs originally developed for distinct purposes (the lungs for respiration, the mouth and tongue for handling food). Speech only developed out of convenience after these systems were already available. In the same way, disgust (it can be argued) developed as a preadaptation from distaste, a form of food rejection based on the properties of food items (taste, texture) (Bloom, 2004; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999, 2008).

The clearest evidence that disgust’s earliest origins trace to food rejection comes from looking at its outputs. For example, the disgust face involves activation of muscles in the mouth and nose, both of which appear aimed at rejecting ingestion: the nose wrinkle tightens the nasal

passage, preventing bad smells from entering, while the mouth gape prepares for spitting or vomiting (expelling contents from the mouth or stomach) (Rozin, Haidt, and McCauley, 2008). Darwin (1876) notes that the disgust face is characterized by “movements round the mouth identical with those preparatory to the act of vomiting” (p. 258), and further observes that infants, when experiencing distaste, reflexively express it by shaping their lips to allow “the contents to run or fall out quickly” (p. 364). Also, disgust is associated with neural activity in the anterior insula, the brain region associated with the subjective feeling states of hunger and taste (Rozin, Haidt, and McCauley, 2008). Also, the felt intensity of disgust is affected by the level of intimacy of contact with the elicitor, with the site of the most intense anxiety being the mouth, the site of food intake.

As further support, the classic physiological output of disgust—nausea—also points to food rejection. Nausea is key in the development of food aversions in animals and humans. In the 1950s, psychologist John Garcia discovered that rats reject foods associated with nausea. Garcia et al. (1955) presented rats with a novel food, and then artificially induced nausea in them (by either drugging them or shooting them with X-rays). Result: the rats developed an aversion to the novel food. This effect came to be known as the “Garcia effect.” The Garcia effect is also observed in humans; people who have bad experiences with a particular food item generally develop a lasting aversion to it (Oaten, Stevenson, and Case, 2009).

The Garcia effect is potent. Conditioning an animal to have a “Pavlovian response” (e.g. salivation) to an otherwise neutral stimulus (e.g. the sound of a bell ringing) requires several trials. In contrast, the Garcia effect can be obtained in one trial. Animals that eat a novel food and become nauseated often develop a food aversion from that one experience (nausea-induced aversion to familiar foods can be conditioned, but the effect is less strong). Humans are the same

way. People who eat sushi and get sick often become lifelong sushi haters (Oaten, Stevenson, and Case, 2009).

These food aversions are stubborn, too. They don't require conscious thinking to develop. If you eat sushi and then get nauseous from riding a roller coaster, you may develop a food aversion to sushi, even if you know it was the roller coaster, not the sushi, that made you sick.

Additionally, while aversions are easy to acquire, they are relatively hard to shake (Rozin and Fallon, 1987). Once an aversion is conditioned, it resists revision through reconditioning, and it is comparatively more resistant to reconditioning than fear- or distress-related aversions (Olatunji, Forsyth, and Cherian, 2007). In other words, one bad experience with sushi will be more memorable than two or three good ones.

There is increasing evidence that a special relationship exists between disgust and memory. Disgusting experiences are particularly memorable, more so than experiences involving other emotions (fear, sadness). People remember disgust-related words better than fear-related or neutral words (Ferré, Haro, and Hinojosa, 2018); people remember disgusting stories better than scary or neutral stories (Heath, Bell, and Sternberg, 2001); and people remember disgusting images better than scary or neutral images (Chapman, 2018; Chapman et al., 2013).

For example, Chapman et al. (2013) showed subjects images of disgust stimuli (cockroaches, feces), fearful stimuli (riots, animal threats), and neutral stimuli (coffee makers, books) and then evaluated subjects' recall of the images after intervals of 10 minutes, 45 minutes, or one week. They found that recall of disgusting images was "significantly higher" than recall of both fearful and neutral images. Moreover, the disparity between recall of disgusting and fearful images increased as the interval got bigger; that is, subjects recalled

disgusting images better than fearful ones after 10 minutes, but after one week, the superiority of their memory of disgusting images was even more pronounced.

The memory advantage associated with disgust has been found in children as well as adults, men as well as women, and has been observed across different types of memory tasks (Scheinle et al., 2021).

There is some uncertainty in the literature regarding why disgust enhances memory. Is it because (1) disgust itself enhances memory, or (2) some other factor (e.g. attention, arousal) enhances memory, and disgust enhances this factor? Some authors seem to prefer the second explanation (for example, see West and Mulligan, 2021). But Chapman et al. (2013) supports the first; this study controlled for attention, arousal, and several other factors, and disgust continued to enhance memory. In contrast, when controlling for arousal and attention, fear ceased to enhance recall altogether. The authors note that disgust's memory advantage "may depend on specialized neural mechanisms that cannot be easily understood in terms of general memory principles" (Chapman et al., 2013: 1109).

All these points—the quick, subconscious, and rigid nature of the Garcia effect and the memory advantage associated with disgust—make sense in light of what some researchers call the "omnivore's dilemma" (Bloom, 2004; Curtis and Biran, 2001; Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 1999, 2008). The omnivore's dilemma refers to the idea that humans, at some point in our evolutionary history, stopped eating only plants and started experimenting with meat. When this transition occurred, a new wave of microscopic threats was introduced. For this reason, humans were forced to balance curiosity about new foods (to maximize calories) with a healthy dose of caution (to avoid novel pathogens and parasites). Since humans live in very different environments, what good food looks, smells, or tastes like

isn't encoded into our genes; we have to figure it out by trial and error. Given these conflicting goals (curiosity and caution), combined with the life-or-death stakes, it makes sense that disgust and memory would have a particularly close relationship. In short, the peculiar outputs of disgust, when viewed as a solution to the omnivore's dilemma, start to make sense.

3.3. *The entanglement thesis*

A third evolutionary origin story for disgust, proposed by Daniel Kelly (2011: Chapter 2), holds that the food-rejection system and the disease-avoidance system are both central features of disgust, but neither is more central than the other. Rather, these systems developed independently, as solutions to independent problems (unsafe foods and transmissible sickness, respectively). However, through evolutionary happenstance, they became entangled in humans, and it was only once these systems became entangled that what we now call “disgust” in humans truly emerged. The conditions were ripe for this entanglement to happen, says Kelly. After all, upon becoming omnivores, humans incurred greatly increased poison- and pathogen-risks, and this would have demanded a proportionate solution.

This “entanglement thesis,” as Kelly calls it, could explain some otherwise puzzling facts that we observe about disgust. For example: (1) We gape and feel nauseous—two ingestion disincentives—towards things we would never eat anyway (e.g. festering wounds, or people with a hacking cough). (2) We physically distance from, and feel viscerally offended by, the prospect of touching certain objects that it would be sufficient to simply not eat (e.g. moldy cheese). The entanglement thesis readily explains these two observations: *disgust is a kludge*²². Two distinct

²² By “kludge,” Kelly (and other researchers) mean that a system is being used to serve a function for which it was not designed and, with no additional modifications, is not well-suited (or even quite *poorly* suited) to serve—like using a wrench to hammer down a nail, or using a knife to cut a piece of paper.

systems—one for food-rejection and one for infection-avoidance—became entwined upon humans’ becoming omnivores and incurring increased poison- and pathogen-risks.

By the way, the entanglement thesis could go some way towards explaining an observation that anyone whose dog eats its own feces has undoubtedly made: *humans alone in the animal kingdom feel disgust*. As previously noted, we share distaste with animals (if you let your dog smell a stink bug, they will reflexively gape). Also, some animals avoid contact with sick conspecifics. However, only humans both refuse to ingest and refuse to touch many of the same objects, and of course only humans feel *offended* by said objects. This is because only humans traversed an evolutionary path that saw adaptive pressures merge the food-rejection and disease-avoidance mechanisms.

4. Beyond Biology

At this point, you might be thinking “Case closed, Darwin explained disgust.” But not so fast. When it comes to understanding disgust—its inputs and outputs, its peculiar psychology—focusing on biological selection only explains so much.

Some researchers suggest it is necessary to separate disgust inputs into two broad categories: “core” and “elaborated” (Haidt, McCauley, and Rozin, 1994; Rozin, Haidt, and McCauley, 1999, 2008). Core disgust triggers are related to disgust’s biological function of protecting us from infections and unsafe foods. But most disgust triggers *do not* fall into this category. For example, only about a quarter of things that gross out Americans and Japanese people are “core” triggers (Imada, Yamada, and Haidt, 1993; see also Rozin, Haidt, and McCauley, 2008). The remaining triggers—which researchers call “elaborated” triggers—represent extensions beyond disgust’s original biological function. In simple terms, if you walk into an international airport—as Curtis and Biran (2001) literally did—and start asking people to

name things that gross them out, most of the answers you get will be things that cannot make anyone sick: putting ketchup on ice cream, betrayal by a friend, deviant sex, menstruation, foul language, deformity, amputees, eating one's pet dog, nose picking, seeing someone chew with their mouth open, used clothing, drunkenness, smokers, overweight people, homeless people, liberals, conservatives, politicians, hypocrites, cheaters, traitors, animal cruelty, and being rude to old ladies (Curtis and Biran, 2001).

Many individual and cultural aversions to specific foods are not products of core disgust. Several authors note that most human cultures are extremely selective about meat (Angyal, 1941; Bloom, 2004; Rozin and Fallon, 1987; Fessler and Navarrete, 2003; Rozin, Haidt, McCauley, 2008). Of all the millions of known animal species on earth, humans eat a tiny fraction. In Western cultures, almost all insects, invertebrates, reptiles, and amphibians are rejected, and so are most mammals (Rozin and Fallon, 1987). For example, in the U.S., there are around 400 mammal species, and Americans only eat two, maybe three (cow, pig, lamb). Bat, squirrel²³, possum, beaver, mole, coyote, manatee, seal, bear, raccoon, gopher, mouse, cougar, and chipmunk are off the menu, despite being nutrient rich.²⁴ Also off the menu are most *parts* of acceptable animals: skin, fat, guts, genitals, brains, blood, eyeballs, lips, tongues, noses, and ears. We stick mainly to skeletal muscle.

²³ About a year ago, my brother and I went fox squirrel hunting in North Texas. We bagged one, cooked it in a crockpot according to a recipe we found online, and ate it. The taste and texture were not all that different from other, more familiar meats. Nevertheless, my brother gave up after a few bites, and I gave up not long after. Despite assuring each other that the meat was totally safe, and despite all the mental preparations we had undergone planning for the hunt, when push came to shove, we found the idea of eating squirrel off-putting. Later, I told my mother-in-law about this experience and she was appalled—disgusted—that I had even contemplated eating squirrel.

²⁴ Interesting social dynamics can come into play here, given cultural variation in meat taboos. For example, in the U.S., the common carp is widely viewed as a “rough fish” (also trash fish or dirty fish), not fit for eating, on account of being too bony and muddy-tasting. Many U.S. state agencies use the term rough fish—which is derogatory—to refer to carp, including the Texas Department of Parks and Wildlife and the Oklahoma Department of Wildlife Conservation. Meanwhile, the common carp is widely accepted as a food fish in Eastern Europe and Southeast Asia. It is easy to imagine a case in which an individual immigrates from Eastern Europe, or Southeast Asia, to the U.S. and feels embarrassed or unwelcome for eating carp—i.e. they have eating habits viewed as disgusting by the cultural mainstream.

Picky eating is not exclusive to people in developed nations with Netflix subscriptions and mini Purell bottles in their purses; hunter-gatherers are also highly selective about meat. For example, the !Kung people of southern Africa reject rodents, insects, and carnivores (Rozin and Fallon, 1987; Rozin, Haidt, McCauley, 2008).

Food taboos vary by culture, but there are noticeable patterns: monkeys, pets (dogs, cats), scavengers, and mammalian carnivores (though not carnivorous fish) are frequently taboo (Rozin and Fallon, 1987). Meanwhile, mammalian herbivores are least likely to be taboo. What explains these patterns? For some, fairly straightforward explanations are available. Carnivores, being animal-eaters, produce putrid-smelling feces, and this makes eating them off-putting (Angyal, 1941; Rozin, Haidt, and McCauley, 2008). Similarly, scavengers are considered gross because they themselves eat disgusting stuff (feces, dead animals). So, it could be that carnivores and scavengers are off-putting as food sources because their own eating habits cause humans to associate them with core disgust triggers.²⁵ As for other food taboos, the explanation is less obvious. Rozin and Fallon (1987) suggest that eating monkeys is gross because they resemble humans too closely, and eating pets is gross because they are treated as family. Of course, these explanations just raise further questions; why should eating family, or things that look human, be gross?

While picky eating may seem funny, it has implications of potentially global significance. Some authors have recently noted that adopting an insect-based diet would contribute to the push toward environmental sustainability (Plakias, 2025; Huis and Rumpold, 2023; Ruby, Rozin, and

²⁵ Two additional theories for carnivores: (1) The taste and texture of mammalian carnivore meat may simply be generally inferior to that of herbivores (e.g. carnivores are generally leaner, and leaner meat is tougher to chew), and gustatory pleasure may have factored in ancestral hunters' prey choices (Koster et al., 2010). (2) Mammalian carnivores may generally be richer in zoonotic diseases (diseases that can transmit from animals to humans), insofar as they accumulate the zoonotic diseases of their prey (Han et al., 2016). Primates, however, are not generally richer in zoonotic diseases than other animals, so this theory would not explain the taboo on primates.

Chan, 2015). Livestock production is responsible for a significant chunk of annual global greenhouse gas emissions. Meanwhile, edible insects technically have the same nutritional value as animal meat pound-for-pound (maybe more), and farming insects would both slash emissions and boost tree and plant health. If citizens in developed, high-emissions nations (like the U.S.) could convince themselves to switch to an insect-based diet, it would greatly benefit the environment. Unfortunately, the primary practical obstacle standing in the way of this policy gaining traction is disgust (Ruby, Rozin, and Chan, 2015). Huis and Rumpold (2023) note that when consumers are informed of the “adequate nutritional value and low environmental impact” of insect-consumption, “This information is often not enough to make consumers less sceptical about consuming insects” (p. 7). In other words, it’s unlikely that Americans will stop eating Big Macs and start eating bugs.²⁶

4.1. Disgust involves magical thinking

By rejecting various meats and edible insects, we are leaving good calories off the table. This seems maladaptive (Rozin and Fallon, 1987). Accentuating this appearance of maladaptation is the fact that disgust is triggered not only by taboo foods but also by a whole network of otherwise neutral items that become offensive via “psychological contamination.”

Psychological contamination refers to a phenomenon whereby a secondary object comes to be viewed as offensive as a result of contact or association with a primary disgust object, whether or not the secondary object is actually rendered dangerous (e.g. poisonous) or distasteful (e.g. bitter) by this contact or association (for general discussion of this phenomenon, see Rozin

²⁶ Disgust towards eating bugs has been recruited into culture war driven politics. In a recent NPR podcast, journalist Huo Jingnan explains, “Although the idea is marginal, various right-wing media outlets use insect eating as a punchline to mock the climate movement.” <https://www.npr.org/2023/03/31/1167550482/how-a-conspiracy-theory-about-eating-bugs-made-its-way-to-international-politics>

and Fallon, 1987; Rozin, Haidt, and McCauley, 1999; 2008). An example of psychological contamination is when a glass of juice makes brief and traceless contact with a sterilized cockroach and is thereafter rejected. Rozin and Fallon (1987) write,

in all cultures there are some substances that can generate psychological contamination by physical contact. Although it probably has some adaptive value as protection against microbial (physical) contamination, psychological contamination seems on balance to be maladaptive, in that in most instances it motivates rejection of nutritive substances (p. 29).

The qualifier “psychological” is included because secondary disgust objects are often not literally contaminated. There are no literal traces of the cockroach in the juice, and subjects may even be convinced of this, but it makes no difference to their rejection or negative appraisal of the juice.

Psychological contamination comes in two varieties: contagion and similarity (Rozin and Nemeroff, 1990). Contagion refers to contamination by contact. Examples of contagion include: many people reject juice that has made contact with a cockroach, refuse to wear Hitler’s sweater (even after it has been washed with detergent and boiled in water), refuse to eat soup that has been stirred with a sterilized flyswatter, and refuse to eat from a sterilized bed-pan (Rozin and Fallon, 1987; Rozin, Markwith, and McCauley, 1994; Rozin and Nemeroff, 1990). In all these cases, a secondary object made contact with a primary disgust object, thus rendering the secondary object offensive, despite the absence of any literal contamination.

Next, psychological contamination can happen by similarity. If a neutral item resembles a primary disgust object in some way, it can become offensive. For example: many people reject gourmet chocolate shaped like dog doo, refuse to place plastic vomit from a comedy store between their lips, and avoid blowing their nose with toilet paper (Rozin and Fallon, 1987). In all

these cases, people reject or avoid a secondary item that superficially resembles a disgusting item, despite the secondary item not being literally dangerous or distasteful itself.²⁷

Rozin and Fallon (1987) point out that psychological contamination is theoretically limitless. That is, if primary disgust items can render secondary items offensive by contact or similarity, then seemingly no secondary item is safe. We breathe the same air as others; at restaurants, we use utensils that may have been in someone else's mouth hours or minutes before; we use public toilets that have been touched by strangers' butts; and so on. Cultures apparently overcome this (potential) problem by setting (arbitrary) limits on perceptions of contamination via customs and manners.

4.2. *Animal-reminder disgust*

So, not all disgust is core disgust. Some disgust has also become more “elaborate,” extending via psychological contamination to secondary objects associated, physically or imaginatively, with primary disgust objects.

But disgust is even more complex than this. Not all disgust objects are merely associated with core disgust objects. Some have no connection to core disgust at all. Paul Rozin proposes four distinct “disgust domains”: (1) core, (2) animal-reminder, (3) interpersonal, and (4) moral (Rozin, Haidt, and McCauley, 1999, 2008). Core triggers tend to be physical substances that are notorious sources of pathogens or parasites. In contrast, animal-reminder disgust is triggered by things that remind us of our animal nature (e.g. viscera, death, sexual fluids). Rozin explains that we reject these reminders because we are uncomfortable with our animal nature and prefer not to

²⁷ I tried this experiment with the students in my introduction to philosophy class. One day, I brought a large bag of barbecue-flavored chips and asked if anyone wanted some. Several students said yes. I then mentioned that it would be unsanitary for everyone to reach into the bag, and suggested we find a better way to distribute the chips. I then pulled out a brand-new roll of dog poop bags that I use to clean up after my dog, Kona, and began placing chips into these bags. When I asked again who wanted some, the girls in the class responded with a chorus of “ewws.” Two boys, however, were still interested.

think about or focus on it (Rozin, Haidt, and McCauley, 1999, 2008). Interpersonal disgust is directed at people, limiting our intimate contact with others. The further someone is from our “inner social circle,” the more offensive intimate contact with them may feel (e.g. strangers, people of lower social class). Finally, moral disgust arises in response to certain moral transgressions, such as hypocrisy and cheating.

According to Rozin, animal-reminder disgust encompasses four subdivisions of disgust triggers: death, poor hygiene, deviant sex, and body envelope violations (Rozin, Haidt, and McCauley, 1999). Basically, if you ask people in an international airport what sorts of things disgust them, you can bet they will mention things related to these four areas. For death, they may say the smell of death, seeing a human corpse, or touching cremated ashes; for deviant sex, they might mention incest, bestiality, or minority sexual practices such as BDSM or homosexuality; for poor hygiene, they might say smelly feet, bad breath, or body odor; and for body envelope violations, they might say gore, seeing people with amputated limbs, or watching depictions of surgery in movies.

What do death, poor hygiene, deviant sex, and body envelope violations have in common? Answer: they are evocative of the fact that humans are deeply vulnerable and, more importantly, mortal. Rozin, Haidt, and McCauley (2008) suggest that it is no coincidence that research has found that disgust sensitivity and death anxiety are positively correlated. We are the only animals on earth cursed with the knowledge that we will one day die, and this knowledge can be a debilitating source of anxiety. We are uncomfortable with being animals and do not like being reminded of or focusing on it. Consequently, we psychologically reject and avoid such reminders. Body envelope violations reveal our bloody, veiny, squishy insides, which resemble

those of animals, so we look away; corpses remind us of death, so we look away; sex and menstruation remind us of our unwilling dependence on carnal reproduction, so we look away.

On a cultural scale, humans establish customs and norms that prop up the human-animal boundary and conceal as far as possible clues and suggestions of our animality (Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 2008). Although we must eat, poop, and have sex, culture prescribes the proper way to do all these things: whom you can have sex with and how, how and what to eat, where and how to excrete, how to maintain a human body, etc. “People who ignore these prescriptions are reviled as disgusting” (Rozin, Haidt, and McCauley, 2008: 761).

Even some core disgust objects, such as body products (feces, sexual fluids, spit), may be offensive more in their capacity as animal reminders than in their capacity as pathogen sources. People are highly offended by bodily emissions, beyond the point at which this offense serves as a guard against danger. Bloom (2004) points out that people are not offended by letting a sealed vial of hemlock touch their food, nor are they offended by eating detoxified mushrooms, yet people *are* very offended by bodily emissions of all kinds, even if they are clearly not toxic. For example, people are disgusted by their own spit seconds after it has left their mouths (Rozin and Fallon, 1987). Studies have found that people will reject a bowl of otherwise acceptable soup if they themselves have spat in it. Additionally, many people dislike re-eating food that has fallen out of their own mouths.

These sensitivities do not enhance our biological safety, so why do we exhibit them? Because, says Rozin, they help us suppress reminders of our animality. Chewing with one’s mouth open, exposing the mushy contents inside, is considered the height of rudeness at a dinner party. Why? Because it reminds all the guests wearing tuxedos and ball gowns that every bite of

food is sliding down your gullet into your guts, where it will be processed into goop. The valuable goop will be distributed as fuel to your bloody, veiny, squishy animal body, while the non-nutritious remainder will eventually exit in an icky and stinky fashion, potentially accompanied by obnoxious flatulent noises. Thus, if you chew and chomp and smack loudly, and re-eat bits that fall into your beard, napkin, or plate, you will receive aghast facial expressions, and these expressions will basically say “Eww! Stop reminding us that we are animals!”

Interestingly, tears are the only bodily effluvium not widely viewed as disgusting (Angyal, 1941; Rozin and Fallon, 1987). This might be because they are a uniquely human fluid, and thus are not evocative of animality. Other animals do not leak fluid from their eyes, or if they do, it is a reflex, not an expression of emotion (joy, sadness).

4.3. *Interpersonal disgust*

Next, distinct from core and animal-reminder disgust is an additional domain: interpersonal disgust. Interpersonal disgust is an aversion to direct or indirect contact with unfamiliar, unknown, or undesirable others. Returning to our international airport, if you ask people what disgusts them, they might mention used possessions (eating utensils, clothing, beds, cars) and people of inferior social class (Rozin, Haidt, and McCauley, 2008; see also Curtis and Biran, 2001; Rozin, Markwith, and McCauley, 1994).²⁸ These triggers are distinct from core and animal-reminder disgust because the source of anxiety is not potential pathogens or reminders of one’s vulnerability or mortality, but rather (intimate) contact or association with certain types of people.

²⁸ One of the funniest examples of interpersonal disgust, to me, is that many people report feeling disgusted by the sensation of sitting on a warm toilet seat, if said warmth is residue from the body heat of the previous user.

According to Rozin, Markwith, and McCauley (1994), interpersonal disgust comes in four varieties: aversion to contact with (1) strangers, (2) people who have suffered a misfortune, (3) people with diseases (e.g. AIDS, tuberculosis), and (4) people known to have committed a moral transgression. In their study, subjects were asked to rate on a scale of 0 (“something that you would dislike extremely”) to 100 (“something that you would like extremely”) how they would feel about wearing or using various items—a sweater, a car, a hotel bed—that had been previously used or owned by various people but had been thoroughly cleaned (e.g. the sweater had been “thoroughly laundered”). There were seven previous owners examined: (1) a perfectly healthy male stranger, (2) a man who contracted tuberculosis through contact with an infected individual, (3) a homosexual man without AIDS, (4) a homosexual man with AIDS, (5) a man with AIDS who had contracted the disease through a blood transfusion, through no fault of his own, and who is neither homosexual nor a drug user, (6) a man convicted of murder, and (7) a man who lost his leg in a car accident that was not his fault. Subjects’ ratings were then compared to their ratings of a brand-new (not previously owned or used) sweater, car, and bed.

The researchers found that all pre-used items were significantly less attractive compared to the new items. While all pre-used items were aversive, some were more aversive than others. In particular, the perfectly healthy male stranger was comparatively least aversive; the man who lost his leg in an accident and the homosexual man without AIDS were next; and the man with AIDS contracted through a blood transfusion, the homosexual man with AIDS, the man convicted of murder, and the man with tuberculosis were about tied for most aversive. Additionally, individual aversions correlated across scenarios and items. That is, if someone rated the man with AIDS contracted through a blood transfusion very negatively, they were also

likely to rate the man convicted of murder negatively, suggesting that individual interpersonal disgust sensitivities are a stable personality trait.

A different study by Rozin and Nemeroff (1990) had similar results. In this study, subjects were asked to rate, on a scale of +100 to -100 (0 being neutral), how they would feel about wearing a sweater that had been worn for half an hour the previous day by various people. These included a positive set (e.g. “a lover,” “a good public figure”) and a negative set (e.g. “someone with AIDS,” “an evil person,” “someone who had fallen into dog doo”). The study found that positive contact sources had a minimal effect on the ratings; the idea of wearing a sweater previously worn by one’s spouse or favorite politician, for example, was only slightly more appealing than neutral.

In contrast, negative contact sources drove ratings way down. In particular, subjects reported the strongest aversion by far to wearing the sweater if it had been worn by someone deemed “evil.” What was causing this distinct aversion in subjects? In follow-up interviews, subjects seemed to suggest that they conceive of evil as a contaminating source. For example, one subject explained, “It’s the fact that he came into contact with it. Creepy. That he could somehow transmit — uh, somehow the object would pick up some negativeness. I’m not saying it would smell or have dandruff on it, but it would be creepy because he’s a creepy person” (p. 178). Researchers offered subjects various methods of physically cleansing the sweater, such as washing it in warm water and detergent or even submerging it in boiling water. These cleansing methods only “very slightly” reduced subjects’ negative ratings (1990: 178).

4.4. *Moral disgust*

The final domain of disgust is moral disgust. If you ask people at an international airport what sorts of things disgust them, they will list moral transgressions and morally bad people

(Bloom, 2004; Chapman and Anderson, 2013; Curtis and Biran, 2001; Haidt, McCauley, and Rozin, 1997; Imada, Yamada, and Haidt, 1993; Rozin, Haidt, and McCauley, 1999, 2008; Tybur et al., 2013). For example, Curtis and Biran (2001) found that subjects in India frequently named “betrayal” as disgusting; subjects in the Netherlands named “politicians” and “insulting behavior”; and people in the United Kingdom named “rude people,” “man beating a woman,” “cruelty to a horse,” and “wounding of an old lady.”

Extending disgust language to the moral domain is not a quirk of English-speakers (see Imada, Yamada, and Haidt, 1993; Haidt et al., 1997; Rozin, Haidt, and McCauley, 1999; 2008). Many languages—including French (*dégoût*), German (*Ekel*), Russian (*otvraschenie*), Spanish (*asco*), Hebrew (*go-al*), Japanese (*ken’o*), Chinese (*aw-shin*), and Bengali (*ghenna*)—have a word whose semantic domain covers “concerns about the body as well as concerns about other people’s social behavior” (Rozin, Haidt, and McCauley, 2008: 763). Moreover, the extension of disgust language to moral wrongs is not exclusive to adults; beginning at a young age, children also describe some moral wrongs as disgusting (Aznar, Tenenbaum, and Russel, 2023; Danovitch and Bloom, 2009). It would seem that “People of diverse cultures and languages apparently *feel* some similarity in their emotional reactions to feces and sleazy politicians” (Rozin, Haidt, and McCauley, 2008: 763).

Some authors are skeptical that “moral disgust” is a genuine emotional state (see for example Bloom 2004, Nussbaum, 2004). Some of these skeptics prefer to think of extensions of disgust language to the moral domain as mere instances of metaphorical speech. For example, Nussbaum (2004) writes,

Some cases [of moral disgust language] are probably best understood as loose or careless usage, explained, to at least some extent, by the fact that English has no affectively strong adjective with which to express anger. (“That’s outrageous!” seems prissy and bland, so “that’s *disgusting*!” sometimes substitutes) (p. 104).

Similarly, Bloom (2004) writes,

My hunch is that in these statements “disgust” is a metaphor. Saying that we are disgusted by a tax plan is like saying that we are thirsty for knowledge or lusting after a new car (pp. 172-173).

In short, people who say or act like they are disgusted by wrongs or bad people are, contrary to what they may be *saying*, really feeling some other negative emotion (anger, contempt, indignation).

In the skeptics’ defense, there is some empirical evidence that “moral disgust” and other types of disgust (core, animal-reminder, interpersonal) may not be exactly the same emotion. For example, Nabi (2000) found that people said they were “disgusted” by certain moral wrongs, but were reluctant to say that they were “grossed out.” In contrast, they were willing to use both phrases to describe their feelings towards prototypical objects of core disgust, such as feces. This suggests that fecal disgust and moral disgust don’t feel exactly the same.

Perhaps it would be dubious to infer that moral disgust is a genuine emotion, not a mere metaphor, solely based on how people use words (see Fiske, 2020 for a discussion of the “lexical fallacy” in emotion research). However, there is other evidence supporting the hypothesis that moral disgust is genuine disgust (for an excellent review, see Chapman and Anderson, 2013). Specifically, people claiming to be in the grips of moral disgust exhibit prototypical disgust outputs (gape face, parasympathetic arousal, distancing behavior, neural activity, contamination psychology).

One source of evidence is neural imaging. Neural imaging studies show that people who play the “Ultimatum Game”—an experimental economics game—and receive an unfair offer experience activity in the bilateral anterior insula, a region of the brain associated with the emotional experience of disgust (Sanfey et al., 2003). The Ultimatum Game involves two

players: a proposer and a responder. The proposer decides how to split a pot of money. For example, if the pot is \$10, the proposer could offer a fair \$5/\$5 split or an unfair \$9/\$1 split in their own favor. The responder then decides whether to accept or reject the offer. If they accept, the pot is split as proposed; if they reject, no one gets any money. Sanfey et al. (2003) found that responders reject most \$9/\$1 and \$8/\$2 offers when the proposer is a computer, and when the proposer is another human, the likelihood of rejection increases. Sanfey et al. (2003) also found, using fMRI, that responders given an unfair offer showed increased neural activity in the bilateral anterior insula, with activity increasing in correlation with the unfairness of the offer. In response to this finding, the researchers write,

Though studies of disgust have largely focused on physical sensations of taste and odor, it has been suggested that emotion-based disgust (as perhaps induced by an insultingly unfair offer) may be conceptually similar. The recruitment of similar neural structures, namely the anterior insula, in both physical and moral disgust gives some credence to this notion (Sanfey, 2003: 1756).

In other words, the researchers believe that their study lends support to the hypothesis that moral disgust is a genuine form of disgust, not a metaphor for some other emotion.

A second source of evidence comes from facial expressions. In a different study on the Ultimatum Game, Chapman et al. (2009) examined the facial expressions of responders receiving unfair offers and compared them to the faces of people experiencing distaste (in response to a bitter food) and core disgust (in response to feces). They found that “all three states evoked activation of the levator labii muscle region of the face, characteristic of an oral-nasal rejection response.” The researchers concluded: “These results suggest that immorality elicits the same disgust as disease vectors and bad tastes.”

Additionally, there is evidence that children associate the disgust face with disapproval of immoral behavior, even before they know the term “disgust.” Danovitch and Bloom (2009)

found that when presented with a vignette of immoral behavior (e.g. someone being mean or rude to a friend), fourth-graders frequently use disgust language to express disapproval, while younger children (second grade and kindergarten) do so much less frequently. To test whether this discrepancy was due to a language barrier, the researchers adopted a different approach: instead of focusing on words, they examined whether children believe the disgust face “goes with” moral wrongs. It turns out they do. When children were presented with vignettes depicting immoral behavior and then shown images of various facial expressions (anger, sadness, surprise, disgust), they frequently chose the disgust face to match the story. Most importantly, when face-matching tests were used instead of language-based tests, the ability gap between kindergarteners and fourth graders nearly disappeared. This suggests that children as young as kindergarten do feel or associate disgust with some moral wrongs even before they know the word “disgust.”

A final source of evidence: behavior. It turns out that people claiming to be in the grips of moral disgust exhibit distancing behavior (don’t look, don’t touch) and contamination psychology. For example, in Rozin, Markwith, and McCauley (1994), subjects reported that wearing a thoroughly laundered sweater formerly owned by a convicted murderer would be highly unpleasant—comparable to wearing a sweater formerly owned by a diseased person. Similarly, Rozin and Nemeroff (1990) found that subjects were highly averse to wearing Hitler’s sweater. In both studies, the sweaters were imagined to be “thoroughly laundered” or even cleansed in boiling water, but this made little difference to the subjects, who continued to view the sweaters as tainted by their contact history. In a similar vein, Haidt and Hersh (2001) asked subjects to consider sexual taboo violations, such as a man masturbating while being licked by his dog and a pair of adoptive siblings having consensual sex, and then asked if they would feel

comfortable being a dinner guest at the violator's home. Result: many subjects were strongly averse to dining with the violators.

Some instances of so-called “moral disgust” seem entangled with core or animal-reminder disgust. For example, many people are disgusted by gory murders. Here, is the disgust directed at the wrongness of the murder, or simply the blood, dismemberment, exposure of viscera, etc.? If the latter, then calling this instance of disgust “moral” sounds inaccurate; more accurately, the gore evokes core disgust, which is a non-moral sentiment. If all alleged examples of “moral disgust” simply involve elements of core disgust, this would threaten to undermine the claim that moral disgust is a genuine emotion (for helpful discussion here, see Fileva, 2021).

However, some objects of moral disgust *do not* involve any potential physical disgust elicitors (Chapman and Anderson, 2013). For example, as noted above, disgust is often triggered by unfair offers in the Ultimatum Game, which does not involve any sticky or gooey substances, bodily emissions, or animals. Additional examples include hypocrisy, cheating, and betrayal. So, it doesn't seem like the disgust in this case can be attributed to a core trigger.

While disgust responds to some types of immorality, it doesn't respond to all types. For example, people do not typically report being disgusted by murder, or robbing a bank. What types of immorality, then, do trigger disgust? Different views on this question have been proposed, which I will discuss in more detail in the next chapter (Chapter 3, Section 3). For now, let me mention two prominent, and contrasting, proposals.

One prominent proposal is that disgust is triggered by violations of norms related to “purity” (or sanctity, or divinity). These are transgressions that are distinctive in being—like cockroaches—intrinsically offensive: they are dirty, twisted, and debased; they pollute the self; they evoke fears of social decay; they suggest that the violator lacks self-control or is ruled by

carnal passions; or they violate the ‘natural’ teleology of sex or the family. Whether purity violations are harmful, unfair, or unjust is of secondary relevance to our disapproval of them; often they are harmless (e.g. two adopted siblings having one-time, protected sex is widely morally shunned).

Another proposal—I’m agnostic as to whether it is competing or supplementary—is that moral disgust is triggered by “anti-social violations.” These are transgressions (or vices) that erode reciprocity norms (shared expectations of cooperation and trust) and are prone to spread via social contagion if not contained. Why did disgust, the poison/parasite police, get recruited by cultures to police reciprocity norms, despite the availability of other emotions (e.g. anger)? Joshua Tybur (2009) suggests the following story²⁹: Ancestral human cultures were vulnerable to “social parasites” (hypocrites, cheaters, liars, etc.), who imposed significant costs on the group by undermining norms. But these costs were indirect and dispersed across a pool of victims. Thus, it was in no specific person’s self-interest to confront parasites; no victim would recoup much value from e.g. fighting off a parasite (not enough to outweigh the risk of personal injury, anyway). Disgust presented itself as an adaptive solution to this problem; its characteristic motivation and behavioral pattern (distancing, avoidance) provided a low-cost method for policing parasites: social exclusion (if you don’t play fair, we won’t fight you, but we will preferentially not interact with you).

5. Disgust Acquisition

5.1. The trajectory of disgust development in children

When does disgust develop in children? The timeline is a little fuzzy, but one thing is clear: for the first roughly two years of life, young children show no signs of disgust (Bloom,

²⁹ I discuss Tybur’s theory on this in more detail in Chapter 3, Section 3.

2004; Rozin, Haidt, and McCauley, 1999; Rozin, Haidt, and McCauley, 2008; Tybur et al., 2013). It is possible to trigger the gape face in babies via a distaste reaction, e.g. by giving them bitter or unpleasant foods. However, babies do not gape at objects or smells considered offensive by adults (Darwin, 1876; Rozin and Fallon, 1987).

The most illustrative example here is feces. While feces are universally reviled by adults, children under age 2 show no sign of offense at feces. Rozin et al. (1986) found that 62% of children under 2 were willing to taste imitation dog poop (chocolate shaped like dog poop and laced with cheese that imitates a fecal scent). If anything, very young children are curious about, even fascinated by, feces, as evinced by their tendency to play with their own feces at some point (Rozin and Fallon, 1987). Moreover, children love “potty humor” (jokes about poop, pee, barf, and farts) (Bloom, 2004).

Feces aside, there are other adult disgust elicitors to which children under 2 show high willingness to put in their mouths. Rozin et al. (1986) found that many children were willing to taste or eat a grasshopper, a whole, dried, small fish, and a sterilized lock of human hair—items rejected by adults.

There are widely documented disparities in how children and adults react to bad smells. In some cases, many children do not dislike smells that virtually all adults dislike. For example, Schmidt and Beauchamp (1988) found that 29% of 3-year-olds (almost a third) rated the smell of spoiled milk as “pleasant,” compared to 0% of adult subjects. In other cases, children and adults may regard the same smells as bad, but children are more tolerant of them (Bensafi et al., 2007). The disparity between children and adults regarding bad smells diminishes as children get older. For example, Brothánková, Vodička, and Bažant (2021) found that 45% of younger children

(aged 6-7) rated the smells of butter, fish, and onion as unpleasant, compared to 63% of older children (aged 11-12)—a sharp increase with age.

While young children may lack a fully mature sense of disgust, they do reject lots of foods starting around age 3 or 4. In Bloom's (2004) verbiage, many children around this age become "neophobes" (highly picky eaters), suspicious of any unfamiliar foods. However, their reasons for rejection, until they get older, tend to be based on distaste or a sense of danger, not disgust (Rozin and Fallon, 1987).

When do children exhibit contamination sensitivity? Sensitivity to contamination is a hallmark of adults' disgust; for adults, an otherwise neutral object can become tainted by contact or association with a primary disgust object. Some researchers believe that contamination sensitivity appears as early as age 3 or 4 (see Siegal and Share, 1990). Other researchers place it in the 6-7 year-old range. For example, Rozin, Fallon, and Augustoni-Ziskind (1985) examined the contamination sensitivity of 4- to 12-year-olds by presenting them with a glass of their favorite drink, dropping a contaminant into it (e.g. a fly, a used comb), and then progressively removing the contaminant in stages (removing the contaminant with a spoon, spilling and then refilling the glass, and washing the glass entirely and refilling it). The researchers note that all children rejected the drink as long as the contaminant was present. At what point were children willing to drink again? This depended on age. Children under 7 were generally willing to drink again after the contaminant was removed with a spoon. In contrast, children older than 7 tended to wait until a later stage, such as spilling the glass and refilling it or washing the glass entirely.

It makes sense that contamination sensitivity would be delayed in childhood. After all, it requires fairly advanced cognitive prerequisites (Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 2008). For example, to view a secondary object as contaminated, children need to

recognize the distinction between appearance and reality; even though two tooth brushes, one used and one new, may look exactly the same, one is different due to its contact history (Rozin, Haidt, and McCauley, 2008). Additionally, worrying about *literal* contamination by toxins or pathogens (the primary concern of core disgust) may require an “atomistic” worldview—an understanding that there are things in the world that can affect people that are so small as to be invisible (Rozin and Fallon, 1987). Contamination sensitivity can only develop after these abstract thoughts emerge.

Overall, the foregoing paints a fairly clear picture of the developmental trajectory of disgust: (1) First, disgust is entirely absent during the first few years of life. (2) Then, it goes through a developmental period, perhaps between the ages of roughly 3 to 8, during which time it is immature and only partly resembles adult disgust. (3) By early adolescence, disgust has fully matured.

Of course, as has been noted, there are different types of disgust—core, animal-reminder, interpersonal, and moral—and they do not develop in children at the same time. According to Paul Rozin, they appear incrementally and in a predictable order: core disgust comes first, followed by animal-reminder, then interpersonal, and moral disgust comes last (Rozin, Haidt, and McCauley, 2008). Other researchers agree that different types of disgust emerge at different times, and core (or generic, or physical) disgust comes first (see Oaten, Stevenson, and Case, 2009; Stevenson et al., 2010; Tybur et al., 2013).

It is worth noting that across domains, people tend to be neophobes (Bloom, 2004; Haidt et al., 1997; Rozin, Haidt, and McCauley, 1999). As a general rule, children learn from parents and culture not what is prohibited but, rather, what is permitted. Everything else is gross by default. Core disgust rejects *most* animals and animal products for consumption; sexual disgust

rejects *most* potential sexual partners (family members, children, old people, animals, an entire gender for hetero- and homosexuals, etc.); and interpersonal disgust rejects *most* people as partners for social intimacy (sharing food, clothes, linens, casual touch). Rozin, Haidt, and McCauley (1999) write, “In all three cases, disgust makes us extremely selective, critical, and judgmental” (p. 435).

5.2. *Acquisition of specific triggers*

How do young children acquire a sense of disgust, and which specific objects come to be seen as disgusting? Certainly, socialization—transmission of attitudes from one person to another—plays a role. The first socialization experience is potty training (Rozin, Haidt, and McCauley, 1999; Rozin, Haidt, and McCauley 2008). During this period, parents teach, consistent with local norms, where and how to excrete and also attempt to cultivate in the child an attitude of disgust towards feces. Parents communicate disgust via verbal and nonverbal expressions (Oaten et al., 2014).

The role of parents in the socialization of disgust in children appears large. Research shows that parents and their children have strikingly similar disgust sensitivities, as do sibling pairs (Haidt, McCauley, and Rozin, 1993). Of course, similar disgust sensitivities within families may be due to shared genes rather than socialization. Tybur et al. (2018) argue, based on data from identical twins separated at birth, that genes are the more likely culprit.

An important nonverbal cue in the parent-child transmission of disgust is facial expression (Oaten et al., 2014). Facial expressions in general are known to trigger physiological arousal in onlookers. The disgust face is no different; it can trigger parasympathetic arousal (nausea, lowered heart rate) in onlookers. The causal mechanism here may involve mirror neurons (Oaten, Stevenson, and Case, 2009; Rozin, Haidt, and McCauley, 2008). Because seeing

someone make a disgust face at an object can arouse vicarious disgust in oneself, one can literally internalize rejection of that object without any further experience with it. Thus, children may learn what to consider disgusting by observing the faces of adults and having vicarious disgust experiences. Repacholi (1998) found that children as young as 14 months reacted negatively to an object at which an adult made a disgust face.

If adults universally revile feces, but young children do not (indeed, they often show interest in feces), this raises the question: Is fecal disgust innate or enculturated? Freud thought fecal disgust was fully due to enculturation. In his view, children are born with a positive attitude towards feces. Then, during potty training, this attraction is repressed. More recently, many authors suggest that fecal disgust, and core disgust more generally, is somewhat plastic—cultures can attune individuals' disgust sensitivities to local conditions and customs—but it is also constrained by biology. Humans exhibit a biological “preparedness” to find certain things disgusting (Bloom, 2004; Curtis, 2001; Oaten, Stevenson, and Case, 2009; Rozin and Fallon, 1987). Teaching disgust toward things outside this prepared list is hard and may be impossible. For example, it may be impossible to deliberately teach a child to be disgusted by sweet-tasting candy (Curtis and Biran, 2001).

Why does the sense of disgust appear relatively late in childhood? This might seem puzzling, given that core disgust serves an important biological function (protecting the body from infections and unsafe foods). Tybur et al. (2013) suggest three reasons, from an evolutionary standpoint, why this might be. According to the first two suggestions, disgust did not develop early because there was no immediate need for it. First, in an ancestral context, children breastfed for roughly the first two years of life, thus obtaining their nutrition exclusively from their moms. This mitigated the need for an unsafe-food-rejection system. Further mitigating

the need for such a system was the fact that children were mostly carried around by their moms. Thus, as long as mom is sufficiently disgust-sensitive, children are safe from toxins and pathogens.

This goes well with the widely documented fact that women, on average, are significantly more disgust-sensitive than men (see Haidt, McCauley, and Rozin, 1993). This heightened sensitivity may have helped women—the primary caretakers for children in ancestral contexts—avoid infections and unsafe foods for both themselves and their offspring. In other words, a higher infection-avoidance workload required higher motivation (in the form of increased disgust sensitivity).

A third suggestion from Tybur et al. (2013) is based more on the omnivore’s dilemma mentioned in §3.2. Given humans live in very different environments, it is necessary for children to be born with an initial openness to new foods and substances. This provides time to learn what local food, hygiene, and mating customs are practiced. Eventually, this learning window closes and disgust sensitivities solidify, at which point children exhibit the same pickiness and general neophobia as adults.

6. The Role of Culture in Disgust Acquisition

Several authors suggest that culture influences the acquisition of disgust (see for example Curtis, de Barra, and Aunger, 2011; Rozin, Haidt, and McCauley, 1999, 2008; Tybur et al., 2013). There are a few reasons to think this. First, as the previous section notes, the formation of disgust in children depends partly on socialization (transmission of attitudes and ideas from one person to another). For example, during potty training, parents convey local attitudes and norms surrounding excretion to their children. This opens the door to culture having an influence, insofar as parents influence children’s attitudes, and culture influences parents’ attitudes.

Second, some cross-cultural variation in the domain of *core* disgust has been observed. Core disgust serves the purpose of guarding against toxins and pathogens in the environment, but human environments differ (Curtis, de Barra, and Aunger, 2011). As such, different environmental pressures can lead to slightly different sets of core disgust elicitors and food taboos. For example, in Fiji, it is taboo for women to eat shark during pregnancy (Henrich, 2015). It turns out that shark meat contains a toxin that is dangerous to fetuses. So, this taboo serves a biological function. In contrast, in the U.S., this taboo doesn't exist. Why? Answer: because shark isn't part of the normal diet anyway.

Third, several authors agree that, to the extent that disgust has been elaborated beyond a core biological domain to include triggers related to animals, strangers, and moral transgressions, culture must be responsible for this, causally speaking. Evidently, in the distant past, cultures that took disgust, an existing system, and extended it to serve certain social ends reaped adaptive benefits. This is a further example of preadaptation: an existing system, which evolved for its own purpose, was extended to serve an additional function, out of convenience.

Some authors stress that most cross-cultural variation in disgust inputs is concentrated in the moral domain (see for example Haidt et al., 1997; Imada, Yamada, and Haidt, 1993; Rozin, Haidt, and McCauley, 2008). These differences cannot be explained by biology or individual disgust sensitivities, leaving culture as the best explanation. Consider two examples:

First, consider the moral disgust experiences of Americans and Japanese people. Imada, Yamada, and Haidt (1993) asked American and Japanese college students to describe three disgusting experiences that they have encountered. Unsurprisingly, both groups listed certain socio-moral experiences as disgust-triggering, with some overlap in specific inputs (e.g. cheating, betrayal by a friend, interacting with a disliked person). However, there were

significant divergences. The authors write that “the content of the socio-moral examples appeared very different in the two samples” (p. 155) and “the discrepancy seems much greater than those in the core disgust and animal-reminder disgust domains” (p. 173).

In a later analysis, Haidt et al. (1997) describe the divergence thusly:

For Americans, socio-moral disgust is a kind of character judgment of others, especially of people who violate the basic dignity of other human beings. The largest single class of events, accounting for 21 per cent of the total pool of responses, described acts of senseless violence or cruelty, especially towards people who are weak or defenceless. Examples included Serbian atrocities, a recent Chicago mass murder, children who tease homeless people, and “a father who shook his child to death because he didn’t like something that he saw during a football game.” The second largest class, accounting for 19 per cent of all responses, included ugly or offensive beliefs and attitudes, especially racist attitudes. The American respondents occasionally mentioned situations in which they themselves were mistreated, but for the most part they mentioned evil people whom they had met or heard about (p. 118).

For the Japanese, in contrast, *ken’o* was experienced during their everyday social interactions, in situations where things were not going right. They mentioned everyday frustrations in which other people failed to meet their needs, or even worse, where other people abused or shamed them (e.g. “When I was criticized for my driving with very harsh words”; “When punks tried to pick a quarrel with me”). They also mentioned situations where they themselves failed to live up to certain standards (e.g. “When I found myself not to be the person that I should be”; “When I did not find my name on the board where names are posted of people who passed the entrance exam”) (Haidt et al., 1997: 118-119).

In short, American moral disgust tends to be triggered by brutal or cruel acts that strip people of their basic humanity, as well as by character flaws like sleaziness, hypocrisy, and insincerity. Meanwhile, Japanese moral disgust tends to focus on failures to achieve good fit in social relationships. Rozin, Haidt, and McCauley (2008) concur with this interpretation of the results.

A second example: American liberals and conservatives. Self-identified liberals and conservatives exhibit differences in their judgments of moral disgust, specifically regarding homosexuality. On average, conservatives believe that homosexuality is morally wrong at much higher rates than liberals (Haidt and Graham, 2007; Haidt and Hersh, 2001). This is probably linked to the disgust response (Haidt and Hersh, 2001). There is some evidence that conservatives are more disgust-sensitive in general, compared to liberals (Inbar and Bloom, 2009). Moreover, conservatives report being disgusted by homosexuality at much higher rates than liberals (Haidt and Hersh, 2001).

It is worth mentioning that conservatives and liberals do not exhibit significant differences in religious strength (Haidt and Hersh, 2001). Moreover, Haidt, McCauley, and Rozin (1997) found that people high in religious strength are generally more disgust-sensitive, but the authors note that this effect was “marginal” (p. 709), much smaller than the effect of gender. In short, saying something like “conservatives are more moral-disgust-sensitive than liberals because conservatives are more religious” would not seem correct.

So, here we have two examples where culture, not biology or individual experience, seems to influence people’s disgust response. The two examples are quite different: one compares attitudes in two different countries—the United States and Japan—while the other compares two different political groups in the U.S.—liberals and conservatives. Despite being different, we can draw insights about both examples from a single theory in moral psychology: Moral Foundations Theory.

Moral Foundations Theory (MFT), advocated primarily by Jonathan Haidt, holds that all people are born with certain moral “taste buds,” or concerns that they are prepared to care about (Haidt and Graham, 2007; Haidt and Joseph, 2004). Which taste buds will be activated, and how

sensitive they are, is determined by one's surrounding culture. Haidt originally proposed five taste buds, or moral foundations: (1) harm, (2) fairness, (3) ingroup loyalty, (4) authority, and (5) purity. All humans are born prepared to be sensitive to these foundations, but as you grow up, local specifications will influence the development of your own moral psychology. Your culture might amplify one foundation while downplaying others. As such, the intensity of your emotional reactions to violations in each of the five domains will track with the emphasis or priority that your culture places on them.

MFT also holds that moral judgments are highly influenced by emotional reactions (Haidt, 2001). Thus, if you have strong emotional reactions to harm violations but not purity violations, you are more likely to condemn harm violations in your moral reasonings, and less likely to condemn purity violations. The intensity of your emotional reactions to moral violations in each of the five domains will largely track with the emphasis that your culture places on them.

Haidt and Graham (2007) note that MFT is derivative of an earlier theory, namely anthropologist Richard Schweder's theory of the Three Ethics of Morality (TEM). Shweder et al. (1997) argue that most cultures' ethical codes can be subdivided into three distinct codes or moral languages: (1) the ethics of autonomy, which focuses on rights, fairness, and justice, (2) the ethics of community, which focuses on duty, hierarchy, and execution of social roles, and (3) the ethics of divinity, which views the body as a "temple" and focuses on protecting the spark of divinity in human beings from degrading or polluting acts.

TEM, it is argued, can be used to analyze cross-cultural differences in morality. Most cultures invoke aspects of each code—autonomy, community, and divinity—but the specific strength or rank order of the codes may differ. In the words of Rozin, Haidt, and McCauley (1999), "All three ethics may be available to individuals in all cultures, but cultures appear to

show preferences and differential patterns of use of the three ethics” (p. 438). For example, in Hindu India, the ethics of divinity is highly salient; in Japan, the ethics of community is highly salient; and in America, the ethics of autonomy is highly salient.

Moral disgust, in general, seems like an emotion that has been co-opted by cultures to police moral violations in the domain of the ethics of divinity. Violations in this domain are known as “purity violations.” Haidt and Graham (2007) describe the character of purity violations as follows:

Those who seem ruled by carnal passions (lust, gluttony, greed, and anger) are seen as debased, impure, and less than human, while those who live so that the soul is in charge of the body (chaste, spiritually minded, pious) are seen as elevated and sanctified. From this point of view, a philosophy that says “if it feels good, do it” is the philosophy of the devil (Haidt and Graham, 2007: 106).

Similarly, Rozin et al. (1999) suggest that, to understand purity violations, “think about things like sin, the natural order of things, sanctity, and the protection of the soul or the world from degradation and spiritual defilement (p. 576).” Empirical evidence suggests that the specific emotion that (perceived) purity violations evoke in people is not anger or contempt but *disgust* (Landmann and Hess, 2017; Rozin et al., 1999).

For example, Landmann and Hess (2017) gave subjects vignettes depicting moral violations related to each moral foundation in MFT. For example, for harm/care, one vignette was “You see a boy placing a thumbtack sticking up on the chair of another student”; for authority, one was “You see a girl repeatedly interrupting her teacher as he explains a new concept”; for purity, one was “You see a man searching through the trash to find women’s discarded underwear”; and so on. Subjects then rated their emotional response to each vignette. The researchers found that the care, authority, fairness, and loyalty vignettes evoked a range of emotions (e.g. anger, resentment, fear, contempt). In contrast, disgust was “highly specific” and

“clearly associated with” purity violations (p. 8). They further note “disgust elicited by purity violations was significantly stronger than disgust elicited by other violations” (p. 4). The researchers conclude that their study offers qualified support for MFT, with disgust/purity being a clear candidate for a moral foundation, alongside intentional norm violations/anger and suffering/compassion.

TEM and MFT help make sense of broad differences in American and Japanese moral disgust, as well as differences in liberal and conservative moral disgust. Regarding the former, American and Japanese moral disgust seem to be policing boundaries that local morality deems inviolable, but “local morality” is different. In the words of Rozin, Haidt, and McCauley (2008),

American moral disgust may be guarding against threats to individualist, rights-based social order, whereas Japanese *ken'o* may be guarding against threats to a more collectivist, interdependent social order.

In short, cross-cultural divergences in moral disgust judgments seem driven by broader cross-cultural divergences in moral code or moral worldview (Rozin, Haidt, and McCauley, 1999, 2008).

Likewise, in the conservative-liberal case, it seems like the divergences in each group’s moral disgust can be explained by thinking of the groups as cultures with distinct moralities that emphasize (and downplay) different moral foundations. Liberal morality focuses on the harm/care and justice/fairness foundations (Haidt and Graham, 2007). Meanwhile, conservative morality has a broader base; it recognizes not only the harm/care and justice/fairness foundations, but also purity/sanctity/divinity, which is why self-identified conservatives are more prone to disgust (Haidt and Graham, 2007).

Haidt and Hersh (2001) note that conservatives, in their condemnations of homosexuality, incest, and bestiality, expressed sentiments evocative of the

purity/sanctity/divinity domain. Some subjects expressed “a teleological view of social roles,” such as “Sex and love are ‘supposed’ to be between a man and a woman” (response to gay anal sex vignette) and “family relationships are ‘supposed’ to be nonsexual” (response to incest vignette) (p. 206). Conservatives also invoked “a teleological view of the body, in which each part of the body must play the role for which it was ‘designed’,” and expressed concern about “misuse of the ‘temple of the body’” (p. 208).

Interestingly, in Haidt and Hersh’s (2001) study, they found that liberals were just as likely as conservatives to be disgusted by, and to morally condemn, incest between close biological relatives. Moreover, both groups overwhelmingly favored criminalizing incest. Haidt and Hersh write, “The liberal insistence that people have a right to do whatever they choose, so long as they don’t hurt anyone, did not extend to the powerful taboo against incest” (p. 213). However, liberals were noticeably less likely to support criminalization in the cases of incest between half-siblings and adopted siblings. Why?

A partial answer may be that liberals and conservatives both have disgusted reactions to incest, but “liberals often separated their own emotional reactions from their moral judgments” (p. 211). They may feel “a flash of disgust or discomfort” when thinking about certain forms of incest, but because liberal morality holds that “sex is entirely a matter of personal choice,” they feel compelled to override their intuitive disgust (p. 212). Meanwhile, many conservatives simply embrace their disgust as justification: “In the course of searching for justifications, conservatives often turned to their own affective reactions of shock, disgust, and discomfort” (p. 211). Here is how one conservative subject condemned the bestiality vignette: “I just think it’s disgusting and, um, I don’t think anyone should be doing that because it’s disgusting and it’s strange” (p. 211).

Other authors offer corroborating evidence for this. Feinberg et al. (2014) found that liberals were more likely than conservatives to feel intuitive disgust at a practice, like gay marriage, initially but to engage in emotional reappraisal before rendering a moral judgment. One subject in Haidt and Hersh (2001) seems to be doing this very thing in response to a vignette about homosexual anal sex:

I'd have to say it would make me feel weird. That's my first instinct. Yeah, I mean there's a weird kind of thing going on in my head there. I feel like there's a way I'm going to act and a way to act based on this preconceived ideology I have about sex being okay between people, no matter what sex they are. You know, instantly a warning bell comes up, and that bugs me. I quickly discard it; I mean I recognize it and then discard it because it makes me feel silly.

In short, liberals, like conservatives, may feel intuitive disgust at various sexual practices, but unlike conservatives, they resist letting disgust influence their ultimate moral judgments. Why do liberals resist the temptation to let disgust influence moral judgment? Miller (1997) suggests liberals are uncomfortable with disgust because it is often at odds with an egalitarian ethos:

It is easy to see the costs of letting disgust and contempt order our moral and social lives without their being subordinated to or constrained by countervailing sentiments and principles. We end up punishing the stigmatized, who may have no justifiable cause for feeling guilt for their stigma, although they often internalize the social judgments of their stigmatization as shame, self-loathing, self-disgust, self-contempt, self-hatred. We fear that contempt and disgust may violate norms of fairness and justice, of a liberal respect for persons; that they may maintain brutal and indefensible regimes (p. 202).

Rozin, Haidt, and McCauley (1999, 2008) agree with Miller's assessment that liberals work to reappraise their intuitive disgust when these intuitions conflict with the harm/care or justice/fairness foundations.

7. Conclusion

The descriptive account of disgust laid out in this chapter can be summarized as follows. Disgust is revulsion at the prospect of intimate contact with an offensive or contaminating object, such as bodily emissions, bugs and animals, death and corpses, or unsavory properties (slimy, sticky). It is negative (unpleasant to feel) and visceral (felt in the body). When triggered, it manifests in a characteristic set of outputs: (1) the gape face (wrinkled nose, gaping mouth, raised upper lip), (2) neurophysiological arousal (nausea, increased activity in the anterior insula), and (3) distancing behavior (don't look, don't touch, stay away). Importantly, disgust triggers are avoided not for being distasteful or dangerous but for being (interpreted as) *offensive* (dirty, soiled, polluted). These triggers are also (interpreted as) contaminating, capable of transferring offensiveness to otherwise neutral objects via contact or association. Phenomenologically, it often feels (perhaps irrationally) as if intimate contact with a disgust trigger will debase or degrade one's body, self, life, or moral purity.

Disgust is a basic and universal emotion. People from entirely different cultures are disgusted by many of the same things and express their disgust in strikingly similar ways. This universality is explained by the fact that disgust evolved to serve a crucial biological function: aiding humans in avoiding infections and unsafe foods. Its biological crucialness is especially visible in how nausea-inducing experiences with novel foods can automatically create stubborn aversions, and also in the general memory advantage associated with disgust (disgusting experiences, images, words, and stories are especially memorable).

But disgust is not solely a biological signal. Cultures have extended it to additional, non-biological domains. People feel disgust toward (i) animal-reminders—things evocative of our animality, vulnerability, and mortality; (ii) interpersonal objects—unknown, unfamiliar, or

undesirable others; and (ii) certain moral transgressions—racism, unfairness, sexual immorality, cruelty, hypocrisy, and other acts deemed dirty, vile, or socially polluting.

There is some cross-cultural variation in disgust triggers, mostly in the moral domain. These variations are driven largely by divergences in group morality. General theories of human moral psychology—such as Shweder’s Three Ethics of Morality and Haidt’s Moral Foundations Theory—can help explain these intergroup divergences. Specifically, humans seem born prepared to care about—and to be emotionally invested in policing—certain moral foundations (e.g. community, purity, autonomy). But your culture determines the relative sensitivity you have to each foundation. The specific mechanisms by which these sensitivities are transmitted to you involve your parents and upbringing. In early childhood, you soak up the disgust sensitivities of your parents by paying attention to their reactions and facial expressions to various objects and behaviors.

The moral foundation most closely tied to disgust is purity—the domain that views the body as a “temple” and is concerned with protecting it from taint. Some cultures amplify the purity domain more than others. This may partly explain why conservatives and liberals tend to have divergent views on certain issues, such as sexual morality. Liberals and conservatives belong to distinct subcultures, which influence their members’ moral psychologies: liberals focus on harm/care and justice/fairness, while conservatives care about these foundations as well as the purity/divinity domain.

7.1. The upshot

So, where does this descriptive account of disgust leave us, with regard to my dissertation’s larger question—whether moral disgust should influence criminal law?

One takeaway is that we now have a better idea of what is going on inside people's minds when they feel disgusted by a moral transgression, whether it be unfairness in economic dealings, racism, or deviant sexual behavior. Disgust responds to offensiveness and contamination. Insofar as moral disgust is a species of this more general disgust system, it tracks moral offensiveness and moral contamination. Disgusting moral transgressions, such as racism and incest, are seen as wrong *in a distinctive way*. They are not merely harmful (e.g. murder is harmful, but does not make us gag); racism and incest *reek*. They have an ugliness about them that induces the heebie-jeebies, even in cases where they are harmless (e.g. Leonardo DiCaprio's character in *Django Unchained* making incredibly offensive comments about black people, or siblings Mark and Julie having consensual, safe sex). Disgusting moral transgressions also feel contaminating, as if too much exposure to them will somehow risk letting them get *in* you or get *on* you. We instinctively respond to this contamination threat by distancing—whether literally, attentionally, or symbolically.

Maybe it is irrational to be disgusted by moral offenses—to mentally construe them as icky and contaminating—but it is hardly *surprising* that this happens. We are embodied beings whose evolutionary history has saddled our brains with programs and modules that were necessary to keep our hominid ancestors alive in a world teeming with invisible, intangible, difficult-to-detect threats. When we turn our minds to moral questions, we're using the same brain. It seems natural that we would intuitively feel anxious about invisible, intangible, difficult-to-detect *moral* threats, to ourselves and our communities.

The question is whether these intuitions are philosophically defensible. Should we regard any moral transgressions as inherently offensive and contaminating? Is there a real sense in which ourselves or our communities can be morally “polluted”? Is distancing ever a morally

appropriate or necessary policy for handling any wrongs? And if so, what, if anything, would this mean for political discourse, public institutions—or criminal law?

Disgust advocates and skeptics will disagree about this, and their arguments need to be evaluated and compared against each other. That is the task ahead.

However, before we go there, there is one more preliminary issue that must be addressed: whether there really is a genuinely *moral* species of disgust. It seems intuitive to me (and others) that there is one, if only based on introspection (aren't you simultaneously morally offended *and* grossed out by incest, hypocrisy, or racism?). However, in the psychological and philosophical literature, the ontological status of "moral disgust" is increasingly contested. Some authors believe talk of "moral disgust" is all hype; it is a mere metaphor, or a misnomer, or reducible to other states. If these authors are right, then the entire disgust advocate/skeptic debate—about whether moral disgust holds any normative significance—is moot. If there is no such psychological state as "moral disgust," then it can't carry normative significance. Things that don't exist don't have properties.

The next chapter takes up this issue. Is "moral disgust" a genuine psychological state—or not?

Chapter Three

Can Iniquity be “Icky”? Vindicating the Existence of Genuine Moral Disgust

There’s no place in America for this kind of violence. It’s sick. It’s sick.

—Joe Biden on Trump rally shooting

Almost everyone would acknowledge that these amusements of this polite, pleasant-looking young man from Milwaukee were “sick! sick! sick!” (At least three reiterations are needed to adequately convey the speaker’s disgust.).

—Joel Feinberg on Jeffrey Dahmer, *Problems at the Roots of Law*

Murder most foul, as in the best it is; But this most foul, strange and unnatural.

—Ghost, *Hamlet*

Woe to you, teachers of the law and Pharisees, you hypocrites! You are like whitewashed tombs, which look beautiful on the outside but on the inside are full of the bones of the dead and everything unclean.

—Jesus, Matthew 23:27

1. Introduction

At first blush, moral disgust seems like a strong candidate for a moral emotion. In ordinary speech, the extension of disgust language to morally problematic actions, practices, and people is ubiquitous. This is illustrated in the epigraphs—political violence is “sick” (serial killers triply so), fratricide is “foul,” and hypocrites at their core are “unclean.” Additional examples are easy to think of: traitors are “rats,” liars “reek” of dishonesty, racists propagate “bile,” and so on. This phenomenon—calling immorality “disgusting”—exists in many languages besides English and may be universal (Rozin et al., 2008: 763). Even children, starting

relatively young, can be observed extending disgust language to bad behavior (Aznar et al., 2023; Stevenson et al., 2010; Danovitch and Bloom, 2009).

It is no wonder, then, that so many contributors to the disgust advocate/skeptic debate—about whether disgust, in itself and as such, carries any normative significance—assume that “moral disgust” is a genuine emotional state. This assumption is especially prevalent on the advocates’ side. In the 1990s, several philosophers began leveraging the disgust expressions in ordinary moral discourse as ammunition in ethical and policy debates. Most infamously, in 1997, bioethicist Leon Kass wrote a paper—“The Wisdom of Repugnance”—advocating for universal bans on human cloning and other emerging biotechnologies which, he worried, would radically alter essential features of human life. The crux of his argument: these technologies are “repugnant,” and repugnance is “the emotional expression of deep wisdom, beyond reason’s power to fully articulate it” (p. 687). In other words, Kass attributes a sort of *epistemic* significance to disgust; the emotion can signal the (serious) wrongness of actions or practices whose wrongness we might otherwise overlook.

Post-Kass, philosophical debate about the potential normative significance of disgust blossomed, expanding well beyond bioethics into political philosophy, philosophy of criminal law, and discussions of the moral emotions more broadly. More authors began attributing epistemic significance to disgust (see for example Midgley, 2000; Kahan, 1998). Beyond epistemic significance, some philosophers argued that disgust is *morally* significant, serving as a fitting and appropriate emotional response to certain moral wrongs (Plakias, 2018; Kumar, 2017; Hauskeller, 2006). Other authors attributed *political* significance to disgust, seeing it as indispensable to a properly functioning criminal law system (Kahan, 2000; 1998; Miller, 1997; Kekes, 1992). Of course, these claims sparked significant criticism and debate (see for example

Bollard, 2022; Giubilini, 2016; Gert, 2015; Kelly and Morar, 2014; Nussbaum, 2004; Knapp, 2003). And thus, the disgust advocate/skeptic debate was spawned.

It is frustrating to get halfway into making a cake and realize you're out of eggs. Unfortunately, disgust advocates may be in for a similar type of frustration. A crucial ingredient in their arguments is the standard assumption that "moral disgust" is an emotional capacity that humans really possess. However, recently, several philosophers, foremost among them Mara Bollard (2022), have recently developed a sort of ontological skepticism about so-called moral disgust (see also Fileva, 2021; Giubilini, 2016; Gert, 2015). They argue that the disgust advocate's standard assumption (moral disgust is a genuine emotional state) is vulnerable to two skeptical challenges, which I call the Metaphor Challenge and the Coincidence Challenge. The Metaphor Challenge holds that "moral disgust" is not really *disgust*; rather, much disgust language applied to moral offenses functions as a mere metaphor for anger (or a neighboring state like irritation, or resentment)—not genuine revulsion. The Coincidence Challenge, in contrast, holds that much so-called "moral disgust" is not really *moral*; rather, the sense of revulsion is directed at extraneous features of the offense (e.g. disgust at a bloody murder is directed at the bloodiness, not the murder *per se*). In such cases, non-moral disgust merely coincides with an independent moral judgment.

These challenges, if correct, threaten to undermine disgust advocacy altogether. If "moral disgust" is either not disgust or not moral, then *ipso facto* it can't carry the normative significance that advocates attribute to it.

The aim of this chapter is to refute the above challenges and vindicate the view—which I call "Genuine"—that moral disgust is a genuine emotional state. My response to the Metaphor Challenge is that it represents an excessive and unreasonable form of skepticism. While scientific

inquiry into Genuine is ongoing, the view rests on a strong theoretical and empirical foundation. In response to the Coincidence Challenge, I argue it is rooted in questionable empirical and metaethical assumptions. Proponents of this challenge tend to impose substantive constraints on what counts as a “moral judgment” (e.g. they must be justified in terms of harm, rights, or injustice)—raising, I hold, serious empirical and philosophical problems. Instead, we should recognize only formal constraints on what counts as a “moral judgment” (e.g. universalizability and categoricalness). When we do so, we find that the empirical evidence favors Genuine: many people’s disgust reactions and negative judgments of various behaviors are genuinely moral. The upshot of these arguments is that, *pace* ontological skeptics, Genuine is a strong hypothesis. The Metaphor and Coincidence Challenges fail, and disgust advocacy remains a contender in normative debates about disgust.

This chapter does *not* argue that any particular judgments involving moral disgust are justifiable, only that moral disgust is a genuine emotional capacity that humans possess. For many people, certain moral offenses are wrong *in a distinctive way*—that’s why they evoke disgust rather than anger (or some other state).

The plan for the chapter is as follows. Section 2 explains the Metaphor and Coincidence Challenges to Genuine, and why these threaten to undermine disgust advocates’ normative claims. Section 3 carves out a distinction between two traditions within disgust advocacy, which have different conceptions of “moral disgust,” and argues that these traditions face distinct ontological challenges. Namely, the Metaphor Challenge threatens a tradition I call “new advocacy,” and the Coincidence Challenge threatens “classic advocacy”—but neither challenge threatens *both* traditions (each is only vulnerable to one). Section 4 defends new advocacy against the Metaphor Challenge. Sections 5 and 6 defend classic advocacy against the

Coincidence Challenge. The chapter concludes that Genuine is a strong hypothesis and suggests some preliminary remarks on what this means for the dissertation's larger project.

2. Ontological Skepticism

2.1. Trends in the psychological literature

One would think that if there exists a distinctive emotional state of genuinely moral disgust, this means an emotional state can be identified that constitutes both (1) a genuine form of disgust and (2) a genuine moral sentiment (Bollard, 2022: 3). However, some empirical researchers claim no such state can be found. The candidates fail to satisfy one or both of the above criteria.

For example, regarding the first criterion (that the state must be genuine disgust), several psychologists over the years have suggested that disgust language applied to moral offenses serves as an instance of figurative speech and is not intended to be interpreted as a precise description of the speaker's emotional state. For example, Paul Bloom (2004) has written

My hunch is that in these statements 'disgust' is a metaphor. Saying that we are disgusted by a tax plan is like saying that we are thirsty for knowledge or lusting after a new car (pp. 172-173).

There is some experimental evidence to back this up. For example, Nabi (2002) found that people are reluctant to report feeling "grossed out" by moral offenses. Herz and Hinds (2013) achieved a similar result, leading them to conclude "moral disgust is not visceral (gross) but rather appears to be representative of anger even though autonomy violations are often labeled as 'disgusting'." More generally, Fiske (2020) argues that in general, much emotion research falls prey to the "lexical fallacy"—the mistaken assumption that colloquial speech maps neatly onto the emotions. Emotions are natural psychological entities existing independently of vernacular. As such, emotion researchers should be wary of taking lay usage of emotion terms at face value.

Regarding the second criterion (that the state must be a genuinely moral sentiment), a number of psychologists have suggested that, rather than disgust itself being a moral emotion, it is an emotion that can amplify or even induce moral judgments. For example, regarding amplification, Schnall et al. (2008) found that subjects who experienced extraneous disgust (e.g. walking through a room that smells of flatulence) just before rendering a moral judgment about a vignette involving a moral offense (e.g. sex between first cousins) judged the offense more harshly than control groups. Amplification aside, other authors have suggested that disgust can induce a moral judgment outright. For example, Inbar et al. (2009) found that intuitive moral disapproval of homosexuality correlates positively with individual “disgust sensitivity” (a dispositional proneness to be easily disgusted), suggesting that disgust “moralizes” homosexuality. In this case, disgust is shown to have a pernicious effect on human moral psychology; it injects noise into the moral reasoning process, ultimately causing distortion.

Experimental evidence aside, one might look at the scientific literature on disgust and conclude that, conceptually, disgust at moral offenses sounds out of place, like anger at the letter “B” or sexual attraction towards the Andromeda Galaxy. From an evolutionary standpoint, disgust originally developed to protect our hominid ancestors from parasites and pathogens. It’s often triggered by likely sources of biological danger—feces, rats, corpses, rotting meat, tactile properties evocative of bacterial growth (being moist or slimy)—things that may infect you if eaten or touched. The outputs of disgust—its characteristic facial expression (the “gape face”), neurophysiological profile (e.g. nausea), and behavioral tendency (distancing, rejection)—also point to these evolutionary origins. In crude terms, disgust is the “don’t eat, don’t touch” police. If it can be triggered by bad actions or bad people, this would be a “great departure” from its original function (Giner-Sorolla et al., 2018: 256).

Early disgust advocates, such as Leon Kass (1997), William Ian Miller (1997), John Kekes (2000; 1992; 1985), Dan Kahan (2000; 1998), and Mary Midgley (2000), did not engage much with empirical work on disgust in psychology. This was largely due to timing. Before the 1990s, little research on disgust existed (Olatunji and Sawchuk, 2005; Rozin, Haidt, and McCauley, 2008; Strohminger, 2014). Since then, however, research on disgust has exploded, with figures like Paul Rozin leading the charge. These early advocates couldn't have anticipated that empirical work on disgust, now a flourishing field, would sow doubt that the central emotion operating in their arguments is a metaphor, misnomer, or myth.

2.2. *Bollard's two ontological challenges*

Mirroring trends in the psychological literature, a number of philosophers have begun to question whether “moral disgust” exists (see Bollard, 2022; Fileva, 2021; Giubilini, 2016; Gert, 2015). This ontological skepticism threatens to undermine or seriously dilute advocates' normative views.

The clearest articulation of this view is found in Mara Bollard (2022). Bollard points out that the philosophical literature on moral disgust tends to assume, without argument, that it is a genuine emotional state, and the only question is whether or to what extent the emotion should guide or inform legislation, policy, or moral decision-making (for a notable exception, see Kelly, 2011: 128-132). However, this assumption is dubious, in her view. She suggests that we ask the following ontological questions: (1) Is so-called “moral disgust” really a form of *disgust*? (2) If so, is it a genuinely *moral* sentiment?

Using these questions as a guide, Bollard ultimately formulates two ontological challenges to the claim that moral disgust is a genuine emotional state: the Metaphor Challenge and the Coincidence Challenge.

The Metaphor Challenge holds that extensions of moral disgust language to moral offenses often function as a metaphor for some other emotional state (e.g. anger). In Bollard's (2022) words,

when people describe moral violations as 'disgusting' or use the term 'disgust' to refer to their emotional response to those violations, we are not licensed in inferring...that they are literally disgusted (p. 16).

On the contrary, an examination of the empirical evidence reveals that "the use of disgust language in response to moral violations is primarily indicative of anger" (p. 17).

Of course, for other proposed examples of "morally disgusting" offenses, it will seem implausible to deny that they evoke genuine disgust. For example, consider some of the cases suggested by John Kekes (1992) (my apologies in advance): "slowly disemboweling a person, dismembering someone with a chainsaw, slaughtering babies and bathing in their blood, or being drowned in excrement" (p. 30). To say that we are sickened by these "is not a metaphor," writes Kekes; on the contrary, they are literally nauseating (p. 433). Bollard is dubious about these examples as well, but for a different reason. Here, she introduces the Coincidence Challenge, which holds that disgust in response to some putative moral offenses is not really moral at all but, rather, physical or aesthetic. In her view, all of Kekes' proposals are "extremely physically disgusting." However,

it's not clear from these examples that what Kekes is calling moral disgust is being elicited by the distinctively moral properties of the wrongful actions rather than by the sheer volume of generically disgusting stimuli (p. 11).

Bollard goes on to examine two prominent accounts of moral disgust—those of John Kekes (1992) and Victor Kumar (2017). Ultimately, she finds that each fails to provide adequate support for a positive answer to her ontological questions outlined above. Kekes' account

succumbs to the Coincidence Challenge, while Kumar's succumbs to the Metaphor Challenge.

The upshot of all this is that disgust advocacy as a whole is significantly weakened:

[D]iscussions about [disgust's] normative relevance such as those between disgust advocates and skeptics about its fittingness or appropriateness will at best need to be reinterpreted in terms of generic disgust or anger (or both), or, at worst, significantly undermined (Bollard, 2022: 21).

2.3. *Ontological skepticism among other philosophers*

Bollard's ontological skepticism is the most clearly developed, but she is not alone in her skepticism; other philosophers have also, at times, raised doubts about the ontological status of moral disgust. For example, regarding the Metaphor Challenge, Martha Nussbaum (2004) also doubts that disgust language applied to moral offenses is anything more than a metaphor for anger. Rather, in her view, it is probably a case of "loose or careless usage," which commonly occurs simply because there is no alternative "affectively strong adjective with which to express anger" in English (p. 104).

Likewise, regarding the Coincidence Challenge, several philosophers have suggested that disgust advocates' favorite examples of "morally disgusting" offenses are really just cases of physical or aesthetic disgust coinciding with an independent moral judgment. For example, mirroring Bollard, Iskra Fileva (2021), Alberto Giubilini (2016), and Kelly (2011) criticize Kekes' examples of putatively morally disgusting offenses (e.g. "being drowned in excrement") on the basis that they are merely *physically* disgusting. Fileva (2021) dismisses Kekes' examples in a mere two sentences, calling them "non-ideal" and concluding "The acts described have features that can, all by themselves, provoke disgust, without any help from moral properties. (putrefying corpses, for instance)" (p. 23). Kelly (2011), in a brief footnote, dismisses Kekes' entire argument as "unhelpful rhetoric" (p. 163).

Focusing exclusively on Kekes' examples may seem like attacking a strawman. Surely there are stronger examples of moral disgust that can't be easily dismissed as the merely coincidental experiencing of two separate sentiments (one moral, the other non-moral). For example, consider the violations of food, sex, and death taboos that Kass includes in his list of "horrors": incest, bestiality, cannibalism, and corpse-mutilation. Don't these acts elicit genuine disgust (for some people) in themselves (not by virtue of involving revulsive extraneous features), and aren't they also widely judged to be wrong in themselves, independent of harm-, consent-, or injustice-based considerations?

Ontological skeptics are unconvinced. They tend to suggest that disgust at these practices is merely extraneous (directed at non-moral features) or moralizing (irrationally induces a moral judgment).³⁰ For example, Giubilini (2016) writes, "cannibalism and bestiality represent moral taboos in many cultures. In such cases, the disgust they elicit is a moralizing emotion: disgust is a fitting emotion and can cause people to moralize about them" (p. 233). Likewise, Gert (2015) writes

it is amazing how often the psychological literature on moral psychology makes use of cases of incest and bestiality as *paradigm* instances of immoral behavior. Justifying the classification of prohibitions on these behaviors as *moral* prohibitions because of their link to disgust has, among other problems, the following: it grants homophobia and other forms of disgust-linked prejudice a false respectability. (p. 54)

2.4. *Disgust advocacy is in trouble*

Bollard suggests that if "moral disgust" is either (1) not genuinely disgust or (2) not genuinely moral, then disgust advocacy is "significantly undermined" (p. 21). I agree. Not only does this reveal that advocates, in throwing around the phrase "moral disgust," are employing

³⁰ Bollard, Fileva, Gert, and Giubilini explicitly question whether disgust at incest and homosexuality is moral in nature. Bollard, Gert, and Giubilini explicitly question whether disgust at bestiality is moral in nature. Giubilini explicitly questions whether disgust at cannibalism is moral in nature.

muddled terminology, but also it eliminates any chance that “moral disgust” bears normative significance.

For example, suppose the Metaphor Challenge is correct. If “moral disgust” is actually just anger (or a neighboring state) in rhetorical disguise, this clearly undermines *epistemic* forms of disgust advocacy à la Kass (which posit that widespread public disgust at X is evidence that X is seriously wrong). Kass et al. still bear the burden of justifying their anger at X. After all, anger doesn’t carry epistemic or evidential authority (as if one’s anger at X is sufficient evidence that X is wrong); it can be an unreasonable emotion (e.g. someone accidentally bumps into me in the hallway and I fly into a rage). If a practice—such as embryonic stem cell research—makes you angry, and you encounter someone who doesn’t share your anger, you can’t win them over to your side by saying, “Trust me, embryonic stem cell research is wrong—I know this because it makes me angry.” That would be circular reasoning (citing the fact that you are angry as evidence that your anger is reasonable). Your interlocutor is entitled to know who or what, in your view, embryonic research wrongs, and how.

Next, suppose that the Coincidence Challenge is correct. If moral disgust is not *moral*, then *ipso facto* disgust lacks normative significance. After all, physical disgustingness is not a bad- or wrong-making property. Most people find putting ketchup on vanilla wafers disgusting (Rozin et al., 1999)—that doesn’t make it wrong. So, at best, being “morally disgusted” by a practice (e.g. stem cell research) means (1) one judges the practice to be wrong and (2) coincidentally, one is disgusted by it. But this gives no added weight to the moral judgment.

In summary, the Metaphor and Coincidence Challenges suggest that disgust advocates, in using moral disgust language as ammunition in ethical and policy debates, are making fallacious arguments. Either (1) they’re trying to legislate merely aesthetic preferences (e.g. one’s personal

distaste for stem cell research) by gaslighting us into thinking they're actually moral preferences (as per the Coincidence Challenge), or (2) they're trying to give their moral arguments against cloning and other biotechnologies (arguments which, considered alone, are not very persuasive) an additional air of authority by convincing us that disgust represents deep and inarticulate wisdom worth adhering to in advance of, or even absent, argumentation (as per the Metaphor Challenge). Either way, disgust advocacy is revealed to be problematic.

3. The Classic/New Advocacy Distinction

So, Bollard's two reasons for doubting the ontological status of "moral disgust" are the Metaphor and Coincidence Challenges. If these challenges hold, disgust advocacy is undermined. The question is, do they hold? Do we have reason to doubt that moral disgust is a genuine emotional state?

I think Bollard and the other ontological skeptics are wrong. There is sufficient evidence to believe that moral disgust is a genuine emotional state. However, to explain why I think this, I first need to detail a distinction—overlooked by Bollard—between what I call "classic disgust advocacy" and "new disgust advocacy."

In the literature, the label "disgust advocate" refers to a range of views—really anyone who endorses one or more of the following claims: (1) *Moral significance*: Disgust is a fitting or appropriate response to some moral offenses. (2) *Political significance*: Disgust is indispensable to a properly functioning criminal law system. (3) *Epistemic significance*: Disgust can alert us to the (serious) wrongness of actions or practices whose wrongness we might otherwise overlook.

While this mapping of the different views within disgust advocacy is helpful, it is incomplete. Disgust advocates not only attribute different kinds of normative significance to disgust but also have different conceptions of the emotion itself. Classic disgust advocates

conceive of moral disgust as primarily triggered by what we can call “purity violations”—actions that are intrinsically offensive insofar as they threaten to corrupt one’s moral character, violate “natural” boundaries, or accelerate social decay (more on this momentarily). Among those I consider to be classic advocates are Leon Kass (1997), John Kekes (2000; 1992; 1985), Dan Kahan (2000; 1998), Mary Midgley (2000), and William Miller (1997).

In contrast, new disgust advocates, such as Victor Kumar (2017) and Alexandra Plakias (2018; 2013), conceive of moral disgust as primarily triggered by “anti-social violations”—moral offenses that pollute beneficial social interactions.

Once we recognize this division within disgust advocacy, we see that each side has strengths and weaknesses. Namely, classic advocacy is not vulnerable to the Metaphor Challenge because purity violations elicit genuine disgust. However, it may be exposed to the Coincidence Challenge, as some authors question whether purity violations are really *moral* in nature. On the other hand, new advocacy is not vulnerable to the Coincidence Challenge, because anti-social violations typically have no extraneous disgust features and are clearly morally wrong. However, new advocates may be exposed to the Metaphor Challenge; some authors are dubious about the empirical evidence suggesting that anti-social violations trigger genuine disgust, rather than anger or some other state.

Therefore, in asking “Is Bollard right?” we’re really asking two different questions that need to be addressed separately:

1. Does classic advocacy succumb to the Coincidence Challenge?
2. Does new advocacy succumb to the Metaphor Challenge?

Below, I will briefly explain the difference between purity violations and anti-social violations and how this distinction informs different conceptions of moral disgust within disgust advocacy.

3.1. Purity violations

In moral psychology, a purity violation is an action or practice perceived as intrinsically offensive—dirty, debased, or twisted—that threatens to debase oneself or pollute the moral fabric of society. Examples include consensual incest, harmless cannibalism, receiving a blood transfusion from a child molester, and having sex with a dead chicken (Rozin et al., 1999). Purity violations are different from other types of moral offense, like harm or injustice; they are wrong *in a particular way*.

The concept of a purity violation originates in Richard Shweder's (1997) theory of the Three Ethics of Morality, which proposes that three distinct languages or codes underlie the moral systems of most cultures: (1) the *ethic of autonomy*, which focuses on rights, fairness, and justice; (2) the *ethic of community*, which focuses on duty, hierarchy, and proper fulfillment of social roles; and (3) the *ethic of divinity*, which views the body as a “temple” and focuses on protecting the spark of divinity in the human being from degradation or pollution.

Shweder originally developed this theory as a tool for cross-cultural analysis of moral systems, an area where it has produced fruitful insights. As Rozin et al. (1999) explain, “All three ethics may be available to individuals in all cultures, but cultures appear to show preferences and differential patterns of use of the three ethics” (p. 438). For example, in Hindu India, the ethic of divinity is highly salient; in Japan, the ethic of community is highly salient; and in the U.S., the ethic of autonomy is highly salient (the Japanese and American cases were discussed in Chapter 2, Section 6).

Building on Shweder, Paul Rozin et al. (1999) developed the CAD Triad Theory, which suggests that the three “other-condemning emotions”—contempt, anger, and disgust—are associated with different moral domains, roughly corresponding to Shweder's three ethics. Contempt is triggered by violations of the community ethic (e.g. disobeying or disrespecting an

authority figure); anger is triggered by violations of the autonomy ethic (e.g. harming or exploiting someone); and disgust is triggered by violations of the divinity ethic (e.g. consensual incest). Experimental evidence supports the claim that people primarily feel disgust toward violations of the divinity ethic and not other types of violation (Landmann and Hess, 2017; Rozin et al., 1999).

Jonathan Haidt's Moral Foundations Theory (MFT) further develops these ideas (see Haidt, 2012; Haidt and Graham, 2007; Haidt and Joseph, 2004). As explained in the previous chapter (Chapter 2, Section 6), MFT proposes that humans have innate sensitivities—"moral tastebuds"—to discrete moral domains ("moral foundations") that cultures use as building blocks to construct moral systems. As children grow and mature, the surrounding culture shapes the development of these sensitivities, amplifying some and downplaying others. This process ultimately explains why there is cross-cultural moral diversity (e.g. U.S. vs. India) and intracultural moral diversity (e.g. U.S. liberals vs. conservatives) (Haidt, 2012; Haidt and Graham, 2007; Haidt and Joseph, 2004).

Haidt identifies several moral foundations, including (1) harm/care, (2) fairness/reciprocity, (3) authority/respect, (4) loyalty/reciprocity, and (5) purity/sanctity. According to Haidt, each foundation has roots in biological evolution. For example, the harm/care "tastebud" originally evolved to help mothers detect signs of suffering in their offspring. The purity/sanctity tastebud originally evolved to protect humans from pathogens and parasites. However, over time, cultures "elaborated" these systems, making them sensitive to new concerns. The harm/care tastebud expanded to all kin, which is why people today are sensitive to the suffering of "other people's children, starving adults on television, images of baby seals being clubbed to death, and our pets dogs" that moan and whine when we leave for

work (Haidt and Joseph, 2004: 60). Likewise, the purity/sanctity domain was expanded from purely biological concerns to regulate “a great many bodily functions and practices,” such as “menstruation, eating, bathing, sex, and the handling of corpses” (Haidt and Joseph, 2004: 60). This has led to the “profound moralization of the body and bodily activities” in cultures such as “Judaism, Hinduism, Islam, and many traditional societies” as well as “American political conservatives” (Haidt and Joseph, 2004: 60).

Because the purity/sanctity foundation has roots in the evolutionary need to guard the body from pathogens, its extension into the moral domain retained much of this sense of embodiedness (purity violations often involve “the body and bodily activities”). However, the (perceived) threat is less to the body than the self or soul contained within. For example, Haidt has found that many people morally condemn consensual incest (even if pregnancy is virtually impossible), not because it threatens physical illness, but rather because it threatens to debase the self, violate the natural teleology of family relationships, or evoke fears of social decay (Haidt and Hersh, 2001; see also Rozin et al., 2008; 1999).

3.2. Anti-social violations

While Rozin, Haidt, and others claim that purity violations elicit a morally tinged form of disgust, other researchers challenge this. Chapman and Anderson (2013), for example, suggest that purity violations elicit “classic physical disgust”:

Examples of purity violations in the literature include engaging in consensual incest, receiving a blood transfusion from a child molester, and eating rotten meat. It is not surprising that people find such scenarios to be disgusting, given that they describe classic physical disgust stimuli. However, because physical disgust and immorality are confounded, it is not clear whether any of this disgust can be attributed to the immoral nature of the actions described.

Chapman and Anderson (2013) are more interested in whether “transgressions that do not reference physical disgust stimuli can elicit signs of disgust experience” (p. 304).³¹ In their meta-analysis, they compile evidence that moral disgust can be triggered by “anti-social violations” (p. 304).³² Examples include cheating, lying, stealing, free-riding, and other “reciprocity violations” that breach shared expectations of trust and cooperation.

As evidence that anti-social violations trigger disgust, Chapman and Anderson (2013) highlight data showing that subjects respond to these violations by exhibiting facial, neurophysiological, and behavioral outputs characteristic of disgust. Physical disgust—the kind you feel in response to stimuli like feces and rotting meat—is consistently associated with (1) *a facial expression*—the “gape face”—characterized by wrinkling along the bridge of the nose, retraction of the upper lip, and slight protrusion of the bottom lip, (2) *neurophysiological arousal*, particularly nausea and increased activity in a specific brain region called the anterior insula, and (3) *a behavior pattern*—called “distancing”—which involves creating space between oneself and the disgust-object, be it physical or attentional (see Rozin et al., 2008; Bloom, 2004; Rozin and Fallon, 1987; Angyal, 1941).

Since physical disgust is reliably associated with these measurable outputs, researchers can test whether subjects’ reported “disgust” at anti-social violations reflects genuine internal disgust by exposing them to moral stimuli and observing how they respond. If subjects respond with these characteristic outputs (e.g. they display the gape face and show increased activity in

³¹ Anti-social violations are self-consciously defined to not be confusable with purity violations. For example, Chapman and Anderson (2013) define “physical disgust stimuli” as actions or practices that “violate moral codes related to purity or divinity, which forbid behavior that is polluting, filthy, inhuman, or profane.” (p. 304). It seems unfair, to me, to equate violations of “moral codes related to purity or divinity” with “physical disgust stimuli,” but alas.

³² Their actual term is “pure moral violation.” I’ve changed this to “anti-social violation” because (1) when disgust researchers talk about “pure moral violations,” they’re almost always talking about anti-social behavior anyway, and (2) using the word “pure” to define a concept that contrasts with “purity” is needlessly confusing.

the anterior insula), this is strong evidence that their internal state is one of genuine disgust. Chapman and Anderson (2013) proceed to highlight studies that do precisely this. For example, they reference research showing that subjects who receive unfair offers in economic games³³ exhibit (1) activation of the levator labii superioris (the facial muscle responsible for the wrinkling along the nose and the upper lip retraction involved in the gape face) and (2) increased activation in the anterior insula (a neural correlate of disgust associated with processing aversive stimuli).

If Chapman and Anderson are right that anti-social moral violations can trigger disgust, this represents an example of a “pre-adaptation,” where an existing system (physical disgust) was extended to serve a new function it wasn’t necessarily selected for in evolution, namely, enforcement of group norms (p. 301, 391; see also Rozin et al., 2008). This extension was likely driven by cultural, rather than biological, evolution. But why would cultures in the distant past have needed to co-opt disgust to police norms? Why wouldn’t other emotions, like anger, suffice for norm-enforcement? They suggest,

This co-option may have been driven by the association between disgust and withdrawal motivation, which could have been a valuable addition to the behavioral repertoire of our highly social species (Chapman and Anderson, 2013: 322).

In a related line of inquiry, Joshua Tybur (2009) proposes a sort of game theoretic explanation for why cultures co-opted disgust to police social norms. He explains that humans evolved in “close-knit groups in which interdependence and social exchange were the norm” (p.

³³ The game in question is the “Ultimatum Game.” In this game, one player is given a pot of money and instructed to offer a portion of it to a second player. The second player can either accept or reject the offer; if they reject, no one gets any money. In purely financial terms, the second player is better off taking any offer, even an unfair one (90/10 is better than 0/0). Nevertheless, unfair offers are typically rejected, and the likelihood of rejection increases in proportion as the offer is unfair (e.g. 90/10s are rejected more frequently than 70/30s). This suggests that the second player expects to receive a fair offer and views the first player as a cheater for not giving one. I.e., unfair offers in the Ultimatum Game are viewed as reciprocity violations.

39). In these groups, social interaction introduced benefits to cooperators that couldn't be achieved individually. However, these interactions, which depended on reciprocity norms—shared expectations of cooperation and trust—were vulnerable to “social parasites” (cheaters, hypocrites, liars, thieves) who gained the benefits of social interaction without reciprocating, thereby imposing “substantial costs” on the group (p. 39). Therefore, to ensure that the benefits of reciprocity outweighed the costs, a punishment system was needed to deal with social parasites and police reciprocity norms.

Punishment could take two forms: second-party and third-party (p. 39-40). Second-party punishment refers to when the victim of a norm violation personally punishes the violator (e.g. someone steals from you, and you confront them to recoup your property). Third-party punishment refers to when someone not directly harmed by a norm violation shows willingness to punish the violator, thereby incurring personal costs without receiving direct benefits (e.g. someone steals from your neighbor, and you assist your neighbor in confronting the thief). Tybur notes that the benefits of third-party punishment to oneself are lower than those of second-party punishment; you might get the indirect benefit of deterring future theft, but the direct benefits (recovery of stolen property) accrue to the actual victim (p. 40). Given the benefits of third-party punishment are low and the costs potentially high—direct confrontation carries the risk of injury—individuals may be demotivated from engaging in it.

To properly motivate third-party punishment, it would be necessary to lower the costs, perhaps by dispensing punishment *indirectly*. Here, social exclusion—preventing someone from reaping the benefits of social interaction—would be the better option for dealing with social parasites.

This is where disgust comes in. Anger and disgust have contrasting motivational and behavioral tendencies. Anger is an approach emotion; it arouses the sympathetic nervous system (e.g. increases heart and respiratory rates), preparing the body for exertion and confrontation (p. 40-41; see also Rozin et al., 2008; Chapman and Anderson, 2013). Meanwhile, disgust is a withdrawal emotion; it arouses the parasympathetic nervous system (preparing the body for relaxation and recovery) and motivates avoiding contact with the disgust object. Therefore, if the group needs an emotion that motivates collective exclusion of social parasites, disgust fits the bill. It offers an alternative to angry fighting—passive rejection—thereby expanding groups’ repertoire for social regulation.

On this view, the “ultimate function” of disgust evolving to police social parasitism was not to protect the self from moral taint but to motivate the social coordination of a low-cost punishment system: social exclusion (Tybur, 2009: 41). In other words, the target of co-optation was disgust’s motivational and behavioral tendencies; any other features of the emotion that carried over from its biological origins (e.g. the tendency to feel viscerally repulsed and offended by triggers) are byproducts.

3.3. Classic advocacy is about purity violations

With this distinction between purity violations and anti-social violations on hand, it becomes evident, I believe, that disgust advocacy—a family of views attributing some form of normative significance to disgust—encompasses different conceptions of the emotion itself. One group—which I call classic advocates—conceive of moral disgust as primarily associated with purity violations, while another group—new advocates—see it as primarily associated with anti-social violations.

Among those I consider classic advocates are Leon Kass (1997), John Kekes (2000; 1992; 1985), Dan Kahan (2000; 1998), Mary Midgley (2000), and William Miller (1998; 1997). The most obvious sign that these advocates focus on purity violations is the examples they propose of putatively morally disgusting offenses. These include: (1) *emergent biotechnologies* (human cloning, enhancement drugs, “designer babies,” medical chimeras), (2) *taboo-breaking behaviors* (cannibalism, bestiality, necrophilia, incest, marketizing human organs, the use of cadavers in automotive crash testing), and (3) *instances of extreme violence* (genocide, torture, rape, overkill³⁴).

The adjectives classic advocates use to describe these acts indicate that they view them as intrinsically offensive. For example, Kass (1997) describes cloning as “Frankensteinian” (p. 686) and “grotesque” (p. 687). Midgley (2000) describes medical chimeras (e.g. a mouse genetically modified to grow a human ear on its back) as “monstrous” and “deeply uncanny” (p. 10). Kahan (2000) suggests that for some wrongs, such as genocide, “it is not enough to become angry when one hears of these atrocities; one must want to retch” (p. 66). In a particularly candid passage, Miller (1998) explains his opposition to cloning here:

We were not meant to be duplicated across generations, only represented vaguely. I know saying such a thing makes me out to be something of a know-nothing, who gives ready credence to an immutable human nature and who believes in some divinely or satanically ordered telos. But I mean no such thing. All I mean to say is that there are certain large constraints on being human and we have certain emotions that tell us when we are pressing against those constraints in a dangerous way. This is part of the job that disgust, horror, and the sense of the uncanny do; they tell us when we are leaving the human for something else; either downward toward the material, mechanical, and bestial, or upward toward the realm of spirit or the world of pure hokum (pp. 15-16).

³⁴ “Overkill” refers to a specific type of murder where the number of inflicted injuries far exceeds the number necessary to kill the victim. It can be an aggravating factor in sentencing decisions.

Classic advocates resist linking the wrongness of their examples to direct harm or injustice to persons. (When harm is involved, as in the cases of extreme violence, the offensiveness of the act inheres in the *manner* in which harm is inflicted, or the *cruelty* of the perpetrator.) For example, Kass (1997) explicitly rejects the suggestion that the wrongness of consensual incest is reducible to the genetic risks of inbreeding, and he is suspicious that such suggestions are disingenuous attempts to “rationalize away our horror” (p. 687).

Equally illustrative is Kahan’s (2000) discussion of the case of Dennis Beldotti. Beldotti was a murderer who strangled a female victim to gratify his sexual desires. While serving life in prison without parole, he requested that various items related to his crime (pictures of the victim, dildos, female undergarments) be returned from state custody to his representatives outside of prison. The state denied this request on the grounds that granting it would “justifiably spark outrage, disgust, and incredulity on the part of the general public” (quoted in Kahan, 2000: 67). The state added that “This has nothing to do with free expression. It has to do with the degradation of a young woman by a depraved individual” (quoted in Kahan, 2000: 67).

Kahan reports this ruling approvingly, stating that the decision will “strike nearly everyone as indisputably correct,” adding that “there is no viable basis for [the ruling] other than the one the court gave—namely, the disgustingness of Beldotti’s request” (2000: 67). Given that there is no other basis, Kahan argues that banishing disgust-based justifications from the law risks making it “morally blind” (2000: 65).

For classic advocates, the functional goal of disgust is to guard the sanctity of the soul or moral fabric of society from taint and decay. This becomes evident if you examine classic advocates’ writings and ask, “Hypothetically, what would happen if one did morally disgusting thing, X?” Regarding human cloning, Kass (1997) seems to worry that its proliferation would

cause people to become desensitized to something that, he believes, ought to evoke profound disgust. In several places, he laments (what he sees as) the decline of moral standards in America:

Today, one must even apologize for voicing opinions that twenty-five years ago were nearly universally regarded as the core of our culture's wisdom on these matters. In a world whose once-given natural boundaries are blurred by technological change and whose moral boundaries are seemingly up for grabs, it is much more difficult to make persuasive the still-compelling case against cloning human beings. As Raskolnikov put it, "man gets used to everything—the beast!" (p. 682).

The remark that "man gets used to everything," combined with his lament about declining moral standards, suggests that Kass worries human cloning will become normalized, and this normalization in itself is bad.

Likewise, Kekes (1992) argues that it is necessary for citizens in a civilized society to be not just angry about taboo violations—a term he uses for the worst crimes—but also disgusted. If citizens fail to be disgusted—if they become desensitized—this will not necessarily cause direct harm or injustice to anyone; rather, it will gradually lead to the erosion of our collective commitment to the basic norms upon which civilized society depends. From there, society risks sinking into "barbarism"—a form of life in which "moral considerations have no foothold" and there looms "the dreadful possibility of the removal of all limits" (p. 442).

3.4. New advocacy is about anti-social violations

Contrasting with classic advocates are new advocates, who conceive of moral disgust as primarily associated with anti-social violations. Two authors I consider new advocates are Alexandra Plakias (2018; 2013) and Victor Kumar (2017). Again, the most obvious sign that these authors focus on anti-social (not purity) violations is their examples of putatively disgusting moral offenses. These primarily include reciprocity violations—breaches of shared

expectations of cooperation and trust—like dishonesty, hypocrisy, cheating, stealing, free-riding, unfairness, disloyalty, betrayal, exploitation, and fraud.³⁵ For new advocates, these acts fittingly elicit disgust because they “pollute,” “contaminate,” “infect,” “circulate,” “spread,” and are “contagious.”

Kumar (2017) suggests that reciprocity violations “pollute social relationships by eroding shared expectations of trust” and can “spread easily through a population” (p. 9). He marshals empirical evidence that these violations, left unchecked, can spread throughout communities like a disease. For example, he highlights studies on public goods games where it’s in everyone’s best interests to contribute money to the collective pot; however, if some individuals defect and aren’t punished, “defection rapidly spreads across the pool of participants” (p. 10). Thus, defection is “polluting” and calls for a “collective response” (p. 9).

Enter disgust. It is a “fitting” response to reciprocity violations because it evolved to track and mitigate the spread of contaminants, and these violations *are* genuine contaminants—albeit they pose a threat “not to the physical body but to the social body” (p. 4). Disgust mentally represents its object as contaminating, and this representation is accurate for these violations. Additionally, disgust motivates withdrawal—the best social policy for mitigating the spread of contaminants. Echoing Tybur (2009), Kumar writes,

Anger motivates blame, confrontation, and threats, and thus functions to halt and discourage immoral behavior. But disgust motivates a less risky form of punishment that is often neglected: social exclusion. (p. 10)

³⁵ There’s one moment in Kumar (2017) where he suggests that incest—a paradigm purity violation—is morally disgusting (p. 5). However, according to Kumar, the reason incest elicits disgust is that it is “exploitative.” It’s unclear what he means by this. Maybe he means incest is always non-consensual sex—e.g. it involves a more powerful family member using their power to pressure another family member into unwanted sex. But in such cases, the wrong-maker is lack of consent. Therefore, on Kumar’s view, incest is wrong for the same reason that rape between non-genetically related individuals is wrong (lack of consent). This sort of reasoning is evocative of the harm/care or fairness/reciprocity domains, not the purity/sanctity domain. That is, Kumar doesn’t seem to see incest as *intrinsically* offensive (it is, after all, possible to imagine cases of perfectly consensual incest). Therefore, I stand by my claim that Kumar’s (2017) account of moral disgust is not concerned with purity violations.

He adds,

It often makes sense to avoid and exclude people who commit reciprocity violations. . . . Imagine you discover someone in your social circle is a hypocrite. He has been urging you not to do something, but has been doing that same thing behind your back. Often, an appropriate response is to sever the interpersonal relationship by creating physical and social distance—to avoid him, to no longer respond to appeals to engage with him, and to encourage other people to behave similarly. (p. 10)

Alexandra Plakias' (2018; 2013) proposes a similar account of moral disgust. Like Kumar, she suggests that some moral wrongs have “contamination potency”—potential to rapidly spread throughout a community. Her examples include heavy drinking, smoking, obesity, suicide, and divorce. She cites empirical evidence that an individual is more likely to develop these conditions or behaviors if their friends or others in their social circle develop them. For example,

An individual is 50% more likely to drink heavily if a friend does so as well, but the effect remains strong even at two degrees of separation—if a friend of a friend drinks heavily, the individual is 36% more likely to do so (2013: 275).³⁶

The “truly striking” point, Plakias says, is that the effect remains even at two degrees of separation (p. 275). Harmful behaviors “spread not just through direct contact but through contact with those who have come into contact with it” (p. 275). Therefore, behaviors and attitudes “spread through a group in much the same way a disease does” (p. 275).

From here, Plakias suggests that disgust is a fitting response to wrongs that possess contamination potency. After all, disgust “tracks contamination,” and its motivational and action tendencies are characterized by withdrawal—the “most effective way to guard against” contaminants, whether biological or moral (p. 274).

³⁶ Plakias notes that the same effect holds for the other behaviors mentioned (smoking, obesity, suicide, divorce).

Plakias emphasizes that contamination potency *per se* is not a bad- or wrong-maker; good things, like cooperative behavior and happiness, can also be infectious (2013: 273, 275-276). What makes disgust (sometimes) a fitting and appropriate *moral* emotion is simply that some wrongs happen to be contaminating, and disgust happens to motivate the best policy for dealing with contaminants (withdrawal).

3.5. *Distinct ontological challenges for classic/new advocacy*

With the distinction between classic and new advocacy on hand, we are now positioned to see that, while both may face challenges from ontological skeptics, they are not subject to the *same* challenges. Classic advocacy is not vulnerable to the Metaphor Challenge, but may be exposed to the Coincidence Challenge. On the other hand, new advocacy is not vulnerable to the Coincidence Challenge, but may be exposed to the Metaphor Challenge. I explain below.

Recall that the Metaphor Challenge posits that extensions of disgust language to moral offenses should not be interpreted as literally conveying the speaker's emotional state; instead, disgust functions as a metaphor for another state, like anger. Classic advocacy is not vulnerable to this challenge because it focuses on purity violations, and empirical evidence demonstrates that these violations are primarily associated with disgust, not anger. For example, incest—a paradigmatic purity violation—has been found to cause many people to experience “oral inhibition”—gagging, nausea, and loss of appetite (Royzman et al., 2008). These are unequivocal signs of disgust, not anger or any other state. More generally, Landman and Hess (2017) tested Moral Foundations Theory's claim that different emotions respond to different types of moral violation (for example, anger responds to harm/care violations, while disgust responds to purity/sanctity violations). Their findings supported this claim, and the disgust-purity link was found to be stronger than that between any other emotion-violation pair. Specifically, they found

that violations of the harm/care, authority/respect, fairness/reciprocity, and ingroup/loyalty foundations evoked a wide range of emotions, spanning anger, resentment, fear, and contempt. Meanwhile, disgust was “highly specific” and “clearly associated with” purity violations (p. 8). Moreover, disgust elicited by purity violations was “significantly stronger than disgust elicited by other violations” (p. 4).

While classic advocacy is not vulnerable to the Metaphor Challenge, it may be vulnerable to the Coincidence Challenge. Recall, the Coincidence Challenge holds that many purported examples of disgust at moral offenses conveniently involve features that can elicit physical (non-moral) disgust all on their own. This introduces doubt that the speaker’s condemnation is anything more than physical disgust coincidentally accompanied by an independent moral judgment (Bollard, 2022: 10; Fileva, 2021: 23; Giubilini, 2016: 235). For example, as discussed earlier, Kekes (1992) proposes several examples of allegedly “morally disgusting” offenses, such as “being drowned in excrement.” These examples have been criticized for involving extraneous disgust (e.g. excrement), and this raises doubt about their legitimacy. Similarly, many of Kass’ (1997) examples—incest, bestiality, human cloning—may be seen as extraneously disgusting. That is, an evil person with no regard for morality may still be disgusted by these activities (Fileva, 2021). It doesn’t help classic advocates’ case that their examples often don’t involve direct harm, injustice, or rights-violations to any victims. Indeed, they often deliberately resist linking their examples to direct harm, injustice, or rights-violations.

A similar story as the above unfolds with new advocacy. New advocates’ account of moral disgust is not vulnerable to the Coincidence Challenge because they focus on reciprocity violations, which (1) are clearly genuine *moral* offenses—they harm, cause injustice, setback interests, erode shared expectations of cooperation and trust, and so on—and (2) do not involve

extraneous disgust. For example, liars, despite being said to “reek” of dishonesty, do not literally emit an unpleasant smell that might coincidentally trigger a physical disgust response.

New advocates may, however, be vulnerable to the Metaphor Challenge. As discussed earlier, disgust is a natural psychological entity with measurable indicators: a facial expression (the gape face), a neurophysiological profile (nausea, anterior insula activation), and a characteristic behavioral tendency (distancing and contamination-sensitivity). Are these present in people’s reactions to anti-social violations? Some authors believe the empirical evidence warrants a “Yes” answer (e.g. Plakias, 2018; 2013; Kumar, 2017; Chapman and Anderson, 2013). But ontological skeptics challenge this. For example, Mara Bollard (2022), Iskra Fileva (2021), and Joshua Gert (2015) argue that proponents of the view that anti-social violations genuinely elicit disgust bear the burden of proof, and this burden remains unmet because the evidence that proponents provide tends to be equivocal (open to competing interpretations) (see Section 4.1 below).

3.6. Conclusion

This section argued that there is a crucial distinction between classic advocates, who associate moral disgust primarily with purity violations, and new advocates, who associate it with anti-social violations. I then argued that each form of advocacy has individual strengths and weaknesses: classic advocacy is invulnerable to the Metaphor Challenge but may be vulnerable to the Coincidence Challenge, while new advocacy is invulnerable to the Coincidence challenge but may be vulnerable to the Metaphor Challenge.

The purpose of the rest of the chapter is to examine whether either form of advocacy succumbs to its corresponding ontological challenge. I will argue that both ultimately withstand ontological scrutiny.

4. New Advocacy and the Metaphor Challenge

Let's start with new advocacy. Ontological skeptics like Bollard (2022), Fileva (2021), and Gert (2015) reject the claim that anti-social violations elicit genuine disgust. They propose that self-reports of disgust towards these violations should be interpreted as metaphorical speech—people are expressing anger (or irritation, or resentment) using disgust language (“Your racism is nauseating!”) to increase rhetorical impact.

While I acknowledge that the evidence that reciprocity violations elicit genuine disgust could be more robust, skeptics, in my view, downplay its strength and frame the discussion in a way that's unfavorable to new advocates. They suggest that advocates (1) bear the “burden of proof,” and (2) to satisfy this burden, must supply “unequivocal evidence.”

This is unreasonable. A more reasonable standard would be (1*) tell a compelling theoretical story about why humans might have evolved to feel disgust toward anti-social violations, and (2*) provide good evidence that this is in fact the case. These are standards that the disgust advocate *can* meet.

4.1. *Two interpretations of self-report data*

Self-report data make the following indisputable: people frequently apply disgust language to anti-social violations, both in English and many other languages (Rozin et al., 2008; Curtis and Biran, 2001). The question is how to interpret these self-reports: do they literally convey the speaker's internal state, or are they figurative expressions of anger?

Bollard (2022) favors the latter. She writes,

when people describe moral violations as ‘disgusting’, we are not licensed in inferring, as Kumar [2017] does, that they are literally disgusted (p. 16).

In her view, this inference relies on the dubious assumption that “scholars and laypeople use emotion terms in a way that reflects a shared meaning” (p. 16). She cites empirical evidence

showing that people use the terms “disgust” and “disgusted” differently than “gross” and “grossed out” (see Herz and Hinds, 2013; Nabi, 2002). Such findings, Bollard thinks, count against the literal interpretation.

In short, we have two competing interpretations of the available self-report data:

Genuine: Anti-social violations elicit genuine disgust.

Metaphor: Anti-social violations elicit some other emotional state (e.g. anger), which is analogized as disgust for rhetorical purposes.

In Bollard’s (2022) view, “without more investigation into what participants mean when they use disgust language, we can’t be sure whether self-reports of disgust refer to generic disgust or anger” (p. 16). She rejects the inference from self-report data to Genuine because Metaphor is also compatible with the data.

If self-report data are inconclusive, how else can we test our two hypotheses? One option: “operationalize” disgust (Bollard, 2022: 4; Fileva, 2021: 21). As already noted, physical disgust—the kind you feel in response to feces and rotting meat—is associated with a characteristic set of expressive, neurophysiological, and behavioral outputs. Given this, researchers can test whether subjects’ reported “disgust” at anti-social violations reflects genuine internal disgust by exposing them to moral stimuli and observing how they respond. If subjects respond with these characteristic outputs (e.g. they display the gape face and show increased activity in the anterior insula), this is evidence that their internal state is one of genuine disgust.

Kumar (2017), a proponent of Genuine, cites two studies in this space: Chapman et al. (2009) and Sanfey et al. (2003). Both studies examine how subjects emotionally respond to receiving unfair offers in economic games.

Chapman et al. (2009) used electromyography to record subtle muscle activity in the faces of subjects who received unfair offers and compared this to subjects experiencing distaste

(in response to a bitter food) and generic disgust (in response to feces). They found that “all three states evoked activation of the levator labii muscle region of the face” (p. 1222)—recall, that’s the muscle responsible for the wrinkling along the nose and upper lip retraction in the gape face. They conclude, “These results suggest that immorality elicits the same disgust as disease vectors and bad tastes” (p. 1222).

Sanfey et al. (2003) used fMRI machines to examine the brain activity of subjects receiving unfair offers. They found that subjects experienced increased activity in the bilateral anterior insula, and activity increased in correlation with the degree of unfairness. They conclude,

Though studies of disgust have largely focused on physical sensations of taste and odor, it has been suggested that emotion-based disgust (as perhaps induced by an insultingly unfair offer) may be conceptually similar. The recruitment of similar neural structures, namely the anterior insula, in both physical and moral disgust gives some credence to this notion (p. 1756).

After citing these two studies, Kumar (2017) writes,

disgust is among the emotions that underpins this behavior. ... Empirical evidence suggests that cheating, dishonesty, and exploitation often elicit disgust (pp. 9-10).

In other words, he believes that (1) levator labii activation (LLA) and (2) increased anterior insula activation (IAIA) in response to moral stimuli are evidence of disgust:

- (1) LLA → Genuine
- (2) IAIA → Genuine

Are (1-2) valid inferences? Bollard, Fileva, and Gert think not. They reject (1) because people sometimes make facial expressions only for “communication or signaling purposes” (Bollard, 2022: 17). That is, the gape face may be “performative,” not indicative of “actual disgust” (Fileva, 2021: 22). Thus, facial expression data “lends itself to more than one interpretation and is insufficient to rule out alternative hypotheses such as the ‘signaling’ one”

(Fileva, 2021: 22). In Bollard's (2022) words, the data "do not establish the claim that participants are responding to reciprocity violations with disgust rather than some other emotional response" (p. 17).

They also reject (2) because anterior insula activation is "not uniquely associated with disgust" (Bollard, 2022: 17). It has also been associated with other negative states, including "pain, distress, hunger, thirst, and autonomic arousal" (Fileva, 2021: 23) as well as "fear and anger" (Bollard, 2022: 17). Thus, anterior insula activation "cannot be taken as clear evidence of a disgust response" (Bollard, 2022: 17). In Fileva's (2021) words, this source of evidence is "far from conclusive" (p. 23).

4.2. *Too much skepticism*

I believe Bollard and Fileva are excessively skeptical of Genuine. They each suggest that proponents of Genuine (1) bear the burden of proof and (2) to meet this burden, must provide unequivocal evidence (evidence that leaves no room for doubt). For example, Bollard (2022) writes,

to successfully shoulder the burden posed by the ontological challenge, the moral disgust proponent must provide *unequivocal* evidence that genuinely moral disgust, and not some other psychological state (notably, anger or generic disgust), is being elicited by (features of) relevant moral violations (p. 19, original emphasis).

What does "Proponents of Genuine bear the burden of proof" mean? Presumably, it means Genuine represents an initially implausible view. In a Bayesian sense, even before we examine the evidence, the initial probability of Metaphor is higher than that of Genuine, such that if the evidence turned out to support Genuine over Metaphor, this would be *surprising*. Thus, dialectically, we should treat Metaphor as the default hypothesis.

I don't have a problem with burden-of-proof talk in principle, or with asking one's opponents to provide evidence for their assertions. However, it's possible to be mistaken about where we place the burden of proof, what type of evidence we ask for (e.g. "unequivocal" vs. "good enough for the domain of inquiry"), and how we set the initial probabilities of the relevant competing hypotheses.

For illustration, imagine Ally and Billy are discussing the philosophical problem of other minds. Ally says, "I believe that other people have minds and thoughts, just like I do." Billy responds, "I beg to differ. Other people don't have minds." Billy then adds, "You cannot prove that other people have minds. Of course, *I* don't have to support my claim. *You* bear the burden of providing unequivocal evidence for your claim."

Is this discussion off to a good start? No. First, unequivocal evidence for the existence of other minds is impossible. Any empirical evidence Ally can muster will necessarily leave room for doubt and be open to alternative interpretations. Second, treating "other people don't have minds" as the default hypothesis—as if its being false would be surprising—is backwards. Clearly, people *not* having minds would be surprising (or, at the very least, Ally's and Billy's two hypotheses are equiprobable).

Bollard and Fileva make a similar mistake as Billy. They (1) treat Genuine like a remote possibility and (2) demand unequivocal evidence. Both are mistakes. I will explain why (1) is problematic in Section 4.3, and why (2) is problematic in section 4.4.

4.3. *The initial probability of Genuine*

As discussed in Section 3.3 above, it makes perfect sense why it would have been adaptive for human communities in the distant past to extend the disgust capacities of their members to police anti-social violations. The prosperity of these communities depended on third-

party enforcement of reciprocity norms, but direct enforcement would have been too costly. Fortuitously, disgust—with its built-in motivational tendency of withdrawal—presented a solution to this problem: enforce norms indirectly via social exclusion. So, while disgust initially evolved for a biological purpose, it was conveniently available to be co-opted by cultures to serve a new function (motivating indirect collective punishments).

This is a plausible story. As such, even before we examine the evidence, we have reason to *expect* that Genuine may be true—at the least, its being true would *not* be surprising.

Perhaps skeptics are willing to accept this story for what it is, but they don't see how disgust (the poison/parasite police) could have actually extended to an unrelated domain (morality) just because it would be *convenient*. That is, they want to see the specific mechanisms that allowed this evolutionary process to happen.

In that case, we do have an idea of the specific mechanisms at play. As discussed in Chapter 2, Section 5, empirical work on the ontogeny of disgust shows that cultures already influence the specific triggers we acquire. There is cultural diversity in what triggers physical (non-moral) disgust. In the U.S., we eat stinky, curdled, fermented milk (cheese); in some other cultures (e.g. traditional Chinese), this is considered disgusting (Chapman and Anderson, 2013: 319). In Fiji, eating shark meat is taboo; no such taboo exists in the U.S (Henrich, 2015). Several authors note that cultural variability in what elicits disgust is necessary because (unlike other mammals) humans inhabit highly varied environments; the poisons and parasites that humans have to worry about differ from one place to the next (Kelly, 2011; Rozin et al. 2008). Thus, it wouldn't make sense for humans to be born pre-programmed with a list of disgust triggers. We must figure out what food sources are parasitic or infectious via trial and error.

Fortunately, this experiential learning needn't be done by us personally. Our parents and surrounding community clue us into what's unsafe, often just by signaling that X is unsafe by making the gape face at it (Oaten et al., 2014; Kelly, 2011; Oaten et al., 2009; Rozin et al., 2008). This causes vicarious disgust, so you can actually develop an aversion to something without ever having a bad experience with it yourself. Children are born with no disgust triggers (Rozin, in one study, found that 60% of children under two were willing to taste imitation dog feces) (Rozin and Fallon, 1987). This initial lack of disgust provides a grace period for children to acclimate to local eating customs and sponge up local knowledge (Bloom, 2004). Disgusting stimuli have been found to have a strong link to memory, being more easily remembered than neutral stimuli or even other negative stimuli (e.g. scary things) (Chapman et al., 2013), which makes sense if humans must memorize long lists of biological dangers in the environment. Eventually, around age four, children suddenly become "neophobes," suspicious of all new foods they haven't been introduced to yet (Bloom, 2004). All this indicates that children are born with a "preparedness" to develop disgust, and within limits, what specific triggers they acquire are culturally influenced (Bloom, 2004; Rozin and Fallon, 1987). In other words, disgust comes with a syntax, but the semantics comes later.

This picture of the ontogeny of disgust is favorable to the view that there are mechanisms available that cultures could have used to elaborate disgust. Indeed, some authors have explicitly invoked the ontogeny of disgust to account for cross-cultural variability in moral disgust elicitors, and also to account for the evolutionary extension of people's disgust capacities to the moral domain (see Rozin et al., 2008).

In summary, if humans cultures in the distant past would have accrued adaptive benefits extending their members' disgust capacities to anti-social violations, and if this extension was

really possible (given the ontogeny of disgust), then there is no reason to treat Genuine as a remote possibility that would be surprising if true. On the contrary, it is a strong hypothesis.

In my view, if we consider the evidence for Genuine again, this time with the hypothesis framed as strong, it has an undeniable gestalt. For illustration, consider this passage from Tybur (2009), where he summarizes multiple pieces of evidence:

In a study of the relationship between disgust and morality across cultures, Haidt et al. (1997) found that all languages surveyed use the same word to describe moral violations and infectious entities like feces and cockroaches, suggesting that the existence of moral “disgust” is not simply a quirk of the English language, but is a reliable cross-cultural phenomenon. Jones and Fitness (2008) found that Australian psychology students who read vignettes about crimes involving drug trafficking, conning, fraud, or theft were more likely to form words associated with pathogen disgust in a word-stem completion task relative to controls. Wheatley and Haidt (2005) demonstrated that participants hypnotized to feel disgust toward arbitrary words had more severe moral judgments of morally questionable individuals. Schnall, Haidt, Clore, and Jordan (2008) found that disgusting odors and videos increased the severity of moral judgments relative to controls, and Schnall, Benton, and Harvey (2008) found that individuals primed with “cleanliness” (either through washing their hands or being exposed to semantic primes) judged moral offenses less severely than controls. Chapman, Kim, Susskind, and Anderson (2009) found that the same facial muscles are engaged when participants were exposed to photos of contaminants relative to unfair treatment in an economics game. Even children as young as six years old describe immoral behaviors with the word disgust, and match photographs of disgust facial expressions to verbal descriptions of immoral activities (Danovitch & Bloom, 2009). Finally, multiple fMRI studies have shown overlap in the neural regions activated during moral judgment versus pathogen disgust (Moll et al., 2005; Schaich Borg et al., 2008).

4.4. The virtue-signaling hypothesis

Ontological skeptics tend not to challenge the theoretical point that disgust extending to the moral domain would have culturally adaptive benefits. Instead, they challenge the empirical data suggesting that this has, in fact, happened. The pattern usually goes like this: a piece of evidence, E, is said to support Genuine. Skeptics then introduce Metaphor and claim that E is compatible with Metaphor. This, they argue, undercuts the claim that E supports Genuine. They

then ask supporters of Genuine to provide unequivocal evidence. If such evidence is not forthcoming, this is considered a strike against Genuine.

Given the Tybur passage I shared earlier, there are too many specific pieces of evidence to examine individually. Instead, let me just focus on one example: facial expression data. While this source of evidence is “equivocal” in the sense that it’s open to more than one interpretation, it still provides *good* evidence for Genuine. That is because the “alternative explanations” of the data that ontological skeptics invoke are simply less compelling than Genuine.

The argument for Genuine based on facial expression data goes like this:

Levator Labii Argument

1. Reciprocity violations trigger levator labii activation (LLA).
2. LLA is indicative of the gape face.
3. The gape face is indicative of internal disgust.
4. Therefore, LLA is indicative of internal disgust. (from 2 and 3)
5. Therefore, Genuine (from 1 and 4).

Ontological skeptics Mara Bollard (2022), Joshua Gert (2015), and Iskra Fileva (2021) seemingly accept Premises 1 and 2 of the above argument but reject Premise 3. They argue that the gape face does not necessarily indicate internal disgust because people sometimes make facial expressions, like the gape face, only for “communication or signaling purposes” (Bollard, 2022: 17), or to be “performative,” which is not indicative of “actual disgust” (Fileva, 2021: 22). As evidence for this claim, Fileva cites Gert (2015) and Bollard (2022). Gert and Bollard, in turn, each cite Royzman and Kurzban (2011). Bollard also cites Giner-Sorolla et al. (2018).

So, what is this “signaling hypothesis” proposed by Royzman and Kurzban (2011) and Giner-Sorolla et al. (2018)? These authors both suggest that the gape face can be a *mere* social signal, but they offer different views on what message people are signaling and why. In my view, neither of their suggestions is convincing. Therefore, by citing these sources, Bollard, Fileva, and Gert are not raising a strong objection to Genuine.

The signaling hypothesis according to Giner-Sorolla

Start with Giner-Sorolla et al. (2018). They suggest that people often display the gape face in response to moral offenses as a means to signal their “moral virtue” (p. 228), “even when a different emotion [other than disgust] is felt internally” (p. 224). Here, “expressions of disgust” include “facial expression, tone of voice, bodily signals, and direct verbal assertions” (p. 277)—people really commit to the metaphor! They elaborate,

For a highly interdependent species like humans, it is vitally important to maintain a reputation as a prosocial and moral individual in order to reap the benefits of cooperation and avoid the costs of social exclusion (p. 271-272).

But why would you make the gape face, specifically, to signal your virtue? They suggest two reasons. First, you must signal your virtue somehow. This requires that you “publicly condemn the immoral actions of other people” (p. 272). But *how* should you do this? You could express anger; however, Giner-Sorolla et al. (2018) write, “our research has demonstrated that an expression of disgust is more likely to be perceived as morally motivated, whereas anger is more likely to be perceived as selfishly motivated” (p. 272). Thus, to maintain an honorable reputation, it’s more effective to feign disgust and disguise your anger.

A second reason why the gape face is a useful virtue-signaling tool is that it universally communicates “I don’t associate with X,” and you should definitely avoid being associated with immorality, especially if third-parties are watching: “people do not want to be *visibly* associated with immorality, so they treat immoral objects as if they are contaminating” (p. 274). To support this suggestion, they cite an additional study, Kupfer and Giner-Sorolla (2017). This study asked subjects whether they would prefer to wear a Nazi armband under their T-shirt (hidden from others’ view) or over their T-shirt (not touching their skin). Result: subjects overwhelmingly preferred to wear the armband under their shirt, citing reasons like “I would not want to seem

like I support Nazis” (quoted in Giner-Sorolla et al., 2018: 274). Giner-Sorolla et al. (2018) conclude that this finding contradicts *other* studies—for example, they cite Nemeroff and Rozin (2000), a review of such studies—which have found that clothes and other items associated with an evil person (e.g. Hitler’s sweater) are viewed as disgusting (aversive, contaminated). Since people prefer to wear the armband directly on their skin rather than over the shirt, this shows they are not, strictly speaking, *disgusted* by the armband; it “seems to lack the key cognitive element of contamination” (p. 275).

I do not find either of these proposals convincing. Being disgusted by something is generally a reason to avoid it, but this reason is not absolute. Even in the biological domain, eating potentially contaminated food is preferable to starving; thus, hunger has been found to mitigate disgust (Hoefling et al., 2009). The same principle applies to moral disgust. Nemeroff and Rozin—the two authors whose work Kupfer and Giner-Sorolla (2017) claim to debunk—explicitly cite one of their earlier papers, Nemeroff and Rozin (1994), which states that someone averse to wearing Hitler’s sweater might be convinced to do so if offered sufficient incentive (e.g. money):

It is also possible that even strong contagion-related feelings can be suppressed or disregarded in the face of pragmatic situational concerns. No doubt many of us who initially cringed at the idea would be willing to wear Hitler’s sweater for \$250 (p. 182).

All Kupfer and Giner-Sorolla (2017) did was present subjects with just such an incentive: *avoid being seen wearing a Nazi armband*. There are many disgusting things I (and presumably others) would (begrudgingly) do to avoid being seen in a Nazi armband, including: eat a vanilla wafer

dipped in ketchup, drink a glass of juice with which a sterilized cockroach has made brief and traceless contact, and eat soup from a sterilized bedpan.³⁷

But Giner-Sorolla et al. (2018) face a more fundamental problem. They do not adequately explain *why* people should care about signaling virtue, or *why* specifically disgust—the pathogen/parasite police—would turn out to be the best emotion for signaling virtue. In contrast, Tybur (2009) offers a clear and compelling reason why people display the gape face in response to anti-social violations: they are *genuinely disgusted*, as a result of human cultures in the distant past co-opting the motivational tendencies of disgust (aversion, withdrawal) to solve a collective action problem (namely, how to indirectly and collectively punish social parasites). Another way to state this point: Tybur (2009) explains why disgust would be a useful social tool *from the group's perspective*, whereas Giner-Sorolla et al. (2018) make it seem like signaling disgust at immorality is instrumentally rational on an individual level. But they don't explain why the conditions that make this behavior rational obtain. That is, why would others around us reward us with a reputation boost for feigning disgust outputs (saying “eww,” making the gape face, acting nauseous), as opposed to just saying explicitly “I don't associate with that behavior”? In contrast, Tybur's (2009) account has no such shortcomings.

The signaling hypothesis according to Royzman and Kurzban

Next, consider Bollard and Gert's other citation: Royzman and Kurzban (2011). These authors raise two different objections to Premise 3 of the Levator Labii Argument (the premise which holds “The gape face is indicative of internal disgust”). They begin by laying out two competing accounts for understanding facial expressions in general: (1) facial expressions are

³⁷ Each of these examples were found by Rozin et al. (2008) and Rozin et al. (1994) to be strongly aversive to most subjects.

“primarily direct read-outs of internal state,” and (2) facial expressions are “strategic elements of social performances” (p. 269-270). Next, they argue that the social performance account is clearly superior. The direct-readout account is implausible on theoretical grounds: “from the standpoint of evolved function,” it is generally disadvantageous “to signal to strategic interactants one’s private relevant strategic information” (p. 270). In other words, facial expressions as direct readouts would be maladaptive; we would broadcast to everyone what we’re thinking all the time, even when this would compromise our goals.

Royzman and Kurzban’s second objection is that facial expressions are “eminently self-fulfilling” (p. 270). They state that “a person who emits a particular (metaphoric) movement may actually come to be in a corresponding bodily and phenomenological state” (p. 270). In other words, socially displaying disgust—by feigning the gape face—can *cause* one to feel actual disgust. As such, facial expression data is “more of a liability than an asset” for defenders of Genuine (p. 270). After all, if a subject perceives a moral offense and responds with levator labii activation (LLA), the causal chain could be either of the following:

- (a) perceived offense → disgust → LLA
- (b) perceived offense → LLA → disgust.

Gert (2015) makes this same point:

if we signal this disapproval with a facial expression that is very close to the facial expression characteristic of disgust, this very expression may induce some of the symptoms of genuine disgust. That does not mean that the actual object of our disapproval strikes us as disgusting in any way (p. 41).

In my view, both of Royzman and Kurzban’s (2011) objections are very weak. Regarding the first, we obviously have *some* control over facial expressions. If I eat a meal prepared by my mom and don’t like the taste, I can conceal this by suppressing the urge to make a face.

Nevertheless, facial expressions clearly are reflexive to some degree. Human infants make the

gape face in response to tasting bitter foods (Steiner et al., 2001); this is clearly a reflex (oral rejection of bitterness), not a social signal. Even for adults, suppressing faces arising from internal states can be hard. For example, it can be hard to suppress a grin when someone we love does something stupid (e.g. confidently walking into a glass door), despite our desire not to embarrass them. This sort of thing can be a problem even for skilled actors; in British slang, “corpsing” refers to when an actor loses their composure and laughs or giggles inappropriately in a scene (e.g. while they are playing a dead body).³⁸ Likewise, here is Daniel Kelly (2011) discussing the difficulty of maintaining a good “poker face” while gambling:

Anyone who has played Texas hold ’em or seven-card draw can tell you how hard it is to keep a good poker face. It is extremely difficult to keep from broadcasting to the entire table your opinion of a hand... the amount of control it takes *not* to give away any information about what you are actually thinking or feeling, let alone to convincingly express misleading information, can be staggering. Genuine emotions seem to *want* to reveal themselves, especially in charged situations like poker games, and often at the worst possible moment. All but the most stoic and unflappable of us regularly reveal our emotions, even in spite of our best attempts to keep a lid on them. . . much of the recognition of emotions—appears to be largely automatic and difficult, if not impossible, to completely stop (pp. 61-62).

The same principle applies to disgust. The face that this internal state corresponds to—the gape face—is often reflexive and difficult to suppress.³⁹

It makes sense, by the way, that disgust would reflexively result in the gape face. From an evolutionary standpoint, the gape face serves the function of preparing the mouth for spitting or vomiting (expelling noxious contents from the mouth or stomach) (Chapter 2, Section 3.2). For this system to work, the gape face *needs* to be reflexive, not under fully voluntary or conscious control. So, while facial expressions may not be “direct read-outs of internal states,”

³⁸ Consider, for example, this interview with actors on the British sitcom *The Office*:
<https://www.youtube.com/watch?v=0J1HMxNPUBM>

³⁹ The first time I tasted a Manhattan with my father-in-law, I really wanted to act cool and not show on my face that the taste was incredibly unpleasant to me. Despite my best efforts, I failed, and the face came out.

they are at least partially reflexive. Thus, *pace* Royzman and Kurzban (2011), if someone makes a gape face in response to an anti-social violation, this is indeed reasonably interpreted as an indicator of internal disgust.

Now, regarding Royzman and Kurzban's (2011) second objection: it just seems dogmatic. They remark that facial expression data are "more of a liability than an asset" for defenders of Genuine. This is because, in their view, causal chain (b) (stimuli → LLA → disgust) is more likely than causal chain (a) (stimuli → disgust → LLA). But they give no reason at all to suppose this. Why would people feel the need to socially perform disgust at, for example, incest (their example, see p. 270)? They answer: the signaler may wish to communicate that, much like "feces," they find incest "unapproachable" and want "nothing to do" with it (p. 270). But this begs the question: why would one feel the need to signal that one finds it "unapproachable," as opposed to just signaling that one is angry, like with any other moral offense (e.g. someone robbing a bank or kicking a dog)? We need an explanation for that. In contrast, causal chain (a) doesn't raise this question. On causal chain (a), the subject treats incest as "unapproachable" because the subject finds incest *genuinely disgusting*.

In summary, in order to undercut facial expression data supporting Genuine, Bollard, Fileva, and Gert propose that these data are explainable by the signaling hypothesis. As support, they cite the two psychological studies: Royzman and Kurzban (2011) and Giner-Sorolla et al. (2018). In my view, these psychological studies simply do not make the signaling hypothesis seem very plausible, certainly not more so than Genuine. Therefore, Bollard, Fileva, and Gert are being excessively skeptical in rejecting evidence for Genuine in the form of facial expression data. If people (i) say they're disgusted by anti-social violations (self-report data), (ii) make expressions indicative of internally felt disgust (facial expression data), and (iii) there's a good

theoretical foundation explaining how and why disgust might have evolved to police anti-social violations, it seems obvious that the most parsimonious explanation for this is Genuine.

4.5. Conclusion

This section has argued that the Metaphor Challenge to new advocacy does not hold. There are strong theoretical and empirical reasons to believe that people do sometimes respond to anti-social violations with genuine feelings of disgust. Ontological skeptics, in placing the “burden of proof” on advocates, treat Genuine as if it has a low initial probability compared to Metaphor. This is incorrect; Genuine rests on a strong theoretical foundation. Additionally, there is a voluminous body of empirical evidence suggesting that Genuine is not just a strong theoretical possibility but also a reality. Ontological skeptics attempt to undercut this evidence by suggesting that it is open to “alternative interpretations” (e.g. facial expression data is alleged to be compatible with the virtue-signaling hypothesis). But these, upon inspection, tend to be less parsimonious than Genuine.

5. Classic Advocacy and the Coincidence Challenge: Defending Kekes

I now move on to defend classic advocacy against the Coincidence Challenge. The Coincidence Challenge holds that disgust in response to many (putative) moral offenses is not really moral in nature but, rather, merely physical or aesthetic (e.g. disgust toward a bloody murder is directed at the bloodiness, not the murder *per se*). This is a potential challenge for classic advocates, as explained previously, because classic advocacy focuses on purity violations, which often have features that, the skeptic argues, are disgusting in a merely physical sort of way. In other words, the claim “Purity violations elicit a distinctly *moral* sort of disgust” faces an undercutting defeater: our disgust in response to purity violations can be wholly explained as a reaction to extraneous disgust features, not morally relevant features. With classic advocates’

examples of allegedly morally disgusting offenses deflated, they have no examples left to draw on. This undermines their ontological assumption that *moral* disgust is a genuine, discrete emotional state.

In this section (6) and the next (Section 7), I will defend classic advocacy against this challenge. For organizational purposes, discussion is divided between purity violations that involve the property of harm (Section 6) and those that don't (Section 7). I deem this necessary because whether harm (a morally relevant property) is present or absent affects which issues are raised.

5.1. *Everybody picks on Kekes*

John Kekes (2000; 1992; 1985)—a classic disgust advocate—is a bit persecuted in the literature. As I have noted previously, he suggests the following as examples of “morally disgusting” offenses: “slowly disemboweling a person, dismembering someone with a chainsaw, slaughtering babies and bathing in their blood, or being drowned in excrement” (Kekes, 1992: 30). Several philosophers dismiss these examples, not on the basis that they do not evoke disgust (they surely do!), but on the basis that the sort of disgust evoked is merely *physical* in nature, not morally connotative (see Bollard, 2022: 10-11; Fileva, 2021: 23; Giubilini, 2016: 235-236; Kelly, 2011: 163).

Mara Bollard (2022) spends a bit more time explaining why she rejects Kekes' examples than other ontological skeptics, so I will focus on her arguments. Bollard explicates Kekes' view as follows. Moral disgust is a reasonable response to grievous and unjustified harm inflicted on a human being that outrages the sensibilities of morally committed witnesses. Kekes' view, according to Bollard, has two elements: (1) “the harm,” and (2) “the outrageous (gross, flamboyant, etc.) manner in which the action is performed” (p. 10). She then reviews the

examples from Kekes mentioned above and suggests that they evoke non-moral disgust. She writes, “it’s not clear from these examples that what Kekes is calling moral disgust is being elicited by the distinctively moral properties of the wrongful actions rather than by the sheer volume of generically disgusting stimuli” (p. 11). She further explains,

all of these examples are extremely physically disgusting. A large volume of blood, guts, gore, dead bodies, and feces are all clearly within the purview of generic disgust. The examples also include the morally relevant property of harm, and they do seem to be performed in an outrageous fashion. However, it seems that what makes each violation outrageous, gross, flagrant, and so on just is what’s physically disgusting about it; the generous amount of blood, guts, or feces involved, say. (p. 10)

To be clear, Bollard believes Kekes’ examples are physically disgusting while also including “the morally relevant property of harm” (p. 9). The issue is that the disgust and the morally relevant harm are unrelated elements. What makes each example “disgusting” is various non-moral features (e.g. excrement), while what makes them *wrong* is the presence of harm.

After making this general point, Bollard homes in on Kekes’ examples of torture. She mentions that he compares two cases of torture: (1) torturing captured enemy soldiers by depriving them of sleep and food, and (2) mutilating someone by dismembering them with a chainsaw. Kekes considers the latter viscerally disgusting (the qualia of our reaction is comparable to the qualia of our reaction to feces) but not the former. At this juncture, Bollard concedes two points to Kekes. First, torture by dismemberment is viscerally disgusting (it makes us want to retch), while torture by deprivation is not. Second, *if* our disgust in response to torture by dismemberment is directed at the cruel and callous attitude of the perpetrator, then this does seem like a “genuinely moral way of spelling out” Kekes’ assertion that moral disgust is a reaction to harm inflicted in an “outrageous” (or flagrant, or flamboyant) manner (p. 11). The attitude of the perpetrator is “particularly objectionable” (p. 11).

The problem with Kekes' argument, says Bollard, is that he does not adequately explain why torture by deprivation does *not* elicit moral disgust. This form of torture "also expresses extreme callousness and cruelty" on the part of the perpetrator (p. 11). Yet, we are not disgusted by it. This raises doubt about Kekes' claim that our disgust in the first example (torture by dismemberment) is really directed at the perpetrator's *attitude*. Both cases of torture (dismemberment and deprivation) have this property, but only one elicits disgust.

Why, then, are we disgusted by torture by dismemberment? In Bollard's view, this is "best explained by the presence of (non-moral) generically disgusting elicitors," such as blood and severed limbs (p. 11). She clarifies, the "compresence of moral violations and generic disgust does not suffice to make that disgust genuinely moral" (p. 11).

Stated formally, Bollard's argument against Kekes goes like this:

- (1) If the disgust we feel toward torture by dismemberment is directed at the perpetrator's objectionable attitude (e.g. extreme cruelty and callousness), then any other act exhibiting the same objectionable attitude should also elicit a similar disgust response.
- (2) Torture by deprivation (e.g. depriving captured enemy soldiers of food and sleep) expresses the same morally objectionable attitude of extreme cruelty and callousness as torture by dismemberment.
- (3) Torture by deprivation does not elicit the same disgust response as torture by dismemberment.
- (4) Therefore, the disgust we feel towards torture by dismemberment is not directed at the morally objectionable attitude of the perpetrator but, rather, is best explained by the presence of generically disgusting elements (e.g. blood, severed limbs).

5.2. *Is torture disgusting?*

I agree with Bollard that some of Kekes examples are a bit suspect.⁴⁰ But overall, her analysis of Kekes' torture examples is incorrect. The problematic premise in her argument is

⁴⁰ E.g. "being drowned in excrement" strikes me as (1) morally wrong (killing someone by any means is morally wrong) and also (2) extremely gross (the thought of touching excrement in general is nightmare fuel). In contrast,

- (2) Torture by deprivation expresses the same morally objectionable attitude of extreme cruelty and callousness as torture by dismemberment.

There are two reasons why I reject this premise. First, Bollard does not expound Kekes' position fully accurately. On Kekes' view, what makes an act disgusting is not so much the aesthetic features of the act themselves but, rather, what those features *say* about the perpetrator's attitude. She leaves out that, in Kekes' original paper, torture by deprivation is imagined to be done "in order to extract information from [the captured soldiers]," while torture by dismemberment is imagined to be done "to establish [the perpetrator's] reputation as an enforcer" (p. 32). These motives are very different. The latter motive is evil; the former is still bad, but not evil.

Second, the relationship between the aesthetic features of the act and the motive are *not* contingent but rather intimately related. The interrogator's method of torture reflects their goals (interrogation), and the enforcer's method reflects theirs (sending a message). In other words, there's a counterfactual relationship between the goal and the torture-method: if the enforcer's goal were different, he would not have chosen that particular method of torture. He chose dismemberment because of what that act *says*, and the message the act communicates, says Kekes, will be *universally understood by all* (1992: 438) (more on this momentarily).

I explain these objections more fully below.

5.3. *Kekes' account of moral disgust*

To start, let me briefly rectify Bollard's exposition of Kekes. (Since I've already given a fuller exposition of Kekes in Chapter 1, Section 3.2, I will be very brief.) Kekes is a liberal pluralist. He believes there are many possible forms of the good life, and the state's role in a diverse liberal democracy is to maintain the minimum conditions necessary for citizens to pursue

drowning someone in clean water (1*) wrong and (2*) not gross. What's the difference-maker between these two cases? Clearly, it's the presence or absence of excrement. Is that a morally relevant difference? No.

their preferred form and to generally promote tolerance. However, some lifestyles threaten these minimum conditions—these can't be tolerated and warrant disgust.

Kekes distinguishes between “primary conventions”—which organize and dispense the basic requirements that all good lives need, like safety, health, and nutrition—and “secondary conventions”—which organize and dispense secondary, culturally variable goods that enrich good lives, like religious faith and specific family customs (p. 24). Citizens' conceptions of secondary conventions are bound to conflict in diverse societies, so tolerance is demanded. But primary conventions are invariable; everyone's secondary conventions depend on them. Thus, breaches against them are intolerable and must elicit outrage. It is essential, says Kekes, that citizens not only abstractly understand that civilized life is underpinned by a set of minimum, necessary conditions, but also that they are deeply committed to their maintenance.

However, not all breaches to primary conventions are of equal seriousness. Some not only breach primary conventions but do so *flamboyantly*. Kekes calls these “taboos.” They are “the worst evil that people can do” (p. 432) and the “secular equivalent of sacrilege” (p. 443). While all immoralities are wrong, taboos are of a different order; the perpetrator's inner attitude and motive is so beyond the pale as to be incomprehensible to morally committed persons. For example, a person who is not debased could still agree that murder driven by revenge, albeit condemnable, is within the realm of the comprehensible. In contrast, for something like torture by dismemberment, “no circumstances could make that sort of breach understandable” (p. 443). The perpetrator “shows such a blatant disregard of what we regard as civilized life that anyone who so acts is placed beyond the pale” (443). His breach conveys *disdain* for primary conventions and, by extension, civil order itself. If safeguarding primary conventions requires

that citizens be outraged by breaches against them in proportion as they are flagrant, then taboo-violations must evoke *disgust*—a zero-tolerance emotion.

With Kekes' view now on the table, consider the passage from Kekes (2000) from which Bollard draws her examples of torture:

Murder, torture, and enslavement are certainly violations of required conventions. But it is one thing...to deprive captured enemy soldiers of sleep and food in order to extract information from them and another to mutilate randomly selected victims by cutting off their limbs inch by inch with a chain saw to establish one's reputation as an enforcer... All of these acts violate required conventions, and are therefore evil. But the second act...provokes deep disgust, whereas the first does not... In the first, wicked people use evil means to achieve evil ends. In the second, the evil of the ends is so extreme as to remove it from the range of understandable human options. These ends are monstrous and their pursuit is incompatible with civilized life (p. 32).

The crucial bit that Bollard leaves out in her exposition is the precise attitude of the perpetrator of torture by deprivation verses that of the perpetrator of torture by dismemberment. The former desires to extract information—a “wicked” end, but something an imaginative person can grasp. In contrast, the latter desires to “establish one's reputation as an enforcer.” This attitude makes the perpetrator worthy of disgust; their actions not only breach primary conventions but express disdain for them.

Therefore, Bollard's premise that both forms of torture express the same morally objectionable attitude is false. The attitudes differ significantly. Our disgust toward torture by dismemberment is directed not at its aesthetic features *per se* (blood, severed limbs), but what they convey about the perpetrator's attitude toward the minimum conditions underpinning civilized life: he disdains them. The message his act sends—universally and instinctively understood by all (1992: 438)—is revolting.

5.4. *Integrating Kekes and Rozin*

So, Kekes believes torture by dismemberment conveys an urgently dangerous, and universally understood, message. But (i) what is this “message,” (ii) why does torture by dismemberment convey it, and (iii) what justifies the assertion that it is universally understood? Here, Kekes is less than satisfying. We’re left only with these assertions unaccompanied by further explanation. Perhaps Kekes thought his meaning was obvious, but given ontological skeptics’ reactions, this clearly isn’t the case.

Fortunately, I believe this gap in Kekes’ account can be filled in by bringing in psychological research on disgust by authors like Paul Rozin and Jonathan Haidt. These authors, as I’ll momentarily explain, claim that all disgust, moral as well as non-moral varieties, is *meaning-laden*, as a result of cultural elaboration. Once a reflexive response to biological contaminants, disgust is now deeply intertwined with cultural meanings and symbolisms (which is perhaps the most human thing about it). For example, as discussed in Chapter 2, Section 4.2, much disgust responds to “animal-reminders”: objects and behaviors evocative of the animality, mortality, and vulnerability of human beings—features of ourselves with which we’re uncomfortable. Bloody car wrecks, open heart surgeries, medical amputations, and depictions of extreme violence and gore universally evoke disgust, each for the same reason: they violate the envelope of the body, placing the animality and vulnerability of persons on full display. Beneath a clean, smooth layer of skin lies a brittle skeleton, a web of stringy veins and tendons, and an ensemble of wet, bloody, squishy organs. These features we share with animals. However, unlike animals, we are conscious that they make us deeply vulnerable to danger and, eventually, death.

Rozin emphasizes in his research that body-envelope violations elicit disgust in all cultures, and its symbolism is instinctively recognized. If you ask people in an international airport what sorts of things disgust them, you can bet they will mention reminders of animality

and death: seeing a corpse, touching cremated ashes, bestiality, smelly feet, bad breath, gore, seeing people with amputated limbs or deformities. Research suggests that individual disgust sensitivity and death anxiety are positively correlated (Rozin et al., 2008). Consequently, we psychologically reject, avoid, and turn away from reminders of it. On a cultural scale, humans establish customs and norms that prop up the human-animal boundary and conceal, as far as possible, clues and suggestions of our animality (Rozin and Fallon, 1987; Rozin, Haidt, and McCauley, 2008). Although we must eat, poop, and have sex, the proper way to do all these things is culturally prescribed, and anyone who ignores these prescriptions is subject to social sanction (Rozin, Haidt, and McCauley, 2008: 761).

All this, I believe, bolsters Kekes' claim that torture involving the deliberate dismemberment of a person conveys an urgently dangerous message: the mortality and vulnerability of the victim is put on full display, as if to humiliate them. It's not just the aesthetic features of the torture (blood, severed limbs) that are morally significant and disgusting. The disgust goes deeper than just "Eww, blood." It's the aesthetic features' intimate link to the perpetrator's attitude that disgusts. Despite the deep hesitations that cultures around the world have about violating the body envelope, the torturer who dismembers his victims flamboyantly violates their body-envelope, as if to say "Behold in horror your vulnerability and animality, mere mortals! I care not for your vulnerability; indeed I relish in humiliating you and making you wallow in it!" This is an evil message by which we are justifiably repulsed; it expresses an attitude diametrically opposed to the minimum conditions upholding civil society. In contrast, torturing captured enemy soldiers by depriving them of food and sleep in order to extract information from them does not convey this same message or attitude. Therefore, the two types of torture are expressive of different attitudes, *pace* Bollard.

5.5. *There is no “run-of-the-mill” disgust*

The above story, I think, is compelling. Admittedly, Kekes himself doesn't explicitly connect his views to Rozin's research, despite the fact that they gel. This wasn't an oversight on his part; empirical disgust research took off in the late 1990s, after the bulk of Kekes' views on moral disgust and taboos were already published. Nevertheless, his views resonate with Rozin's research profoundly. Rozin himself at times references Kekes' work favorably (see Rozin et al., 1999: 436), and Kekes has, at least once, reciprocated this scholarly high-five (see Kekes, 2005: 440).

Why don't Bollard, Fileva, Gert, or Giubilini consider this story before dismissing Kekes' account of moral disgust as flawed? The answer is seemingly that they do not accept Rozin's (influential) account of disgust. The most obvious sign of this is their repeated use of phrases like “physical disgust,” “generic disgust,” “pathogen disgust,” and “run-of-the-mill” disgust. Rozin never uses these; he always refers to disgust that is concerned with biological contaminants as “core disgust,” invoking the idea that the emotion has a core biological function, which was later culturally elaborated to police new domains.

I'm not quibbling over minor preferences in word choice. From a Rozinian perspective, it introduces serious confusion to speak of “physical disgust.” Much disgust has a physical component, even *moral* disgust. For example, for Rozin, moral disgust is frequently elicited by purity violations, which typically have bodily or sexual elements. To conflate core and animal-reminder disgust under the label “physical disgust” is to minimize the cultural meanings and symbolisms intertwined into many disgust triggers. Gert (2015) states in a footnote that a different team of psychologists, Tybur et al. (2009) present “powerful reasons” to reject Rozin's concept of “animal-reminder disgust”—without relaying what those reasons are or why they're

“powerful” (see Gert, 2015: 36). Even some new disgust advocates—whose conceptions of moral disgust, as I’ve noted, are self-consciously designed to focus on anti-social violations *rather than* purity violations—explicitly reject Rozin’s account of disgust (see for example Plakias, 2013: 270-272).

In my view, it’s not unreasonable to reject Rozin’s account of disgust. It remains influential, but there are alternatives, and many empirical researchers, such as evolutionary psychologist Joshua Tybur, challenge it (Tybur et al., 2013; Tybur et al., 2009; Tybur, 2009). Nevertheless, I think Rozin has adequately defended his account against criticisms (see for example Rozin and Haidt, 2013). So, ontological skeptics should not *dismiss* the Rozinian paradigm. It’s worth serious consideration, and if we take it seriously, we see that it aligns with Kekes’ account of moral disgust.

Unfortunately, Bollard and Giubilini at times conflate core and animal-reminder disgust in way that prejudices their arguments against Rozin and Kekes. For example, Giubilini (2016) writes:

Originally a physiological reaction protecting humans from pathogens and infections, in the course of evolution disgust became responsive to reminders of our animal nature, such as some of the ‘envelope violations’ of the human body that are likely to be carriers of infections—for example, excrement or open wounds (p. 229).

He then cites one source: Haidt et al. (1997). Unfortunately, Giubilini misunderstands the source he cites. Excrement, for Haidt, is not a body envelope violation. Haidt et al. (1997) define envelope violations as “a forcible breach or alteration of the exterior envelope of the human body” (p. 112). Their examples include “bloody car accidents, mutilated corpses, surgery, wounds, and physical deformity” (p. 112). Meanwhile, core disgust objects (disease vectors) include “foods, animals, and body products” (p. 112). Excrement is a body product (i.e. a core disgust trigger).

Moreover, for Haidt, envelope violations are *not* disgusting in virtue of being “likely carriers of infections,” *pace* Giubilini. Haidt et al. (1997) write,

The massive restrictions that Americans place on eating, sexuality and body modification, and the linkage of all three to disgust, point to a concern about the human body that cannot be based on rational fears about health. ... Rather, Americans seem at times to hold a view of the body observed in other parts of the world: that the body is a temple, housing the self or the soul within. ... Core disgust guards against material contamination.

Meanwhile, “the extended animal reminder concerns about sex, envelope violations and hygiene guard against any undignified use or modification of the temple” (p. 114).

Making a similar mistake as Giubilini, Bollard (2022) at one point writes the following:

Given disgust’s evolutionary function to protect us from contamination, the paradigm elicitors of disgust are objects that present possible threats of disease or infection, such as decaying or rotten foods, body products (e.g. blood, feces, vomit, viscera), body envelope violations, and certain sexual practices (e.g. incest, bestiality) (p. 4).

Thus, Bollard classes viscera, body envelope violations, incest, and bestiality as core disgust triggers (they “present possible threats of disease or infection”). Of course, the Rozinian paradigm rejects this completely. Haidt et al. (1997) explicitly write,

our analysis of core disgust cannot explain two other kinds of events. First, there are frequent references to sexual matters, such as incest, homosexuality, bestiality, or almost any other deviation from the cultural ideal of “normal” heterosexuality. Second, there are frequent references to bloody car accidents, mutilated corpses, surgery, wounds, and physical deformity. This last set of examples all involve a forcible breach or alteration of the exterior envelope of the human body (p. 112).

Haidt and his colleagues then suggest that these spectacles “remind of us our animal nature” and every culture “prescribes the proper *human* way to handle these biological functions, and people who violate these prescriptions are typically reviled or shunned” (p. 112).

In summary, it is likely that many ontological skeptics dismiss Kekes in part because they do not accept Rozin’s descriptive account of disgust—a descriptive account which resonates with

many of Kekes' *normative* assertions. If, however, we take Rozin's descriptive account of disgust seriously—which I believe we should—then Kekes' normative views become more plausible.

5.6. Conclusion

This section sought to defend classic disgust advocacy against the Coincidence Challenge. More precisely, I defended Kekes' examples of morally disgusting offenses (e.g. torture by dismemberment) against common critiques from ontological skeptics, like Bollard, who claim that the disgust they elicit is merely physical, not directed at moral properties. I integrate empirical research from Paul Rozin and Jonathan Haidt into Kekes' account of moral disgust in an attempt to vindicate Kekes' claim that torture by dismemberment conveys an urgently dangerous message and expresses disdain for primary conventions on the part of perpetrator. This message and the attitude of the messenger justifiably evoke a form of disgust that is genuinely moral in nature.

6. Classic Advocacy and the Coincidence Challenge: Harmless Purity Violations

The aim of the previous section was to defend Kekes' account of moral disgust against the Coincidence Challenge. Kekes, in expositing his account, proposed several examples of “morally disgusting offenses” to use as illustrations throughout his argument. Something that these examples have in common is that they “include the morally relevant property of harm” (Bollard, 2022: 10). This might have made it easier to defend the view that they are instances of genuine moral disgust; they represent harms *performed in an outrageous manner*. But what about purity violations that do not include this property?

This section is about *harmless* purity violations and whether the disgust they elicit is genuinely moral. Bioethicist Leon Kass' infamous 1997 paper "The Wisdom of Repugnance" gives a list of "horrors," which he believes are morally disgusting. The list includes several harmless purity violations, such as cannibalism, incest, bestiality, and corpse-mutilation. While these offenses *can* involve harm, this is not always or necessarily the case. Incest can occur between consenting adult siblings for whom there is no chance of pregnancy; cannibalism can be harmless if one obtains permission to consume someone's remains after they die of natural causes; and so on. Kass expressly resists linking the wrongness of these "horrors" to their potential to cause harm. For example, he claims that someone who links the wrongness of incest to the risk of inbreeding is engaging in a disingenuous attempt to "rationalize away our horror" (p. 687). Kass is not alone in this view. As discussed in Chapter 1, empirical research has found that many ordinary people continue to express disgust toward, and disapproval of, taboo-breaking behaviors like incest, cannibalism, necrophilia, bestiality, and the commoditization of human organs even when they are convinced there is no chance of harm, injustice, or violations of rights or consent (see Haidt, 2001; Haidt et al., 2000; Haidt and Hersh, 2000; Haidt et al., 1993).

Arch-ontological skeptic Mara Bollard suggests that harmless purity violations, like bestiality and incest, do not elicit *moral* disgust; rather, they elicit physical disgust, which is accompanied by an independent moral judgment that they are wrong. Bollard is not alone. Other ontological skeptics also suggest that harmless purity violations are disgusting in a physical sense, not a moral sense, and the connection between the feelings is coincidental. Alberto Giubilini (2016) explicitly claims that the disgust that some people feel towards sexual practices like homosexuality, bestiality, and incest is merely *moralizing* (p. 227, 233). That is, these

practices are physically disgusting, and their physical disgustingness *induces* many people to judge them as wrong. In such cases, moral disapproval is indefensible. Physical disgustingness is not a wrong- or bad-making property. Joshua Gert (2015) finds it “amazing” that the moral psychology literature so frequently uses bestiality and incest as “paradigm instances of immoral behavior,” and he laments that this grants a “false respectability” to homophobia (p. 54). Finally, Iskra Fileva (2021) writes that “taboo violations such as sibling incest... are not genuine cases of moral disgust” (p. 30). Instead, disgust toward incest is in principle separable from a moral judgment; extraneous disgust may even cause a moral judgment (i.e. physical disgust towards incest may moralize incest) (pp. 25-6).

In summary, ontological skeptics tend to deny that harmless purity violations are moral violations at all, which undermines the claim that we are morally disgusted by them (one can’t be morally disgusted by merely physical stimuli). This raises the question, why exactly do ontological skeptics believe that harmless purity violations are not genuine moral violations? To be clear, ontological skeptics do not simply argue that the proposition “Incest is immoral” is false (or invalid, or unjustified). Rather, they argue that anyone who is disgusted by incest is *not experiencing a moral sentiment*, or if they believe that they are, they are simply mistaken—their physical disgust is causing them to moralize. Why do skeptics believe this?

6.1. *Bollard’s moral/conventional distinction*

I will focus on Bollard, since she provides the most thorough explanation. In her view, “how to delineate the boundaries of the moral domain” is a “difficult problem” about which philosophers continue to puzzle (p. 5). However, this doesn’t mean “we can’t make *any* distinctions between instances of genuinely moral disgust and non-moral generic disgust” (p. 5). She continues, “moral disgust must be demonstrably differentiable from generic disgust” in a

couple of ways (p. 5). First, for an instance of disgust to be moral, it must be directed at moral stimuli—“violations with certain morally relevant features” (p. 5). Second, it must be directed at a *moral* violation, not a violation of convention. Violations of convention do not count as moral stimuli.

Next, Bollard claims to follow psychologist Elliot Turiel in how she understands the moral/conventional distinction. She writes that moral violations are “defined by their consequences for the rights and welfare of others and typically involve harm being done to a victim,” such as “hitting someone or pulling their hair” (p. 7). In contrast, conventional violations are defined as

violations of the behavioral norms that operate within social systems (e.g. chewing gum at school). Unlike moral transgressions, the wrongness of conventional violations is considered authority-dependent (i.e. it is only bad to chew gum at school because the teacher says so) (p. 7).

In summary, Bollards claims that moral violations differ from conventional violations insofar as they are (1) authority-independent, (2) more serious than conventional violations, and (3) typically associated with harm, injustice, or consent- or rights-violations.

Bollard then applies this to the case of harmless taboo violations, such as incest and bestiality. First, she makes the following concession:

violations [of sexual morality] are perceived (by some) as wrong, tend not to involve harm, and, importantly, elicit disgust. If these actions are perceived as moral rather than merely conventional violations, then we may have the makings for a case in favor of the existence of moral disgust, insofar as the disgust that they elicit could be moral in nature (p. 8).

In short, if incest and bestiality are perceived as moral rather than conventional violations, and if they elicit disgust, then this lends credence to the claim that they elicit moral disgust—but only if the disgust is directed at a morally relevant feature of the act. Of course, Bollard denies that the

disgust is really directed at morally relevant features. Instead, a “more plausible explanation” is that “people are experiencing generic disgust” towards them, which “accompanies the agent’s independent judgment that these sexual acts are wrong” (p. 9).

To support this claim, Bollard proposes that we compare incest and bestiality with another violation of sexual morality: adultery. A comparison of these reveals that incest and bestiality elicit disgust “far more often” than adultery (p. 9). And yet, bestiality and incest are “far less amenable to a harm-based explanation for [their] (putative) wrongness” than adultery (p. 9). In short, (1) adultery does not widely elicit disgust, and (2) adultery is amenable to a harm-based explanation. Combined, these points suggest that there is not a “unified class of sexual moral violations that elicit *moral* disgust in virtue of some wrong-making moral property” (p. 9). After all, adultery is a stronger candidate for an instance of “sexual immorality” than incest or bestiality, insofar as it involves harm—yet, unlike incest and bestiality, it does not widely elicit disgust. Bollard continues, “the violations of sexual morality that most often bring about disgust are the ones with particularly salient physically disgusting (to some) features” (p. 9). Therefore, the “more plausible explanation” is that people are experiencing “generic disgust” and “this disgust accompanies” their moral judgment that incest, and bestiality are wrong (p. 9).

6.2. *Two obvious problems*

To start, let me note a couple of obvious, surface-level problems with Bollard’s argument, before moving on to deeper issues.

First, ample empirical evidence shows that many people (1) are disgusted by incest and bestiality and also (2) morally disapprove of them (see for example Haidt and Hersh, 2001; Haidt et al., 2000). Bollard accepts both of these points. Her contention is that (1) and (2) are not connected—people’s disgust is directed at physical features of incest and bestiality and happens

to accompany an “independent judgment” that they are wrong. But what “independent judgment” is Bollard referring to? She has already admitted that they “tend not to involve harm” (p. 8). Haidt’s research on moral dumbfounding has shown that if you tell people to explain why they morally disapprove of incest and bestiality, they will often say “They are disgusting.” If you ask them to elaborate, they become dumbfounded—unable to explain any further. To me, this suggests that many people have no “independent judgment” that incest and bestiality are wrong, aside from the feeling that they are deeply disgusting.

Second, Bollard suggests that we compare adultery to incest and bestiality. If we do this, we find that adultery is a stronger candidate for a moral violation (it is “far more amenable” to a harm-based explanation), yet it does not elicit disgust. She then interprets this as evidence that many people’s disgust toward incest and bestiality is not morally connotative. However, this comparison doesn’t make much sense. Obviously, adultery *can* evoke notions of purity, if the adulterer is a sleazy lecher who lacks control over their sexual urges, and this causes them to succumb to the temptation to cheat. But if we just focus on the act itself and ask, “Why is it wrong?”, the answer that readily jumps to mind is that *adultery involves a reciprocity violation*. Someone who is married has sex with someone who is not their spouse, presumably without their spouse’s knowledge. There is nothing intrinsically offensive about the act of two people having sex. What makes adultery wrong is the backdrop against which it takes place, namely, that the adulterer is breaking their marriage vows and betraying their spouse’s trust. The wrong-maker is the betrayal. If the adulterer would like to have sex with someone other than their spouse, they can remove this wrong-maker by getting a divorce (formally letting their spouse know that they are no longer committed to sexual monogamy, nullifying the expectation of trust). In contrast with adultery, incest and bestiality tend to be viewed as *intrinsically offensive*. What makes them

wrong (assuming they are) has nothing to do with the harm, injustice, or violations of rights or consent that they pose (indeed, they need not pose any of these). Rather, they are wrong in and of themselves.

6.3. *Turiel's moral/conventional distinction*

A deeper problem with Bollard's argument lies in how she spells out the moral/conventional distinction. She claims to follow Turiel's (1983) influential account but omits crucial components of it. Bollard mentions only three criteria that Turiel identifies for a moral rule: (1) authority-independence, (2) being worse than conventional violations, and (3) being justified in terms of harm, rights, or injustice. However, there are two additional criteria from Turiel's (1983) account that Bollard doesn't mention: (1) universalizability and (2) categoricalness (for a lucid exposition of Turiel's project, see Machery and Stich, 2022). The universalizability criterion holds that to be moral, a judgment must be generalized to other countries or cultures; in contrast, conventional violations are localized. The categoricalness criterion holds that a judgment, to be moral, must describe an obligation that an agent has that cannot be evaded by reference to the agent's personal desires. For example, the common moral judgment "Murder is wrong" is, Turiel's research finds, typically viewed by speakers as universalizable and categorical (people extend this judgment to other cultures, and they also view the judgment as describing an obligation bearing on oneself that is not based in one's goals or self-interest). In short, Turiel recognizes five criteria for moral, as opposed to conventional, judgments: they are (1) universalizable, (2) categorical, (3) more serious than conventional, (4) authority-independent, and (5) justified in terms of harm/rights/injustice.

With these five criteria on hand, we can ask the following. Should widespread condemnation of purity violations—take incest as an example—be interpreted as (i) a

widespread *moral* sentiment or (ii) merely a negative reaction to the violation of a convention?

To operationalize this, we need only examine whether people's judgment regarding incest satisfy Turiel's five criteria.

Bollard doesn't mention this, but Turiel himself examined this question, a few years after his 1983 book, in Turiel et al. (1991). This study examined the judgments and reasoning of American high school and college students regarding "prototypical transgressions" (e.g. killing and rape) and "non-prototypical transgressions" (e.g. abortion, pornography, homosexuality, and incest). The researchers found that "Of all the nonprototypical issues we investigated, incest was most closely assimilated to reasoning in the moral domain" (1991: 78). A substantial number of subjects judged that incest was "not all right," and their judgments were (1) universalized, (2) categorical, (3) authority-independent, and (4) more serious than their judgments regarding conventional violations. The only criterion that these judgments did not meet was (5): they tended *not* to be justified in terms of harm, rights, or injustice. Instead, subjects tended to base their reasoning on what Turiel and his colleagues called "deterministic beliefs," which they define as "standards dictated by psychological normality, biological order or religious order" or "custom and tradition" (p. 24, 78).

So, many subjects' judgments of incest follow the typical pattern of moral judgments, sharing a strong family resemblance with judgments like "Murder is wrong" and "Rape is wrong," only exhibiting one difference: incest is not seen as harmful. This raises the question, should we treat subjects' condemnation of incest as morally connotative or not?

This question, of course, turns on whether we accept criterion (5) as a necessary condition for a judgment's being moral. Right off the bat, it's worth noting that criterion (5) is not like the other four. Criteria (1-4) are *formal* in nature. For example, universalizability and

categoricalness don't impose specific requirements for what *content* a judgment must include to count as moral. They are only concerned with the formal or structural features that moral judgments typically have. In contrast, criterion (5) requires that a judgment, to be moral, must be justified in terms of a specific list of reasons. This is a *substantive* criterion, as opposed to a formal one (Machery and Stich, 2022).

We can now rephrase our question as follows. In defining “moral” for the purposes of psychological research, should we recognize any substantive criteria—such as “For a judgment to count as ‘moral,’ it must be justified in terms of harm, injustice, or rights”—or should we not?

6.4. *Problems with Turiel's substantive criterion*

I'm highly dubious about the particular substantive criterion proposed by Turiel. The most obvious problem with it is that it is vulnerable to counterexamples.⁴¹ For example, it seems intuitive that breaking promises one has made to those who are now dead is wrong (morally, not conventionally). For example, suppose I promise my grandma on her deathbed that I will personally take care of her Cocker Spaniel, Mitch, and not send him to the pound or give him to another family. Then, after my grandma passes, I decide that taking care of Mitch is too much of a hassle and give him to another family. Have I done something morally wrong? Yes. Have I harmed anyone? No. I would add that this isn't an isolated case. It seems to me, Turiel's substantive criterion is in tension with the common intuition that so-called “thick” moral concepts like “dignity,” “courage,” “humility,” “kindness,” and “loyalty” are genuinely *moral* concepts. How can we explain what's morally good about these things solely in the language of

⁴¹ Bollard herself makes this exact point, but for some reason she doesn't incorporate it into her analysis of incest and bestiality. She admits that many people judge incest and bestiality to be wrong, and she also notes that these acts “tend not to involve harm.” She then suggests that the fact that they don't cause harm is a reason to doubt that people's negative judgments about them are moral in nature. She doesn't consider the alternative possibility that people's judgment of incest and bestiality are genuinely moral in nature *despite* not being based in considerations of harm.

harm, justice, and rights? Of course, if someone just wants to stipulate that what they mean by the term “moral judgment” is “a judgment that is justified in terms of harm, rights, or justice,” that’s totally fine. But you can’t turn this stipulation against disgust advocates and say, “Negative judgments of incest are not *moral* judgments, properly speaking.” That would be question-begging.

As far as I can tell, in general, it’s more common for moral philosophers and metaethicists to define “moral” in formal, not substantive, terms. For example, Michael Huemer (2013) is a prominent moral realist, meaning he believes there are at least some “objective moral reasons.” In other words, there are at least some objective facts that make true sentences of the form ‘S has a moral reason to Φ ’ (p. 261). Before giving any arguments for this view, Huemer defines a “moral reason” as a reason for (or against) acting a certain way that is (1) *categorical* (not based on the agent’s desires), (2) *non-selfish* (not based on the agent’s interests), and (3) *objective* (independent of the attitudes of observers towards the action). Note that these are formal constraints. They do not impose specific requirements on what content a “moral reason” must include. Obviously, Huemer intends this definition to be uncontroversial. His argument for moral realism would be off to a bad start if his anti-realist opponents objected to his definition of the term “moral reason.”

Philosophical complaints aside, Turiel’s substantive criterion has also been criticized by empirical researchers, including Richard Shweder and Jonathan Haidt. The crux of Haidt’s criticism is that the predictive relationship between criterion (5) and the other four criteria is unidirectional. That is, judgments that are justified in terms of harm, justice, or rights tend to also exhibit features (1-4) (i.e. 5 reliably predicts 1-4). However, judgments that exhibit features (1-4) are not always justified in terms of harm, justice, or rights (i.e. 1-4 do not reliably predict 5). In

his research, Haidt has found that many people make moral judgments in response to “harmless taboo violations.” Here are some examples that Haidt has studied:

- Eating one’s pet dog.
- A brother and sister who enjoy passionately kissing on the mouth.
- Having sex with a dead chicken.
- Masturbating while being licked by a dog.
- A medical researcher taking home and eating leftover parts of cadavers that would otherwise be discarded.
- Cleaning one’s toilet with the national flag.

Haidt has found that many morally condemn these acts. He considers condemnation “moral” in nature, not merely conventional, if the subject endorses the two following beliefs:

First, people should not be at liberty to perform these acts; they should be stopped or punished. Second, the wrongness of these acts is universal, not contingent on local custom or convention (Haidt et al., 1993: 613).

In Haidt’s view, these are “the hallmarks of a moral judgment” (p. 614).

Interestingly, the rate at which people make moral judgments about harmless taboo violations differs cross-culturally. For example, as noted in Chapter 2, Section 7, Haidt has found that American conservatives moralize about harmless taboos more than liberals (Haidt and Graham, 2007; Haidt and Joseph, 2004; Haidt and Hersh, 2001); Brazilians more than Americans (Haidt et al., 1993); and people of low socioeconomic status (SES) more than high-SES (Haidt et al., 1993). Findings such as these ultimately led Haidt to develop his Moral Foundations Theory, which proposes that people are born with innate sensitivities (tastebuds) to various moral domains (such as harm/care, loyalty/reciprocity, and purity/sanctity), but the culture in which one is raised shapes the development of these capacities. This, for Haidt, is ultimately what explains cross- and intra-cultural differences in moral intuitions between liberals versus conservatives, Westerners versus non-Westerners, and low-SES individuals versus high-SES individuals. For example, Haidt has claimed that liberals tend to only be sensitive to the

harm/care and justice/fairness domains, while conservatives are sensitive to a broader set of domains, which includes authority/respect, ingroup/loyalty—and purity/sanctity. Likewise, Haidt finds that Westerners in general tend to emphasize the harm/care and justice/fairness foundations over other domains (e.g. purity/sanctity), whereas non-Westerners recognize a broader set of domains. In response to findings like this, Haidt has stated that Turiel’s substantive criterion “doesn’t travel well” (Haidt, 2012: 22).

In summary, Turiel’s substantive condition has empirical critics. Bollard does briefly mention Haidt’s view that “purity norms *are* moral norms, and not merely conventional norms” (p. 7). However, she quickly dismisses this view, writing “for present purposes, I take myself to be justified in setting wide/descriptive accounts of morality aside.” Why? She seems to give two reasons, neither of which, in my view, *do* justify setting Haidt aside. First, she interprets Haidt’s account to be that “most (if not all) disgust just *is* moral disgust” (p. 7). She continues, “Purity norms involve food and bodily objects, and the violations of purity norms typically elicits generic disgust” (p. 7). But this is simply not an accurate representation of Haidt’s view. He does not hold that most, or all, disgust just *is* moral disgust. On the contrary, Haidt defines “moral” judgments as those that meet a set of formal conditions (e.g. universalizability, severeness), which he considers the “hallmark of moral judgment.” Some negative judgments towards disgusting acts are universalized and severe (e.g. condemnation of incest, cannibalism, and bestiality), but—crucially—not all of them are (e.g. disgust at dipping vanilla wafers in ketchup). Also, Bollard’s claim that “purity norms typically elicit generic disgust” is incorrect (p. 7). To support this, she cites Haidt (2012) and Rozin et al. (1999). Yet, both of these sources recognize a difference between core disgust (for reasons explained in the previous section, Haidt and Rozin prefer the term “core disgust”) and *moral* disgust. People do not moralize about core disgust.

Bollard's second reason for dismissing Haidt is that psychologists, in circumscribing morality, are "not doing what moral philosophers are doing: we have something more explicitly normative in mind" (p. 7). However, isn't condemnation of purity violations like incest, bestiality, and cannibalism "explicitly normative"? People universalize their judgments about these acts. Moreover, Turiel is also a psychologist. In circumscribing "morality" the way he does, he (like Haidt) is attempting to *describe* people's moral judgments and intuitions, not lay out a normative ethical theory.

Bollard is not alone in setting Haidt aside. Other ontological skeptics, such as Alberto Giubilini, do too—and for equally unconvincing reasons (see also Gert, 2015: 52). Giubilini (2016) writes,

There is no consensus about which judgments and evaluations count as 'moral'. ... According to many people in Western societies, 'morality' is essentially about not harming others, and about respecting some standard of justice; and other types of norms are mostly seen as mere social convention. However, many non-Western cultures, and also many sub-cultures within Western societies (such as religious groups), consider as moral matters certain practices that are not related to harm or to justice, such as eating habits [Haidt 2012: 14-18]. As a consequence, talking of moral disgust in moral psychology is problematic because it is not clear which criteria allow us to say that a certain disgust reaction is connected with a moral evaluation (p. 231).

This claim just seems false. We *can* tell the difference between physical disgust judgments and moral disgust judgments. Namely, we can establish a plausible set of formal constraints on what counts as a moral judgment *simpliciter* (e.g. universalizability and severeness) and then empirically check to see whether people's disgust judgments satisfy these constraints. If they do, this is evidence that their disgust is moral in nature, not merely physical.

6.5. Conclusion

This section sought to defend classic advocacy against the Coincidence Challenge. Specifically, I examined whether disgust towards harmless purity violations, like consensual

incest or cannibalism, can be genuinely *moral* in nature. Apparent cases of moral disgust towards harmless purity violations is often dismissed by ontological skeptics, like Bollard, as merely *physical* disgust that happens to be accompanied by an independent moral judgment. I resisted this suggestion by drawing on empirical research showing that people's moral judgments about incest and other taboo behaviors are deeply tied to their disgust reactions and don't involve independent, harm-based reasoning. Additionally, I challenged Bollard's application of Turiel's moral/conventional distinction, noting that she overlooks certain *formal* features of moral judgments that Turiel proposed, like universalizability and categoricalness. If we impose formal rather than substantive criteria on what counts as a moral judgment, there is a strong case to be made that many people's negative judgments of purity violations are morally connotative, not merely conventional.

7. Conclusion

This chapter has argued that moral disgust is a genuine emotional state. That is, it is possible to identify emotional states that simultaneously constitute both (1) genuine cases of disgust and (2) genuine cases of a moral sentiment. More precisely, I have defended this view against two objections: the Metaphor Challenge and the Coincidence Challenge.

My argument began by distinguishing between two versions of disgust advocacy: classic and new. I suggested that classic and new advocates have different conceptions of moral disgust. Namely, classic advocates view the emotion as a response to purity violations, while new advocates view it as a response to anti-social violations. I then argued that, in light of these different conceptions, classic and new advocates are not equally vulnerable to the Metaphor and Coincidence Challenges. Rather, classic advocacy is vulnerable to the Coincidence Challenge

but not the Metaphor Challenge, while new advocacy is vulnerable to the Metaphor Challenge but not the Coincidence Challenge.

From there, I argued that neither classic nor new disgust advocacy ultimately succumbs to their corresponding challenges. New advocacy, which focuses on anti-social violations, does not succumb to the Metaphor Challenge because there are strong theoretical and empirical reasons to believe that people do sometimes respond to anti-social violations with disgust.

Next, I defended classic advocacy against the Coincidence Challenge. Specifically, I addressed the question of whether purity violations ought to be considered moral violations. This discussion was divided between purity violations that involve harm and those that don't. Regarding harmful purity violations, I argued that John Kekes has convincingly shown why these actions legitimately evoke disgust. Disgust is a legitimate reaction to actions that flamboyantly violate the basic conditions underpinning civilized life, and harmful purity violations—such as deliberate body-envelope violations—are rightly interpreted as such flamboyant violations.

Next, I considered whether *harmless* purity violations evoke a form of disgust that is potentially moral or merely physical. I argued that these violations evoke genuinely moral disgust on the grounds that, first, the most plausible descriptive definition of “moral judgment” imposes only formal, not substantive, constraints. Second, ample empirical evidence shows that many people's judgments of harmless purity violations satisfy these formal criteria. Moreover, people characteristically express genuine disgust toward these violations, and given they are harmless, it follows that their moral disapproval is based in disgust and disgust alone.

7.1. *The upshot*

Bollard herself notes that disgust advocates posit a “special, distinctively moral emotion” (2022: 6). This chapter has argued that this postulate is real. What is the strategic value of this conclusion for the larger project of the dissertation—providing a (qualified) defense of moral disgust’s role in criminal law?

One obvious takeaway is merely a defensive one. If the goal of the dissertation is to defend, at least partially, the (surprisingly large) role moral disgust currently plays in law in many countries, including the United States, then the looming threat of ontological skepticism thwarts this goal. If “moral disgust” is not a genuine emotional state, it can’t carry the normative significance that disgust advocates attribute to it. So, establishing that moral disgust *is* a genuine emotional state, distinguishable from both other states (e.g. moral anger) as well as physical disgust, heads off this particular worry. Disgust advocacy (the view the disgust carries normative significance) lives to fight another day.

Another takeaway is this. I have argued that ontological skeptics raise objections to Genuine that don’t hold up under scrutiny. Part of my argument was that skeptics, at times, advocate a form of skepticism that is simply excessive. They treat Genuine as if, prior to any investigation, it has a low probability of being true. As such, it can’t be established unless a body of unequivocal empirical evidence is provided. Of course, this pre-investigatory skepticism about Genuine, I argued, is not reasonable or defensible. Raising doubts about Genuine is one thing, but treating it as a far-fetched, tin-foil-hat-worthy hypothesis does not line up with what we know from work in moral or evolutionary psychology. Why, then, is there an air of pre-investigatory skepticism about Genuine in the literature?

Far be it from me to question the motives of ontological skeptics. The body of scholarship they have produced has so far been unqualifiedly excellent. Nonetheless, several

authors in the literature—such as Kumar (2017), Clark and Fessler (2015), Demetrious (2013), Kahan (1998), and Miller (1997)—have pointed out an undeniable pattern: it is *liberal* political philosophers who tend to be suspicious of moral disgust. The reasons for liberals’ suspicion are pretty obvious; in the course of the disgust advocate/skeptic debate, much experimental evidence has been marshaled by skeptics to show that disgust (allegedly) has a dark side. Martha Nussbaum (2004) suggests it is the primary emotion involved in racism, sexism, anti-Semitism, and xenophobia. Daniel Kelly (2011) suggests it is prone to false positives and that it may have direct evolutionary roots in policing ethnic boundaries (see also Kelly and Morar, 2014). Other authors note that, if we ever discover that we are disgusted by something that ought not disgust us, our feelings of disgust can be very difficult to uproot (Kurth, 2021). Some experimental evidence suggests that disgust is relatively insensitive to others’ intentions or agency or to mitigating factors (Russell and Giner-Sorolla, 2011a). Being in a state of physical disgust may also amplify the harshness of one’s *moral* judgments (e.g. if you smell a fart, you may judge a bank robber as deserving of a longer prison sentence) (Schnall et al., 2008; Wheatley and Haidt, 2005).

In short, it is often alleged that disgust is inherently unwise or problematic. In that case, *of course* we should be suspicious of it in normative discourse. To concede that moral disgust is a genuine emotional state would be to grant disgust-linked prejudices a “false respectability” (Gert, 2015: 54). Denying its existence, then, is a sort of noble lie.

The question, of course, is whether that lie really is noble. The next chapter takes up this issue. Is moral disgust actually a problematic, dangerous, or unwise emotion? Is that what the evidence suggests? Or, contrarily, might the evidence suggest moral disgust *is* wise—able to point us to moral truths we might otherwise miss?

Chapter Four

Is Repugnance Wise? Moral Dumbfounding, Gene-Culture Co-Evolution, and Disgust Theory

In crucial cases...repugnance is the emotional expression of deep wisdom, beyond reason's power to fully articulate it.

—Leon Kass, “The Wisdom of Repugnance”

The pressure for an acceptable, clear, and explicit reason for doing things is merely a social norm common in Western populations, which creates the illusion (among Westerners) that humans generally do things based on explicit causal models and clear reasons. They often do not.

—Joseph Henrich, *The Secret of Our Success*

1. Introduction

One of the earliest flashpoints in the disgust advocate/skeptic debate—possibly the central moment responsible for its initiation—was bioethicist Leon Kass’ 1997 paper “The Wisdom of Repugnance.” The paper was a response to a successful attempt by Scottish scientists that same year to clone Dolly the sheep via somatic cell nuclear transfer. Kass advocated for universal bans on human cloning, embryonic stem cell research, and other emergent biomedical and biotechnological endeavors that, he worried, threatened to radically alter essential features of human life. The crux of his argument is that these endeavors elicit feelings of “repugnance.” He further asserted that

In crucial cases...repugnance is the emotional expression of deep wisdom, beyond reason’s power to fully articulate it. Can anyone really give an argument fully adequate to the horror which is father-daughter incest (even with consent), or having sex with animals, or mutilating a corpse, or eating human flesh, or even just (just!) raping or murdering another human being? (p. 687).

In other words, in Kass’ view, some practices—from incest and cannibalism to biotechnologies like cloning—are clearly wrong. Almost everyone is repulsed by them. Yet, our

attempts to articulate what their wrongness consists in, in clear philosophical prose, will inevitably fail to be “fully adequate to the horror” that they represent. Of course, this failure of philosophy should in no way make us question or doubt the wrongness of these practices; on the contrary, we know they’re wrong because they make us shudder, say “eww,” and feel icky. In a word, Kass attributes a sort of *epistemic* significance to disgust; the emotion can signal the (serious) wrongness of actions or practices whose wrongness we might otherwise overlook.

While the disgust advocate/skeptic debate covers a range of issues—including whether disgust bears any moral or political significance—the question of whether it bears any epistemic significance is arguably where the debate began, and it continues to receive attention in the literature. A spectrum of views on disgust’s epistemic significance can be found. Some authors, similar to Kass (1997), believe that at least in certain cases, disgust holds at least some epistemic value (see for example Fischer, 2016; Midgley, 2000; Kahan, 2000). By contrast, others hold, for various reasons, that disgust’s influence on moral and legal reasoning is distortive—that it is inherently *unwise* (see for example Kelly and Morar, 2014; Kelly, 2011; Nussbaum, 2004). A third view is that disgust is neither wise nor unwise; perhaps it can be an appropriate emotional response to certain wrongs, but only if we have independent evidence that our disgust at those wrongs is properly directed.⁴²

<i>Epistemic Advocacy</i>	<i>Weak Skepticism</i>	<i>Strong Skepticism</i>
In at least some circumstances, disgust holds at least some justificatory value.	Disgust, in and of itself, is not a form of evidence and should thus be discounted in moral reasoning.	Disgust has a distorting influence on moral reasoning.

Table 2. Spectrum of views on disgust’s epistemic significance.

⁴² Kelly and Morar (2014) discuss a view like this, though they do not ultimately endorse it.

The focus of this chapter is the question referenced in the title: *Is repugnance wise?* More precisely, I am interested in the intersection between the Epistemic question and the Political question: Does moral disgust, in and of itself, hold any epistemic (or evidential, or justificatory) value? Should disgust ever be taken to help justify the judgments that it influences? Should it be permitted to do any justificatory work in public debates about policy or law?

My goal is to defend a view that falls within the territory of Epistemic Advocacy. I will not propose a novel thesis. Rather, I want to return to the set of texts that initiated the disgust advocate/skeptic debate in the late 1990s and early 2000s, in order to (1) make more rigorous and (2) defend an informal argumentative move that appears in them. I refer to this move as the “Appeal to Disgust.” The main texts that I have in mind include Kass (1997), Midgley (2000), Kahan (2000), Kekes (1992), Sandel (2004), and Levin (2007). These authors—call them “epistemic advocates,” for convenience—were engaged in various debates across bioethics, philosophy of law, and social/political philosophy. They were not coordinating with each other and in some cases may not have been aware of each other. Yet, all of them make a similar move: they cite the public’s deep sense of disgust or repugnance towards some action or policy, X, as at least some evidence that X must be seriously wrong, before any arguments for or against have been heard.

Essentially, I argue, the Appeal to Disgust is a burden-shifting move; proponents of X bear the burden of showing that legalizing or destigmatizing X would serve the public interest. This doesn’t mean that having a conversation about X is off-limits. On the contrary, we should conduct a vigorous public debate. However, in order to improve the *procedure* of public debate, we must start with the presumption that the public’s disgust towards X is justified.

The Appeal to Disgust has been maligned as “poetic” and “mystical” (Macklin, 2006: 34, 37), “unhelpful rhetoric” (Kelly, 2011: 162), “religious or crypto-religious” (Bostrom, 2005: 203), and “tinged with a pre-Enlightenment flavor” (Youngner, 2014: 364). In my view, none of these objections hold. On the contrary, as I will show, a number of testable claims fall out of the Appeal to Disgust, and taken together, these claims—if true—provide support for the burden-shifting conclusion sought by epistemic advocates. The two testable claims, and accompanying conclusion, that I have in mind are as follows:

1. *Inaccessible Premise*: Our moral disgust reactions are often rooted in reasons that we do not have direct introspective access to.
2. *Wisdom Premise*: Widespread feelings of moral disgust are sometimes directed at things that really are seriously wrong.
3. *Burden-Shifting Conclusion*: Those who do not share in the disgust of the majority bear the burden of argument.

If epistemic advocates may be accused of anything, it is failing to provide adequate support for the Inaccessible Premise and Wisdom Premise. The goal of this chapter is to fill this lacuna. I will construct arguments for both Premises.

A bit of work needs to be done before I get there, however. Namely, I must refute several arguments for Strong Skepticism (‘moral disgust is an inherently unwise emotion’). That will be the focus of the first half of the chapter, followed by my defense of the Appeal to Disgust.

The specific plan for the chapter is as follows. Sections 2-4 consider four arguments for Strong Skepticism: that disgust (i) moralizes, (ii) amplifies the severity of moral judgments, (iii) is insensitive to intentions, agency, and mitigating circumstances, and (iv) is inherently dehumanizing. Each of these claims faces empirical and philosophical problems and should be rejected. Next, Section 5 is purely interpretive: I explicate the Appeal to Disgust and resolve a number of misconceptions about it. Sections 6-8 detail my defense of the Appeal to Disgust. Specifically, I draw resources from moral psychology, social psychology, and gene-culture co-

evolutionary theory to construct arguments for the Inaccessible and Wisdom Premises, which together entail the Burden-Shifting Conclusion.

In Section 9, I conclude by showing how the arguments of this chapter provide the qualified defense of moral disgust's influence on criminal law that this dissertation has ultimately been working toward. Specifically, I return to the story of the Scottish public petitions committee, mentioned in the Introduction of the dissertation, and explain why, in light of my arguments in this chapter, the committee's swift dismissal of the incest petition was justified.

2. Does Disgust Moralize?

The next three sections (Sections 2-4) canvass some reasons in favor of Strong Skepticism about disgust. This view holds that, although disgust *does* influence moral judgment, it *ought* not to, because it is an inherently unwise or unreliable emotion. Far from being a guiding moral light, it injects noise and distortion into the moral reasoning process, increasing the likelihood that our conclusions will be incorrect. The two arguments I will consider are:

- (1) *Disgust is not a "moral" emotion at all; it is a moralizing emotion.* Extraneous feelings of core disgust (the sort triggered by feces) can (a) induce a moral judgment where one is not deserved or (b) amplify the harshness of a moral judgment.
- (2) *Disgust is inherently dehumanizing.* That is, (a) it is relatively insensitive to others' intentions or agency or to mitigating factors, and (b) it is a major component of racism or the cognitive machinery that gives rise to it, and may even have direct evolutionary roots in policing ethnic boundaries.

Either way, disgust is not wise. It is a bias—a glitch in our moral psychology—and the proper response to this information is to be vigilant against disgust and avoid letting it influence our moral reasonings.

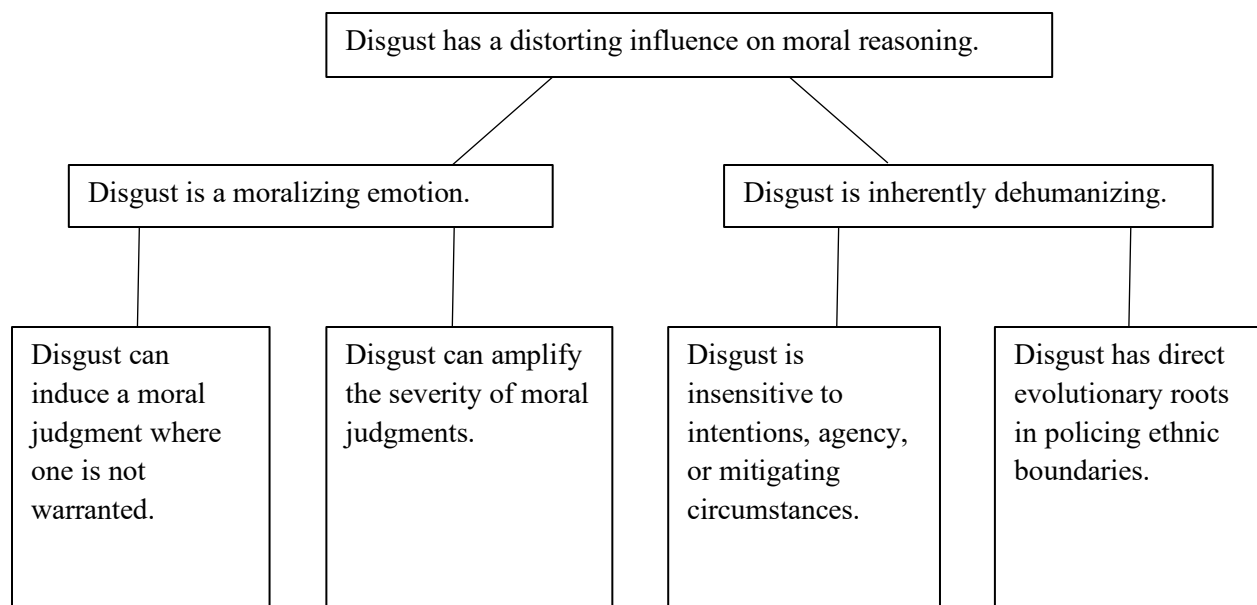


Figure 4. Different views within Strong Skepticism.

None of the above views is persuasive, I will argue. They either (1) rely on unsupported empirical claims about disgust or (2) derive normative conclusions from empirical claims that don't necessarily follow or don't apply to *disgust* more than other emotions.

The following subsections (Sections 2.1 and 2.2) consider two reasons to think that disgust *moralizes*. In both cases, the basic argument is something like this: (1) Disgust causes us to make unwarranted or excessively harsh moral judgments. (2) Any emotion that causes us to make unwarranted or excessively harsh moral judgments should be regarded as a pernicious influence on moral reasoning that we should be vigilant against. (3) Therefore, we should be vigilant against letting disgust influence our moral reasonings. Both versions of this argument to follow fail.

2.1. *Can disgust induce an unwarranted moral judgment?*

The claim that disgust “moralizes” sometimes means “Disgust can induce a moral judgment where none should exist.” This claim is typically rooted in a broader ontological

skepticism about “moral disgust,” according to which there is no emotional state that qualifies as both (1) genuine disgust and (2) a genuinely moral sentiment (see Bollard, 2022; Fileva, 2021; Giubilini, 2016; Gert, 2015; Bloom, 2004; Nussbaum, 2004). Rather, any time a subject appears to express disgust at a moral violation, either (1*) the subject is using disgust language or gestures *metaphorically* (to more effectively express some other state) or (2*) they are experiencing non-moral disgust (the sort triggered by feces) while *coincidentally* morally disapproving of the relevant action (e.g. disgust towards a bloody murder is directed towards the bloodiness, not the murder *per se*).

Because *physical* disgustingness is not a morally relevant property, if someone morally disapproves of a practice because it physically disgusts them, it follows that disgust is causing them to moralize. For example, people tend to judge incest and cannibalism to be morally wrong (Haidt and Hersh, 2001; Haidt et al., 2000), but this is just a case of moralization, says the ontological skeptic. These practices may have disgust-fitting features (e.g. they involve odd culinary and sexual behavior), but that is morally irrelevant.

Disgust might also amplify moral judgments. For example, imagine (1) George Bush tortures enemy soldiers by depriving them of food and sleep, in order to acquire information from them, and (2) Michael Myers tortures a babysitter by slowly severing one of her limbs. Obviously, both cases of torture are wrong⁴³, but the latter seems worse; Michael Myers should go to jail for longer than George Bush. Of course, if you think this because one of the cases of torture makes your stomach turn while the other just makes you angry, then you are succumbing to the “amplifying” influence of disgust.

⁴³ Actually, I’m not sure torturing enemy soldiers by starving them is wrong, depending on what information you’re trying to get (e.g. what if you could save 100 comrades by starving one enemy?). I’m just using this example because other authors do (see Bollard, 2022; Kekes, 1992).

As I explain in the previous chapter (Chapter 3), ontological skepticism about moral disgust is unsupported. There is strong theoretical and empirical reason to think that moral disgust is a genuine emotional state. On the theoretical side, as Joshua Tybur (2009) explains, it would have been adaptive for ancestral human cultures to recruit disgust to motivate a low-cost method of policing anti-social violations (see Chapter 3, Sections 3.2 and 4.3). On the empirical side, there are ways to test whether disgust is elicited by moral stimuli (e.g. hypocrisy or cheating); just operationalize its characteristic outputs, such as its facial expression (the gape) or neural signature (increased activity in the insular cortex or insula). When researchers do this, they find that subjects experience disgust in response to anti-social violations (see Chapter 3, Section 4).

As to the claim that incest and cannibalism, though disgusting, are not *morally* disgusting, authors who advance this claim tend to impose substantive constraints on what counts as a “moral” sentiment, such as “A *moral* judgment must invoke utilitarian considerations or concepts like harm, rights, and justice.” However, defining “moral” this way begs the question (Chapter 3, Section 7). If incest is not harmful, and you define “moral” as “involving harm,” then you’ve only avoided counting disapproval of incest as “moral” by selecting a definition favorable to your position. Moreover, imposing substantive constraints faces two additional problems: (1) It is philosophically unintuitive. It makes more sense to define “moral” judgments by formal criteria, such as categoricalness (“My judgment is not contingent on the agent’s self-interest or desires”) and universalizability (“My judgment applies to people in my own cultural context as well as people in other cultures”). (2) It is empirically dubious, excluding large swathes of values in non-Western cultures (and some Western subcultures) from consideration as “moral” values. (In Jonathan Haidt’s words, the list of substantive constraints that Western moral

psychologists tend to use “doesn’t travel well” (2012: 22).) If we focus on formal criteria only, we see that many subjects’ judgments regarding incest and cannibalism are categorical and universalized.

Lastly, about amplification. It is not the case that our differing intuitive judgments of the two torture methods (George Bush’s food/sleep deprivation versus Michael Meyer’s limb-severing) are unjustified. Rather, our intuitive judgments are responding to a real difference in the *attitudes* expressed by the perpetrators: George’s attitude is bad, but Michael Meyer’s attitude is *evil*. He expresses disdain for the basic moral conditions that underpin civilized life. The message that his chosen method of torture sends—“Behold your animality and vulnerability, mere mortals!”—is universally understood, and we rightly shudder and recoil at it (Chapter 3, Section 6).

2.2. *Can Hitler hypnotize you?*

A second meaning that the claim “Disgust moralizes” sometimes takes on is “Extraneous disgust can amplify moral judgments.” That is, being in a state of disgust while evaluating someone’s bad action can induce you to judge it more harshly than you otherwise would, even if the cause of your disgust is *completely unrelated to the action*.

Two classic studies in this space are Wheatley and Haidt (2005) and Schnall et al. (2008). Both basically do the same thing: have a control group and a test group provide a numerical moral rating (e.g. ranging from 0 = “perfectly okay” to 100 = “extremely morally wrong”) in response to a series of vignettes about people doing bad things (e.g. a student steals library books, a congressman takes a bribe, a trolley driver kills one bystander to spare five others), but induce a state of (completely extraneous) disgust in the test group. Wheatley and Haidt (2005) did this by “hypnotizing” test subjects to be disgusted by an otherwise neutral word (e.g. “often,”

“take”)⁴⁴ and then strategically placing that word in the moral vignette read by test subjects. Schnall et al. (2008) used a number of methods, such as discretely discharging novelty “fart spray” near subjects and (in a different experiment) having them sit at a messy, dirty desk.⁴⁵ The result: test groups judged the bad actors in the vignettes more harshly than control groups. For example, in Wheatley and Haidt (2005), the test group’s average rating of the bribery case was 91.28, compared to the control group’s 78.73, and for the library thief case those numbers were 71.24 and 69.40, respectively (p. 781).

Some authors are alarmed by these findings. Stephen Stich, in a lecture series on YouTube provocatively entitled “Debunking Morality: A Hodge-Podge of Multipurpose Kludges,” calls the findings “very weird and disquieting” and interprets them as showing that the test group’s moral judgments were “much more severe.”⁴⁶ Another disgust critic, Daniel Kelly (2011) discusses the findings of Wheatley and Haidt (2005) in relation to Nazi propaganda, which frequently used “the idiom of disgust” to characterize Jews as “uncivilized, barbaric, animalistic, and dirty” (p. 125). Kelly adds that disgust rhetoric

can be sadly effective, as this language not only expresses disgust but can invoke it in listeners and readers, as well. Once in play, the effect of disgust can be particularly pernicious because of the powerful but subliminal—perhaps powerful *because* subliminal—influence it can have on behaviors and judgments (2011: 125, original emphasis).

⁴⁴ Honestly, I don’t know how psychologists get away with stuff like this. How can you “hypnotize” someone to be disgusted by the word “often”? That sounds like magic. Nevertheless, Wheatley and Haidt (2005) is very widely cited and taken seriously, so I discuss it.

⁴⁵ “For the disgust condition, a workspace was set up to look rather disgusting: An old chair with a torn and dirty cushion was placed in front of a desk that had various stains, and was sticky. On the desk there was a transparent plastic cup with the dried up remnants of a smoothie, and a pen that was chewed up. Next to the desk was a trash can overflowing with garbage including greasy pizza boxes and dirty-looking tissues” (Schnall et al., 2008: 1103).

⁴⁶ <https://www.youtube.com/watch?v=fqFFw9EFzYM>

In other words, we might worry that people's moral reasonings can be manipulated by causing them to feel extraneous disgust. You might even worry that people's disgust reactions could be manipulated by a demagogue to help catalyze moral atrocities (e.g. the Holocaust).

I think the above authors may be overexcited about the research on the amplification effect of extraneous disgust. First, assuming the two classic studies discussed above are replicable⁴⁷, it is important not to overstate their findings. For example, when Stich says the test group's moral judgments in Wheatley and Haidt (2005) were "much more severe," that is an exaggeration. As Joshua May (2018) (a skeptic of the amplification effect) has explained, the study merely engages in null hypothesis testing and gets a statistically significant result. This suggests that the null hypothesis ("Extraneous disgust has *absolutely no effect* on moral judgment") is false. However, "statistically significant" is not the same as "substantial." The size of the effect in Wheatley and Haidt (2005) is actually quite small and, in some cases, barely noticeable; recall, the difference in average moral rating between the test group and the control group for the library thief vignette was only 1.84 points (71.24 versus 69.40) on a scale of 0 to 100.

What about the worry that you can manipulate someone's moral reasonings by using disgust rhetoric to put them in a state of disgust (e.g. make someone an anti-Semite by describing Jews in inherently disgusting terms)? Here, May (2018) notes that the amplification effect, even if it can (slightly) amplify moral judgments, does not change the *valence* of those judgments. That is, it does not *cause* people to judge that (e.g.) stealing a library book is wrong; people are already overwhelmingly agreed about that. Given this, it is implausible that you could make a

⁴⁷ A prominent meta-analysis raises doubt about this (see Landy and Goodwin, 2015). See also Jylkka et al. (2021).

significant impact on someone's reasonings, or change their mind entirely, by exposing them to extraneous disgust features.

For empirical support, May draws attention to a study by Wisneski and Skitka (2018), which measured the effectiveness of disgust-relevant anti-abortion propaganda (such as protest signs with pictures of aborted fetuses). The study found that such propaganda had no persuasive effect; those who were already pro-choice did not change their mind. If anything, the propaganda increased the *conviction* with which the subjects held their pre-existing judgments. Pro-lifers felt more motivated to oppose abortion, while pro-choicers were more motivated to defend it.

Reflecting on this, May (2018) writes:

The science of moral judgment suggests that appeals to repugnance are unlikely to change anyone's mind. Yet they are so common in politics. ... the history of political rhetoric is rife with attempts to portray marginalized groups as disgusting... While such propaganda is often meant to convince others to join opposition to a policy or practice, such attempts can gain traction simply by amplifying existing attitudes... So, while repugnant rhetoric may be futile when it comes to changing minds, the strategy might be useful for purely motivational tactics, such as rallying one's moral or political allies (p. 11).

In other words, Hitler can't hypnotize you into adopting anti-Semitism, but if you *already are* anti-Semitic, he can use disgust rhetoric to get you agitated, which may increase your readiness to coordinate with other anti-Semites to act on your dormant attitude.

An additional point worth making here is that, even if being in a state of disgust can slightly influence your moral reasonings (by causing you to judge a bad actor slightly more harshly than otherwise), it remains unclear that this is a unique feature of *disgust*. The effect could be mediated by negative affect, or just affect in general. That is, disgust is a negative emotional state; like fear and sadness, it is generally unpleasant to experience. Lots of studies claim to find evidence that being in a negative state can slightly influence moral judgment. For

example, hungry Israeli judges are less lenient in their rulings before lunch (when hungry) than after lunch (Danziger et al., 2011); sleepy Norwegian military officers are more likely to rely on strict moral principles when assessing moral dilemmas than well-rested officers (Olsen et al., 2010); and depressed Australian college students are more utilitarian in their judgments of the trolley problem than their non-depressed peers (Laing et al., 2022).

In summary, extraneous disgust may slightly amplify moral judgment or, in some circumstances, increase moral conviction. We should, of course, draw *good* lessons from this and not bad ones. A bad lesson to draw would be something like this:

To avoid prejudicial impact, we would advise steering clear of [providing jurors in criminal trials with] descriptions that can elicit disgust, such as graphic details about sexual behavior, including pictures and verbal descriptions ... [or] gruesome pictures of heinous crimes (Russell and Giner-Sorolla, 2011c: 5).

As I've already argued above in Section 2.1, the disgust we feel in response to heinous crimes is morally relevant (not extraneous) (see also Chapter 3, Section 6). So, it would be crazy to avoid showing jurors pictures or giving them verbal descriptions of heinous crimes (those details are pertinent to their deliberations!).

The correct lesson to draw is something like this: When having juries deliberate about criminal cases, to achieve best results, avoid making them tired, hungry, or depressed—and also avoid sitting them at sticky, stained desks near overflowing trashcans; in a room that smells strongly of farts; while forcing them to look at pictures of armpit tumors, maggots in dog food, and severe cases of ringworm that have nothing to do with the case at hand.⁴⁸

⁴⁸ These examples are taken from Jylkka et al. (2020), a study of both the moralization and amplification effect of disgust. None of the researchers' experiments found support for either effect.

3. Does Disgust Dehumanize?

This section and the next consider two different reasons to think that disgust is inherently dehumanizing. In both cases, the basic argument looks something like this: (1) Dehumanization is wrong (even if the person you're dehumanizing really is bad, such as a pedophile). (2) Disgust inherently dehumanizes its object (e.g. pedophiles are mentally construed as dirty). (3) Therefore, it is wrong to be disgusted by bad actors or bad practices (pedophiles still have rights and deserve moral regard).

Caveat: There may be no one definition or type of dehumanization (for discussion, see Clark and Fessler, 2015: 488-489). You may notice that the specific arguments discussed in Sections 4 and 5 take "dehumanize" to mean slightly different things. Nevertheless, it's not like the concept "dehumanize" is too unclear or confused to talk about meaningfully. The following arguments use the term in ways that are plausible and make intuitive sense, so they are worth considering.

My basic view on the dehumanization question is this: Disgust does often co-occur with racism, anti-Semitism, and other forms of prejudice and dehumanization, and this can have incredibly ugly implications. However, the emotion itself is not *inherently* dehumanizing in any plausible sense. For example, the empirical evidence does not support any of the following claims: (1) Disgust *always* causes us to see other individuals as animals, objects, or dirt. (2) Disgust, unlike (e.g.) anger or fear, is *uniquely* associated with, or the cause or origin of, racism and other prejudices. (3) The *ultimate evolutionary function* of disgust being extended to the social and moral domains was to police ethnic boundaries.

3.1. *Is disgust insensitive to agency?*

One argument that disgust dehumanizes rests on the claim that people's disgust reactions are not sensitive to intentions, agency, or mitigating circumstances. Here are strong skeptics Kelly and Morar (2014):

the connection between disgust and dehumanization is not merely an artifact or historical accident, but flows from the nature and operation of the emotion itself. For instance, psychological studies have shown how people in the grip of disgust are much less responsive to other people's intentions and to their very agency.

The authors then cite the following two studies: (1) Young and Saxe (2011) and (2) Russell and Giner-Sorolla (2011a). Both studies basically do the same thing: give subjects a series of emotion-arousing vignettes (anger or disgust) about someone who does something bad, use different manipulations in the vignettes (sometimes the bad actor intends to cause harm, other times the harm is accidental or unintended), and measure subjects' moral judgments of the vignettes. Both studies find that subjects' anger was responsive to cues related to intentionality. For example, in Young and Saxe (2011), an actor who *unknowingly* poisoned their peanut-allergic friend by giving them a casserole containing peanuts aroused less anger and was judged less harshly than an actor who *knowingly* poisoned their friend. Meanwhile, an actor who had a one-night stand with a long lost sibling aroused equal ratings of disgust and not-equal-but-similar ratings of wrongness whether their actions were intentional (they knew it was a sibling) or accidental (they didn't know it was a sibling).

What should we make of this? One interpretation is that moral anger is responsive to people's intentions, while moral disgust is not. When people knowingly harm others, it makes us mad; when they accidentally mess up, we give them a break. Meanwhile, if someone sleeps with their sister, moral alarm bells start going off no matter their intentions. Thus, disgust induces a sort of blindness to *moral agency*. Seeing others as agents—whose wrongful deeds are

blameworthy but whose honest mistakes can be forgiven—is central to seeing them as fellow humans, worthy of moral regard, as opposed to an animal or object. If you hold it against someone that they made an honest mistake outside their control, you are denying them their humanity and treating them like an object. This makes disgust especially problematic in legal contexts, says the authors of one of the studies (Russell and Giner-Sorolla, 2011a):

In legal philosophy, these findings support arguments that disgust is a less allowable emotion under liberal concepts of the law than anger. Determining intent is a key concept of liberal jurisprudence; the fact that disgust is not responsive to intent shows it to be a less reasonable emotion than anger, which responds to more legitimate concerns of harm and justice. ... in liberal concepts of the law, breaking a taboo is not a crime, unless it is done with the intent to harm (p. 237).

In other words, if someone is on trial for breaking a law, the fact that their actions were accidental should be a valid defense. As a rule, people should not be punished (or should be punished less harshly) for harms that they did not intend or could not control. Seemingly, disgust doesn't acknowledge this basic maxim; incest is viewed as wrong whether it was intentional or not.

I think we need to be very careful to avoid drawing conclusions from Young and Saxe (2011) and Russell and Giner-Sorolla (2011a) that do not follow. First, it seems relevant that Young and Saxe (2011), at least, did not find that moral disgust was unresponsive to intentions; on the contrary, accidental incest was rated as appreciably less wrong than intentional incest (an average rating of 4 versus a 6 on a 7-point scale). More precisely, disgust was *relatively less* responsive than anger (the accidental poisoning vignette received an average wrongness rating of 2, compared to intentional poisoning's 6.5).

Second, we should not conflate moral *wrongness* and moral *blameworthiness*. One obvious limitation of the above studies is that they did not ask the subjects to rate the blameworthiness of the actors in the vignettes; subjects were only asked to rate the wrongness of

what happened. As such, the fact that subjects rated accidental incest high in “wrongness” does not evince that subjects viewed the actor as *personally culpable*. In other words, suppose the actor were on trial for criminal incest, and the subjects were the jurors. The above studies do not provide evidence for the following claim: “The jury will not consider the actor’s lack of intent as a valid defense.” That claim *would* be disquieting, if true—but we do not have evidence that that is how the jury would behave.

The closest that Young and Saxe (2011) came to distinguishing between “wrongness” and “blameworthiness” was in asking subjects to provide separate ratings for “wrongness” and “disgustingness.” In asking this question, they found that (1) accidental and intentional incest were rated equally disgusting, but (2) intentional incest was rated as more wrong. Why? Here, it is possible that the subjects were being sensitive to the actor’s lack of moral blameworthiness, and compensating for this feature by lowering (though not erasing) their rating of wrongness.

In fact, the researchers themselves gesture at this possibility. They write that

“moral wrongness” could take on qualitatively different meanings depending on the context, just as the polysemous term “bank” refers sometimes to a place where money is kept and sometimes to a part of a river (p. 211).

They add, incest is a “purity violation” (not a harm violation), and purity violations are often “victimless”—“no one is directly affected by them” (p. 203). What makes them wrong (in the minds of many) is not harm or the creation of victims but, rather, the presence of intrinsic moral offensiveness. Thus, when subjects rate accidental incest as wrong, this does not imply that they are suddenly forgetting that intentions matter in assigning blame. Rather, they are responding to the presence of intrinsic offensiveness.

In other words, it’s not like incest occurring by accident makes it A-OK. Oedipus did not intend to have sex with his mom; nevertheless, when he finds out, he blinds himself, and Jocasta

hangs herself. One way to read this is “Wow, Oedipus and Jocasta are dehumanizing themselves. They are blaming themselves for moral wrongdoing, yet no one should be blamed for happenings outside their control.” But that would be an unintuitive reading of the moment. Instead, a better reading is “Not all moral emotions are responses to willful harm. Anger is; disgust is not. Disgust is a response to impurity and contamination, and Oedipus and Jocasta regard incest as morally impure. Their moral, or spiritual, lives have been polluted by accidental actions which they wish could be taken back, but cannot.”

I am not here defending the reasonableness of moral disgust towards incest; I am only pushing back against the claim that disgust’s being less sensitive to intent than anger somehow makes it a “dehumanizing” emotion. However, to conclude this section, I might note that the fact that anger is responsive to intent is not a reason in itself to consider it “more reasonable” than disgust. Arguably, anger can be excessively “humanizing” (as the Amazon Alexa in my kitchen that annoys the hell out of me well knows). Consider, for example, the prevalent phenomenon of “vending machine rage,” where people kick or rock vending machines that they believe have stolen their money—to the point of injuring or accidentally killing themselves (by causing the machine to fall on them) (Cosio, 1988)—in an attempt to exact revenge. Obviously, this form of anger is irrational; vending machines are not agents and do not have intentions.

4. Kelly’s Co-opt Thesis

This section considers a version of the “disgust dehumanizes” argument that is due to Daniel Kelly (2011). Kelly argues that the ultimate explanation for disgust’s extension into the moral domain was to police ethnic boundaries. Therefore, moral disgust is inherently dehumanizing; the whole reason it exists is to promote ethnocentrism. Kelly calls this the “Co-opt Thesis.” Below, I show that the Co-opt Thesis is false.

4.1. *Gene-culture co-evolution and the Coordination Problem*

Daniel Kelly's 2011 book *Yuck! The Nature and Moral Significance of Disgust* is essential reading for theorists working on disgust. It makes several invaluable contributions to the literature, not least of which is the (probably correct) "entanglement thesis" (discussed in Chapter 2, Section 3.3), which proposes to unite the most attractive features of the food-rejection and disease-avoidance theories of disgust's origins under a single banner.

Here, I want to take up the most momentous thesis from Kelly's book: the so-called "Co-opt Thesis." The goal of the Co-opt Thesis is to offer an ultimate explanation for why, at some point in human evolution, disgust—originally the pathogen/parasite police—became a socio-moral emotion (one used to enforce social and moral norms). The Thesis basically holds that disgust, as a result of various adaptive pressures, was extended to police ethnic boundaries and de-motivate interacting with out-group members. Importantly, Kelly proposes this thesis in the context of a broader theory of human evolution: gene-culture co-evolution. Specifically, the Co-opt Thesis is designed to develop an aspect of that broader theory by providing a solution to a puzzle that arises within it, which I will call the "Coordination Problem." Toward the end of the book, Kelly uses the Co-opt Thesis to develop a version of the "disgust dehumanizes" argument, as I explain momentarily.

First, let me present a brief overview of gene-culture co-evolution (GCCE), followed by how Kelly's Co-opt Thesis aims to supplement it. Note that the following discussion of GCCE covers eight points, which I have numbered. These eight points are standard fare in the field, not unique to Kelly.

Here's the gist of GCCE: Several of humans' most distinguishable traits (e.g. big brains, elaborate tool use, tribal psychology) are best explained by positing that the evolutionary path of

humans, unlike any other primate, led us to become *cultural* animals (Henrich, 2016; Tomasello, 1999; Richerson and Boyd, 1985). We were able to conquer innumerable diverse habitats and climates (from the tundra to the grassland to the tropical rainforest) not because we are individually smarter than other primates⁴⁹ but, rather, because we engaged in multi-generational efforts to (1) acquire and store up parochial, non-genetic information—cultural information—about how to survive in the local environment and (2) transmit those bodies of information to upcoming generations. The cumulative nature of this information quickly led to a sort of “ratchet effect” whereby (3) the growing size of cultural bodies of information combined with (4) their vital role in human survival (learn or die!) to introduce (5) a new set of selective pressures that shaped the human genome (i.e. it was as if cultural evolution said, “Ahem, Mother Nature, please select for bigger brains, bigger memory banks, and better social learning skills”). (This feedback loop between genes and culture is why the theory is called “gene-culture co-evolution.”) At some point in this process, cultural bodies of information became so complex—encompassing distinct cooking techniques, hunting processes, and social customs, norms and practices—that (7) different tribes of humans emerged with vastly different, and clashing, cultures, or “ethnicities.”

The emergence of distinct ethnicities led to a potential problem, which we can call the *Coordination Problem*: Suppose two ancient humans, Hunter and Fisher, need to make sure to engage in social learning from, and preferentially invest in interactions with, people from their respective cultures. This is because people in Fisher’s culture, F-tribe, have a clear set of expectations about the ‘proper’ form of customs, practices, and norms, and so does Hunter’s culture, H-tribe. Therefore, if Fisher interacts with people from H-tribe or Hunter interacts with

⁴⁹ Chimps and orangutans outperform, or perform equally well as, human subjects on a surprising number of working memory, information processing, and spatial, quantitative, and causal reasoning tasks (see Henrich, 2016: Chapter 2, especially the table on page 15). Meanwhile, human subjects consistently outperform chimps and orangutans on *social learning* tasks; in fact, very young children can blow adult apes out of the water.

people from F-tribe, they will fail to cooperate along a number of important dimensions (e.g. collective hunting) due to their expectations failing to mesh. This would result in equal detriment to all; maybe collective hunting technique A is just as effective as B, but we'll never catch any deer if you're following technique A and I'm following B, and then we'll all starve. In other words, Fisher (or Hunter) will fail to *coordinate* their expectations with people from H-tribe (or F-tribe). Yet, coordination is a necessary precondition for *cooperation*; we can only altruistically or pro-socially work together to collectively hunt deer if we first coordinate.

So, how do we solve the Coordination Problem? Easy: Introduce (8) a system to identify and track ethnic markers—arbitrary indicators that someone is from another tribe, such as dress or dialect—so that we don't get confused about who to socially learn from or cooperate with, and can focus our social investments on co-ethnics. In non-anthropomorphized terms: humans who could effectively track ethnic markers survived, while those who couldn't died, and this process of natural selection shaped the human genome and human psychology in important ways. (This concludes the eight points of GCCE.)

4.2. *The Co-opt Thesis and dehumanization*

This is all well and good. But in terms of the actual mechanisms involved, *which specific system* did ancient humans (as a result of a selection process) come to rely upon to (A) identify ethnic boundaries and (B) derive intrinsic motivation to avoid crossing them? This is where Kelly comes in. Kelly's Co-opt Thesis proposes that disgust—originally a biological system that safeguarded individuals against the dual threats of unsafe foods and transmissible sicknesses—stood out as “almost an ideal match” (p. 118) to be co-opted to perform, in addition to its primary, biological function, an auxiliary function: identifying members of other ethnicities and de-motivating interactions with them (thereby solving the Coordination Problem).

Kelly's use of the term "co-opt" is a reference to a basic evolutionary concept mentioned many times in this dissertation, namely: sometimes, in a quick-and-dirty way, existing traits or mechanisms are extended to serve new purposes for which they did not originally adapt but which, by sheer happenstance, they have features that make them a fitting candidate to be *recruited* to serve (see Chapter 2, Sections 3.2 and 6; see also Chapter 3, Section 3.2). For example, the mouth and tongue adapted to manipulate food; the lungs adapted to respire; but later these systems, being conveniently available, were enjoined to serve an additional function: speech. In the case of disgust, it is ideally fitted, says Kelly, to (A) identify ethnic markers because it was already sensitive to "phenotypic abnormalities"⁵⁰ (2011: 29-30, 49-50, 118)—i.e. outwardly visible deviations from "normal phenotypic health" (p. 49), such as being deformed, disfigured, or handicapped.⁵¹ Thus, insofar as members of other ethnicities exhibit phenotypic abnormalities (relative to the reference point of your own co-ethnics, anyway), they are ripe to trigger disgust. Kelly adds that disgust is also ideally matched to (B) de-motivate interacting with members of other ethnicities, insofar as the motivational tendency of disgust is to avoid or distance oneself from the triggering object.

So, that's Kelly's Co-opt Thesis. How does it lead to a "Disgust dehumanizes" argument? Answer: *disgust is a kludge*. What began as a biological safeguard was later co-opted to serve a new, social purpose (policing ethnic boundaries), but when this happened, it didn't shed any of its original, core outputs, such as the gape face, nausea, a phenomenological feel of offensiveness towards disgust triggers, or a sensitivity to contamination. On the contrary, it brought these

⁵⁰ Kelly doesn't specify what he means by "phenotypic abnormalities," so I'm not sure what he means. I guess he could be talking about racial differences (e.g. different skin color or hair form) or different clothes (e.g. "People in that other group wear funny hats").

⁵¹ An organism's "genotype" refers to the genes it carries; its "phenotype" refers to its observable characteristics, which are often genetically rooted. For example, some phenotypical traits of the category "dog" are "hairy body, long tail, four legs, and has canine teeth."

outputs along as “byproducts” (pp. 114, 133-135). This has resulted in a bit of “mismatch” between the emotion and the auxiliary functions it was recruited to perform (p. 135). Here’s Kelly (2011):

Disgust elicitors evoke the entire cluster of components that make up the disgust response, including the propensity to think about and treat those elicitors as if they were offensive and contaminating even when they are not. In instances where disgust underlies prejudices and ethnocentric attitudes, then, actors will not only avoid members of those tribes but will be more likely to project offensiveness and contamination potency onto them as well, and will thus judge them to be offensive, contaminating, unpleasant—in a word, disgusting. Those with different values and norms, members of other tribes that do things differently and give priority to different moral principles, will be thought of not just as different but as tainted, contaminating, immoral, and somehow less or lower than one’s own tribe—animal or subhuman (pp. 134-135).

In short, when you detect an ethnic marker or boundary, disgust is evoked, resulting in a cascade of automatic and reflexive—and quite ugly!—thoughts about the triggering object, which in this case is *another human being*.

Note that the Co-opt Thesis represents a far stronger version of the “Disgust dehumanizes” argument than the one considered in the previous section (Section 3) (which held that “Moral disgust reactions are relatively insensitive to agency”). The Co-opt Thesis proposes that the *ultimate evolutionary function* of moral disgust—the original reason disgust became a social and moral emotion—was to promote ethnocentrism (preferential interaction with co-ethnics). In other words, we could say—perhaps with a hint of melodrama—the following: It is not that moral disgust has irrational or unsavory downstream effects; rather, it is, by Mother Nature’s amoral and uncaring design, an irrational, kludgy, dehumanizing emotion.

4.3. *Three problems with the Co-opt Thesis*

Here, I explain three reasons why I think the Co-opt Thesis is simply not true. Kelly's argument for the Co-opt Thesis depends, I believe, on three key claims, all of which need to come together for his argument to work:

- (1) *Explanatory Claim*: The Co-opt Thesis is a valuable supplement to gene-culture co-evolutionary theory (GCCE). It offers a solution to the Coordination Problem that GCCE does not otherwise have the resources to solve.
- (2) *Flexibility Claim*: The Co-opt Thesis, which holds that disgust was extended to be triggered by ethnic markers, particularly "phenotypic abnormalities," is *possibly* true because the disgust acquisition system is highly flexible and open-ended. That is, "there is some latitude in what 'abnormal' amounts to" (Kelly, 2011: 118).
- (3) *Empirical Claim*: The available empirical data on the frequent co-occurrence of disgust and dehumanization is predicted by the Co-opt Thesis. Therefore, these data constitute evidence in favor of the Co-opt Thesis.

As I argue below, these claims are false.

The Explanatory Claim is false.

Start with the Explanatory Claim, probably the most important pillar undergirding Kelly's Co-opt Thesis. The Claim holds that GCCE does not have the resources to solve the Coordination Problem—a gap which the Co-opt Thesis can fill in. The problem is, this isn't true. Kelly is correct that theorists working on GCCE widely acknowledge the Coordination Problem, but they also solve it, without referencing any role for disgust.

For example, consider how Joseph Henrich (2016)⁵², a leading figure in the field of GCCE, lays out the Coordination Problem:

Cultural evolution gave rise to a variety of different social norms, so different groups became increasingly characterized by different practices and expectations about such things as marriage, exchange, sharing, and rituals [1]. Then, natural selection acting on

⁵² Henrich's 2016 book *The Secret of Our Success* received glowing reviews from other leading GCCE researchers, including Alex Mesoudi, Robert Boyd, Peter Richerson, and Michael Tomasello, so I take it to be an authoritative source on the subject. <https://secretofoursuccess.fas.harvard.edu/reviews>

genes responded to this world governed by social norms by endowing individuals with the cognitive abilities and motivations to help them better navigate and adaptively learn. The success of persons growing up in this emerging landscape of social norms depended – at least in part – on their ability to acquire the appropriate social norms for their own group and to preferentially target their interactions toward those most likely to share their norms [2]. If learners acquire social norms that don't fit with others in their group, the learners will end up violating the local norms, getting a bad reputation, getting punished, etc. Even if they acquire the appropriate local norms, but then interact with people from other groups, who have different norms, they can end up getting sanctioned, wasting time, or miscoordinating [3] (p. 200, numbers added).

Notice that Henrich captures every aspect of the Coordination Problem that Kelly thinks is important: Point 1 acknowledges the emergence of ethnicities; Point 2 acknowledges that ancestral humans faced selective pressure to “preferentially target their interactions” towards co-ethnics; and Point 3 acknowledges that those who failed to preferentially target co-ethnics for interaction risked “miscoordinating” (a potentially deadly mistake).

So how do humans, according to Henrich, overcome the Coordination Problem? The short answer: not by being *repulsed* by signs that someone is an out-group member, but rather by being *attracted*, and attuned, to observable signs of co-ethnics. One point that Henrich stresses—which Kelly doesn't mention—is that, yes, ancestral humans needed to pay attention to observable ethnic markers. However, the “best markers are those that are difficult to fake” (p. 201). This is because an “easy-to-fake marker, like a distinctive hat, can be simply put on in order to trick or manipulate someone” (p. 201). In contrast, language or dialect are “better markers” because they are “not easy to get right unless one grows up in a certain place or within a certain social group” (p. 201). In fact, “language or dialect might be a *priority cue* for determining who to learn from and interact with” (p. 201, emphasis mine).

Because the markers you are looking for are (1) subtle and (2) hard to get right, it is less important to pay attention to signals of abnormality or otherness (a negatively defined category,

like “Everything that is not an umbrella”) and more important to be finely attuned to markers of genuine co-ethnicity (a positively defined category). When such markers are found, our brains light up: “cues related to...ethnicity fire up our cultural learning psychology and spark our interest in what the model is doing, or talking about, and focus our attention and memory” (p. 45). For empirical support, Henrich references fascinating research on infants and young children who preferentially learn from models who speak their language, even before they are able to speak themselves:

Young children preferentially acquire both food preferences and the functions of novel objects from those who share their language or dialect. This is true even when the potential models speak jabberwocky—English-sounding speech composed of nonsense words. That is, kids prefer to learn from those speaking nonsense *in their own dialect* rather than from someone speaking nonsense in a different dialect... infants preferentially copy the more unusual and difficult action—turning on a lamp with their head—from someone who speaks their language (German) as opposed to an unfamiliar language (Russian) (p. 45, emphasis original).

In short, the Coordination Problem is solved, not by out-group members being repellent, but rather by co-ethnics being magnets.

Henrich also notes that perceived racial or phenotypic markers (e.g. “skin color or hair form”) are not as salient as ethnolinguistic ones:

racial cues do not have cognitive priority over ethnic cues: when children or adults encounter a situation in which accent or language indicate ‘same ethnicity’ but skin color indicates ‘different race’, the ethnolinguistic markers trump the racial markers. ... children pick as a friend someone of a different race who speaks their dialect over someone of the same race who speaks a different dialect (p. 205).

This, I think, is the nail in the coffin for the Explanatory Claim; the Co-opt Thesis posits that disgust was easily recruited to detect ethnic markers because (1) it was already sensitive to “phenotypic abnormalities” and (2) presumably, other ethnicities will exhibit these (e.g. different

skin color or hair form). But humans' ethnic detection systems are more attuned to linguistic than phenotypic features.

In sum, the Explanatory Claim is false because GCCE *does* provide resources to solve the Coordination Problem without reference to disgust. Thus, the Co-opt Thesis is a theoretical postulate that we don't need and should, in the spirit of Occam's razor, reject.

The Flexibility Claim is false

Another problem for the Co-opt Thesis is that it rests on the Flexibility Claim, which holds that disgust's acquisition system is highly flexible. In Kelly's words, the system has an "open-ended database of elicitors" and can be "easily activated by new types of cues" (p. 118). This is an important claim because this is how, for Kelly, disgust was able to *acquire* new sets of elicitors: markers of ethnic-otherness, such as abnormal dress (different hats), abnormal speech (different dialect), or abnormal phenotype (different hair). A point that Kelly does not consider here is that even if disgust's acquisition system is quite flexible (it is!), it is not *limitlessly* flexible. It is bound by certain "prepared" limits. As noted in Chapter 2, Section 5.2, you probably can't teach a child to be disgusted by sugar; you probably also can't teach a child *not* to be disgusted by feces or vomit (see also Chapter 3, Section 4.3). Even if co-opting disgust to police ethnic boundaries had been necessary to solve the Coordination Problem (it wasn't), would Mother Nature have been able to "trick" ancestral humans' disgust systems into being triggered by e.g. other languages or other dialects (paradigmatic ethnolinguistic markers)? Plausibly, no; this goes beyond the prepared limits of the acquisition system. Unfortunately, nowhere in the book does Kelly discuss the concept of preparedness, and his argument suffers as a result.

The Empirical Claim is false

Finally, consider the Empirical Claim. By Kelly's own admission, the Co-opt Thesis is the most theoretical and least empirically supported part of the book. The best empirical data he can marshal to support his theoretical suppositions is that disgust and dehumanization frequently co-occur—which is true—and this is what the Co-opt Thesis would predict. The problem is, the data on the co-occurrence of disgust and dehumanization can be explained by means that (1) are already familiar and rooted in the mainstream empirical literature on disgust and that (2) have nothing to do with, nor require positing, the Co-opt Thesis. Hence, the co-occurrence of disgust and dehumanization, though interesting in its own right, does not constitute *evidence in favor of* the Co-opt Thesis.

Much of the evidence that disgust and dehumanization co-occur is qualitative. Martha Nussbaum (2004) deserves credit here for thoroughly compiling this evidence (see especially Chapter 2, Sections 5 and 6 of her book). For example, she notes that anti-Semitic propaganda in Nazi Germany did not intend to portray Jews as frightening, or deserving of anger; it portrayed them as “soft and porous, receptive of fluid and sticky,” “womanlike in [their] oozy sliminess,” in stark contrast to the “clean safe hardness of the true German man” (p. 108). Hitler himself compared the Jewish population in Germany to “an abscess...a maggot in a rotting body” (quoted in Nussbaum, 2004: 71).

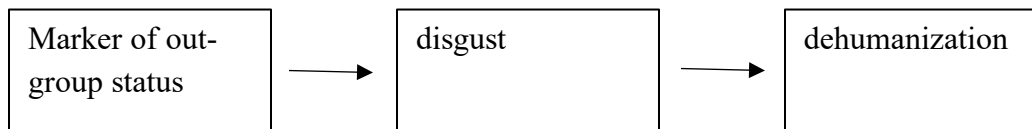
In terms of quantitative data showing a link between disgust and dehumanization, in the entire book (as far as I can tell), Kelly cites just a single pair of neuroimaging studies by Lasana Harris and Susan Fiske (2007; 2006). Both studies found that if you show subjects pictures of homeless people and drug addicts while using an fMRI machine to record their brain activity, you find that, on average, subjects' experience (1) increased activity in the insula and amygdala, a pattern consistent with disgust and fear, and (2) decreased, or no, activity in the medial

prefrontal cortex (mPFC), a brain region necessary for distinctively social cognition. Meanwhile, if you show the same subjects pictures of other types of people—elderly people, disabled people, middle-class Americans, business professionals—subjects’ brains exhibit activity in the mPFC.

What should we make of the qualitative and quantitative data? One takeaway, in my view, is: we probably cannot fully understand the evils of dehumanization, racism, and anti-Semitism without first understanding the science of disgust. This is an important takeaway in its own right.

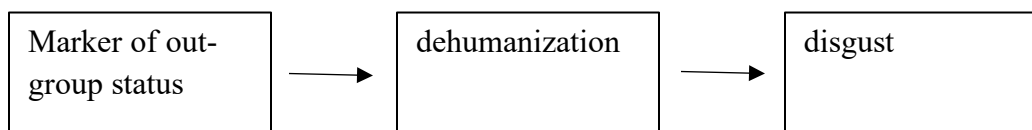
But do the above data, as per the Empirical Claim, support the Co-opt Thesis, insofar as the Co-opt Thesis would predict that out-groups are dehumanized? That particular inference is not deductively valid, in my view. The causal model implied by the Co-opt Thesis looks something like this:

Disgust dehumanizes



The qualitative and quantitative data are *consistent* with this causal model. However, they are also consistent with possible alternatives, such as:

Dehumanization disgusts



In other words, why think that disgust is a *cause* of dehumanization, as opposed to a *consequence*? It is worth noting, as other authors have, that disgust is not *uniquely* associated with dehumanization (see Clark and Fessler, 2015: 487-489). Other emotions that can contribute to dehumanization, and often cognitively underpin dehumanizing attitudes, include anger and

fear (Clark and Fessler, 2015: 488) and hate (Sternberg, 2003). In other words, it is not the case that, if we could press a button and make disgust disappear, we would thereby make dehumanization disappear.

Actually, the “Dehumanization disgusts” model makes quite a bit of sense. A look at the psychology of disgust reveals that disgust objects are viewed as offensive (see Chapter 2, Section 2). “Offensiveness” can be cashed out in a number of ways. One of the common ways is to cash it out as “lowliness” or “debasedness.”⁵³ For example, when cockroaches are mentally construed as offensive, this may partly be because they are seen as *low*; they sneak, scavenge, hide in cold, dark nooks and crannies, and are primarily nocturnal. If we were to organize all the animals on earth into a hierarchy, the cockroach would be at the lowest-ranking tier. To be clear, we do not rank it as low because it disgusts us; we are disgusted by it because it *is* (perceived as) low. In a similar way, it may be that, in societies that have strict social hierarchies (e.g. the Indian caste system or the Antebellum Southern U.S.), those at the bottom of the hierarchy become dehumanized in virtue of being *very* low-ranking, and the disgust of the higher classes is simply a response to this low rank, albeit an incredibly unfortunate one.

I’m not necessarily endorsing the “Dehumanization disgusts” causal model. I’m just trying to raise reasonable doubt about Kelly’s inference from (1) “disgust and dehumanization co-occur” to (2) “disgust *causes* dehumanization”—an inference implicit in the Empirical Claim.

4.4. Conclusion

In conclusion, the Co-opt Thesis depends on three key claims: the Explanatory Claim, the Flexibility Claim, and the Empirical Claim. Each is false. It follows from this that the Co-opt Thesis is false, or at least not evidentially supported.

⁵³ For example, Miller (1997) writes “Contempt, disgust, and shame rank us and order us in hierarchies; these emotions provide the basis for honoring and respecting as well as dishonoring and disrespecting” (p. 202).

At this point, you might be thinking the following. The Co-opt Thesis was Kelly's attempt to offer an ultimate explanation for disgust's becoming a socio-moral emotion. If that explanation doesn't work, what then is the ultimate explanation? I've already addressed this question elsewhere. I am persuaded by Joshua Tybur's (2009) theory that disgust was extended to the moral domain to police anti-social violations. In other words, disgust became moral in order to *enforce compliance with in-group norms*, not to de-motivate interaction with out-groups.

This theory would predict that norm violations committed by (1) one's own co-ethnics would arouse greater disgust than norm violations committed by (2) out-group members. It turns out, there is at least some experimental support for this. Joseph Henrich (2016: 191-195) points out that, when people play the Ultimatum Game with co-ethnics, they almost always reject unfair offers—unlike chimpanzees who play the game, who always accept unfair offers (the most economically rational choice). This, says Henrich, indicates that humans are willing to “pay to punish”: our tribal psychology makes us instinctively willing to incur a personal cost to enforce group norms. However, interestingly, when the Game is played with out-group members, people are much more likely to *accept* unfair offers. As discussed elsewhere in this dissertation, the emotional response typically associated with unfair offers in the Ultimatum Game is disgust (i.e. people are disgusted by unfair offers) (Chapman et al., 2009; Sanfey et al., 2003). So, if people are more likely to accept unfair offers from out-group members, this suggests that they are *less* disgusted. Why might that be? Here's a plausible explanation: Disgust is an emotion that enforces compliance with in-group norms. Someone who is not part of one's in-group cannot comply with these norms anyway. So, there is no sense in becoming disgusted by their unfair treatment of you; your disgust will not go toward enforcing any norms. It would be a waste of emotional investment.

A final point: I challenged the Co-opt Thesis on the grounds that it exaggerates the flexibility of the disgust acquisition system. Could Tybur's theory be challenged on the same grounds? No. There is a strong psychological fit between disgust and anti-social violations, as several authors note (see Plakias, 2018; 2013; Kumar, 2017). Fundamentally, disgust is an emotional system that is sensitive to (1) offensiveness and (2) the spread of contamination. Anti-social violations (e.g. freeriding) do in fact have a tendency to spread via social contagion if not contained (see Chapter 3, Section 3.4). In other words, when Jesus in Matthew 23:27 calls the Pharisees "hypocrites" whom "inside are full of the bones of the dead and everything unclean," he is not metaphorically expressing anger; he's disgusted. His innately prepared moral psychology is detecting contaminating social behavior. Whether it is morally *appropriate* to call hypocrites "unclean" is a separate question; perhaps it is harmful. The point is that, psychologically, Jesus has the cognitive machinery to *see* hypocrites that way. In short, it would make sense, as Tybur's theory maintains, that some ancestral humans were highly disgust-/contamination-sensitive and capable of directing this sensitivity toward forms of social contagion, and over time, evolution (whether biological or cultural) selected for those disgust-sensitive humans. By contrast, Kelly's book nowhere acknowledges that disgust, insofar as it is a contamination-detection system, can in fact be fitted to anti-social violations; instead, he suggests, unpersuasively, that our sense of contamination towards moral violations is a kludgy, irrational, mismatched "byproduct" of disgust's co-optation into the social domain to promote ethnocentrism.

5. The Appeal to Disgust

The preceding sections (Sections 2-4) argue that Strong Skepticism about disgust is false. There is no reason to suppose that disgust distorts moral reasoning. The goal of the rest of the

chapter is to argue that there is at least one plausible sense in which disgust, or “repugnance,” is indeed wise, as per Kass’ infamous claim. In other words, I will argue that Epistemic Advocacy is plausible—in some circumstances, moral disgust holds some justificatory value.

I do not intend to advance a novel thesis here. Rather, as noted in Section 1, I want to defend an informal argumentative move that first began to appear in the 1990s and early 2000s in debates across bioethics, philosophy of law, and social and political philosophy, which I will call the “Appeal to Disgust.” The goal of this section is purely interpretive: to explicate the Appeal to Disgust. The substantive claims of the Appeal will be defended in subsequent sections.

In essence, the Appeal to Disgust involves citing the fact that a majority of people consider an action or practice, X, morally disgusting as at least some evidence that X is, in fact, seriously wrong or should be prohibited. The move is intended to provide pre-discursive or pre-theoretical support for opposing X, relieving opponents of some of the burden of providing a specific, substantive argument. The appeal has been particularly prominent in debates about the ethical and legal status of emerging biotechnologies, such as human cloning, enhancement drugs (to improve humans’ mental or physical capabilities), and genetic engineering (see for example Giubilini and Sanyal, 2015; Mills, 2014; Youngner, 2014; Roache and Clarke, 2009; Levin, 2007; Macklin, 2006; Bostrom, 2005; Midgley, 2000; Kass, 1997).

5.1. Four misconceptions about the Appeal to Disgust

The best way to start explicating the Appeal to Disgust may be to anticipate several misconceptions about it at the outset. In the literature, the Appeal has received significant pushback (see for example May, 2018; Youngner, 2014; Roache and Clarke, 2009; Macklin, 2006; Bostrom, 2005). Some of the objections to the appeal that came up deserve careful consideration and demand a response (Sections 2-4 of this chapter considered and refuted four

such objections). However, a number of complaints about the appeal simply mischaracterize the dialectical function its proponents intend it to serve. These complaints can be organized into four groups:

Misconception A: The Appeal to Disgust holds that disgust is an infallible moral guide.

A number of critics interpret disgust advocates as suggesting that moral disgust is a supremely reliable emotional reaction; whenever a practice triggers it, this is sufficient to believe the practice is wrong or bad. For example, Ruth Macklin (2006), while expositing Kass (1997), attributes to Kass the view that revulsion is “*sufficient* grounds for rejecting any public policy short of total prohibition” (p. 38, emphasis mine). Similarly, Mills (2014) understands Kass as proposing that “repugnance is *enough in itself* to show that a practice is wrong” (p. 368, emphasis mine) and that “some moral intuitions are beyond debate” (p. 369).

Of course, if the Appeal to Disgust is based on a conception of the emotion as infallible, this immediately undermines its credibility, insofar as disgust’s fallibility is easily demonstrated. There are lots of historical examples of societies that once viewed a practice as repugnant but eventually came to accept the practice as innocuous or even good, and rightly so. For example, as Youngner (2014) points out, moral objections to dissecting cadavers in Europe and the U.S. faded away once the medical benefits were acknowledged (p. 360). We might also note that, across Europe and the U.S., disgust was once a common reaction toward interracial or interreligious marriages (Macklin, 2006: 38). Of course, that was an obvious case of misplaced and harmful disgust.

Misconception B: The Appeal to Disgust is a merely religious argument, unsuitable for discourse in a pluralistic democracy.

A number of critics complain that appeals to disgust “derive from religious or crypto-religious sentiments” (Bostrom, 2005: 203) or “spring from deep religious beliefs” (Youngner, 2014: 363). Of course, such private judgments have “little to offer” important conversations on science policy (Youngner, 2014: 363), being non-neutral and centering on premises that many people simply do not share. Macklin (2006) elaborates:

In pluralistic societies, people do not all share the same repugnances. . . Many Christians [abhor abortion and stem cell research], whereas a large number of other Christians, many Jews, agnostics, and atheists do not share that repugnance or religious moral foundation” (p. 39).

Obviously, unshared premises don’t make for persuasive arguments.

Misconception C: The Appeal to Disgust is anti-reason, anti-intellectual, or anti-argument. As such, it stifles productive debate and dialogue.

Several authors complain that the Appeal to Disgust is not just a bad philosophical move; it is not a *philosophical* move at all. Rather, it is “mystical,” “emotional,” and “poetic” (Macklin, 2006: 34, 37); tinged with a “pre-Enlightenment flavor” (Youngner, 2014: 364); or rooted in “some bias or philosophical prejudice” (Bostrom, 2005: 212).

Critics further claim that the appeal “disavows reason” (Macklin, 2006: 24), substitutes “metaphors and slogans” for “empirical evidence and reasoned arguments” (p. 38), and is “invoked as a polemical substitute for clear ideas” (Bostrom, 2005: 209). Because the Appeal to Disgust has this anti-intellectual character, it “precludes reasoned response” (Macklin, 2006: 39). For example, policymakers who favor a proposed policy will not be dissuaded from enacting it by “opponents who cite their profound repugnance or heartfelt intuitions” (p. 39).

Because the appeal is not an argument and cannot persuade, critics complain that it kills dialogue. Macklin (2006) writes, “It is not clear...whether a dialogue between [those who rely on the appeal and their opponents] is possible” (p. 42). Roache and Clarke (2009) agree: many ethicists just want to “move the debate forward” on various issues, yet those who appeal to disgust “do not wish to engage in unacceptably analytic discussions” (p. 7).

Misconception D: The Appeal to Disgust is radical, illiberal, and contradicts the harm principle.

Many critics of the Appeal to Disgust feel that, setting aside the question of whether the move is fair or unfair, the move is at least radical and unusual, one typically not heard in contexts where liberal-democratic norms reign. For instance, the move says nothing about harm, yet “people who live in a liberal, democratic society are free to pursue whatever lifestyle they wish,” provided they are not “significantly harming others” (Bostrom and Roache, 2008: 5). Bostrom (2005) concurs:

A liberal democracy should normally permit incursions on morphological and reproductive freedoms only in cases where somebody is abusing these freedoms to harm another person (p. 210).

Given the reigning norms in democratic societies, those who Appeal to Disgust are challenging “the current default position” and advocating “a radical departure from accepted thinking” (Roache and Clarke, 2009: 8). To compensate for this radicalness, “their case needs to be very strong” (p. 8).

5.2. A how-to guide for appealing to disgust

The above objections would be forceful if the targeted view were construed accurately. As is, however, they get the Appeal to Disgust all wrong—or at least they fail to engage with its strongest possible interpretation.

This confusion may partially be the fault of those who deploy the Appeal to Disgust themselves, insofar as they sometimes fail to be optimally clear about how they intend the move to function within the dialectic. Nevertheless, carefully reviewing the literature in which the Appeal appears will help reveal what it is and what function it serves (in the minds of those who deploy it, anyway). And once we explicate the move better, we see that the above complaints largely dissipate.

The following is not intended to be an argument that the appeal is a fair move (this section remains agnostic on that point). It is a purely interpretive attempt to clarify the move's various elements and its intended dialectical function. While different theorists present the appeal in slightly different ways, all iterations, I believe, follow six basic steps. Those steps are as follows.

Step 1: Target a practice for ethical critique for which there are no obvious or decisive harm-based, consent-based, fairness-based, or utilitarian objections.

The very first thing to notice about the Appeal to Disgust is that it doesn't come up in every ethical debate. It only appears in a special set of contexts: those in which the ethical status of a given action or practice is contested, but the conventional sorts of considerations/moral concepts that ethicists draw upon don't clearly or neatly apply.

For example, in the midst of analyzing the ethics of cloning human beings, Kass (1997) laments that a liberal or libertarian outlook does not provide an adequate vocabulary for critiquing cloning. This is because liberal/libertarian analyses typically center notions of "rights freedoms, and personal empowerment" and aim to ensure that whatever practice is under scrutiny allows for "adequately informed consent" and "the avoidance of bodily harm" (p. 688).

But these notions provide “thin gruel” for bioethical analysis (p. 683). After all, says Kass, it is quite unclear whom cloning harms or whose consent it violates.

Michael Sandel (2004) makes a similar remark in his analysis of the ethics of “designer babies” (babies whose genes are edited to give them traits desirable to their creators/parents). He writes:

In liberal societies they reach for the language of autonomy, fairness, and individual rights. But this part of our vocabulary is ill equipped to address the hardest questions posed by genetic engineering.

In other words, Sandel intuitively feels that using gene-editing to “design” babies seems morally problematic—but the reasons underlying this intuition cannot be readily translated into the moral *lingua franca* of liberal societies.

Step 2: Appeal to the widespread repugnances of humankind, or to the disgust of the ‘ordinary’ (or reasonable, or average) person.

Once you’ve identified a practice to target for ethical critique, the next thing you should do is appeal to the fact that the practice in question *widely* elicits disgust. Appealing to your own disgust isn’t very impressive (what if your interlocutor doesn’t share your feelings?). But cite the fact that most people are disgusted by something, with the exception perhaps of a small minority, and now you are appealing to a sentiment that your interlocutor can’t dismiss as idiosyncratic or outside the mainstream.

For instance, Kass (1997) speaks of “the widespread repugnances of humankind” (p. 689), which arise “from the man or woman in the street and from the intellectuals, from believers and atheists, from humanists and scientists” (p. 686). He further suggests that a cloning ban would be supported by “the overwhelming majority of our fellow Americans, and of all the human race, and of most...practicing scientists” (p. 701). Levin (2007) cites an opinion poll

according to which “over 90 percent of the American people” oppose cloning (p. 58). Kekes (1998) writes that “moral disgust” is the emotional response to actions that “just about everybody in contemporary Western societies would find disgusting” (p. 102); it’s “the normal reaction” (p. 103). Midgley (2000) argues that it is important for ethicists to pay “close and respectful attention” to the “fears and apprehensions of ordinary citizens” and “all those agitated radio callers” (p. 10).

The basic idea here is that the fact that most (if not all) people—or the ‘ordinary’ person—finds something, X, disgusting serves as some amount of evidence that X is wrong. When widespread, repugnance is “wise.”

Step 3: Acknowledge that disgust is not an infallible moral guide.

The next thing you should do is acknowledge that widespread disgust—albeit of some evidentiary value—is not a knockdown, bulletproof, airtight argument. Widespread repugnance can be, and often is, mistaken.

By far the most frequently quoted line from Kass’ infamous 1997 paper “The Wisdom of Repugnance” is this: “In crucial cases...repugnance is the expression of deep wisdom, beyond reason’s power to articulate it” (p. 687). Much less frequently acknowledged is the sentence *directly preceding that one*: “Revulsion is not an argument; and some of yesterday’s repugnances are today calmly accepted...” (1997: 687). Whatever “wisdom” Kass imputes to disgust, he makes sure to acknowledge that the emotion is not incapable of erring. He adds elsewhere, the emotion *can* be triggered by mere “strangeness and novelty” (p. 687), which are hardly morally relevant properties; it’s just that, in the case of cloning (and some other technologies), it is *not* unfamiliarity that elicits disgust, but rather a sense that cloning violates values or crosses boundaries that “we rightfully hold dear” (p. 687).

Like Kass, several others who make the Appeal to Disgust acknowledge that widespread disgust, though evidentially significant, is not infallible. Midgley (2000) explicitly writes that repugnance is “not an infallible guide”; it must be “supplemented by thought,” its meaning analyzed, and articulated into something “that gives us coherent and usable standards” (p. 9). “Unanalyzed feelings sometimes turn out to be misplaced” (p. 9).

Kahan (2000) similarly writes, “The insights of liberalism do indeed give us...much reason to be wary of disgust—indeed, of all emotions—in criminal law” (p. 70). “The proper course for liberalism,” however, “is not to obliterate disgust but to reform its objects so that we come to value what is genuinely high and to despise what is genuinely low” (pp. 70-71). “Even egalitarians hold pedophiles and sadists in low esteem” (p. 70). In short, while there can be properly directed disgust, there can also be *improperly* directed disgust.

Like the preceding authors, Kekes (1992) also acknowledges, quite emphatically, that disgust can be a “morally mistaken reaction” (p. 438):

feeling in general, or disgust in particular, is an unreliable moral reaction. This is...because the moral value to which deep disgust, or any other feeling, may indicate a commitment could be noxious. What if twelve randomly chosen Germans or slaveowners felt deep disgust at the thought of a Jew or a black sleeping with an Aryan or a white woman? ... Is there not such a thing as a vicious society? (p. 438).

Step 4: Anticipate that the reasons underlying our disgust reaction may be difficult to put to words.

The next step—and this is an important one—is to anticipate that even if our widely shared disgust towards a given practice is undergirded by good reasons or good arguments, it may turn out to be surprisingly difficult to articulate those reasons. Even after much contemplation, we may struggle to piece together a strong case. Of course, this need not worry us; we may be struggling not because the disgust is unfounded, but because the reasons

undergirding the disgust themselves are deep, complicated, and nuanced, such that the task of articulating them tests the outer limits of our rational or philosophical capacities.

This is clearly the point Kass (1997) is making here:

Can anyone really give an argument fully adequate to the horror which is father-daughter incest (even with consent), or having sex with animals, or mutilating a corpse, or eating human flesh, or even just (just!) raping or murdering another human being? Would anybody's failure to give full rational justification for his or her revulsion at these practices make that revulsion ethically suspect? Not at all (p. 687).

In other words, on the one hand, it is hard to say what makes (e.g.) incest wrong. On the other hand, it feels quite odd to question or doubt that incest is wrong on the basis that we struggle to articulate reasons for our disapproval of it. This doubt is overpowered by the singularly vivid and urgent sense of moral revulsion incest makes us feel.

Other authors mirror Kass here. Sandel (2004), for instance, writes that “it is not easy to identify what exactly is wrong with screening embryos to choose the sex of our children” (p. 4). He adds,

When science moves faster than moral understanding, as it does today, men and women struggle to articulate their unease. ... The genomic revolution has induced a kind of moral vertigo.

Similarly, Levin (2007) writes:

[A taboo] marks a barrier whose violation would strike so deep that we would not have the words to describe it, but we would understand such a violation fully and at once. This unspoken understanding seems always to surround taboos. . . our stomachs recoil, even if at first our minds do not (p. 54).

And here's Kekes (1992):

People feel [repugnance] because their moral education makes them feel [it] at some acts. And they come to feel it not because they have memorized a list of horrors, but because they have learned what is appropriate and what sort of violation of what is appropriate is tolerable. By knowing that, they can be relied on to recognize the occasions that outrage

normal standards of decency. If challenged to justify their reaction, they may not be able to do so. But that does not mean that their reaction is not justifiable (p. 443).

What should we do when we feel dumbfounded by the task of articulating support for our repugnances? For one thing, keep trying and don't give up. That's Midgley's (2000) recommendation here:

Feelings always incorporate thoughts—often ones that are not yet fully articulated... we need to look for the hidden partners. We have to articulate the thoughts that underlie emotional objections... look below the surface... (p. 8).

Step 5: Offer secondary arguments that do not directly appeal to disgust, but only so as to partially articulate our feelings.

Next, once you have appealed to widely shared feelings of repugnance and anticipated that unpacking the reasons underlying them will be difficult, don't stop there; now it's time to propose some specific, substantive reasons to think that our repugnance is properly directed (i.e. that which repulses us is worthy of repulsion). Importantly, the reasons you offer should not continue to Appeal to Disgust. Instead, you should be offering reasons that draw on moral vocabulary or concepts that ethicists draw on in more typical contexts. For example, you can suggest that the practice that repulses us is harmful, unjust, exploitative, a threat to rights or consent, or an assault on human dignity—if not directly then indirectly; if not obviously then subtly; if not immediately then eventually, at the bottom of a slippery slope; or if not in a single isolated case then when normalized and widespread throughout the population.

Of course, keep in mind that whatever reasons you provide are merely secondary. Our main, initial reason for suspecting the object of our repugnance is seriously wrong or ought to be prohibited is that it is so widely and intensely detested; we're only offering secondary reasons so as to be repugnance's press secretary and give some coherent defense of its (admittedly brute)

pronouncements. We're trying to be cooperative philosophical interlocutors and give opponents something substantive to respond to. Of course, they should be charitable with what we give them and recognize that it is just a preliminary argument—a philosophical down payment—that will require additional work to unpack.

For example, in Kass' widely discussed paper "The Wisdom of Repugnance," which advocates for a ban on cloning, he only argues for the wisdom of repugnance in the first 10 pages of the paper; the subsequent 15 provide several specific reasons to think cloning is seriously wrong and ought to be prohibited, even in liberal societies. The transition from the first 10 pages to the subsequent 15 is mediated by the following passage:

Yet our repugnance need not stand naked before the bar of reason. The wisdom of our horror at human cloning can be partially articulated, even if this is finally one of those instances about which the heart has its reasons that reason cannot entirely know (Kass, 1997: 689).

From there, Kass proposes a torrent of reasons why a reasonable person should be morally uneasy about cloning. For example: (1) It threatens to upend or confuse our "normal understandings" (p. 690) of family and kinship, biologically rooted cultural constructions (e.g. progenitor → mom, offspring → son, mate → spouse) that we depend on not only to provide people with a clear sense of social identity, but also to clearly distribute important "moral relations" (p. 690), duties, obligations, and responsibilities. (2) There are serious concerns that the very act of producing a clone threatens the clone's "consent" or "moral agency," insofar as they will know that they were not "an unbidden surprise, a gift, to the world" but rather "the designed result of someone's artful project" (p. 693). They will now be, willingly or not, compared to their 'parent' (person whose genetic material they were cloned from). And the clone could understandably complain about that (p. 694). (3) It "depersonalizes" and "commodifies" the process of reproduction (p. 691). That is, cloning changes our attitude toward the creation of

children, causing us to adopt a “technocratic mentality” (p. 696). After all, selection of genetic material to clone will inevitably be biased towards genes that are “desirable” (p. 698) (why clone Joe Schmoe when you could clone Mel Gibson or Michael Jordan?). From there, why not seek the *most* desirable, “perfect baby” (p. 700)? If we take that path, children will become the “instruments” of their creators (p. 696). We will turn “human nature” into a “technological project” influenced by the “subjective prejudices of the day” (p. 696). And this will be “inherently despotic” (p. 698), involving an extreme amount of control over others: who they are, why they were made, and what prospects they have in life.

Whether any of these arguments actually work isn’t the point. The point is that Kass attempts to unpack our widely shared, instinctive repugnance towards the idea of cloning with substantive arguments not directly related to disgust. I won’t clutter this subsection with any more specific examples, but note that other authors who appeal to disgust follow the same pattern as Kass: after appealing to disgust, they provide a series of secondary arguments, so as to partially articulate their (and others’) disgust. For example, Midgley (2000), who analyzes the ethics of using various bioengineering techniques (e.g. xenotransplantation) to create new (or blend two existing) species, spends the first 3-4 pages discussing “the role of feeling in morals” (p. 9), particularly disgust, but proceeds to build a substantive ethical critique of the practice in the following 6 pages, centered around the claim that it is hubristic and imprudent to “alter radically our whole conception of nature” (p. 12). Likewise, Sandel (2004), who discusses “designer babies,” “bionic athletes,” and “genetic engineering,” makes several remarks throughout the paper about how these practices induce “moral hesitation,” “moral unease,” or “moral vertigo” that is surprisingly difficult to unpack; nevertheless, after making such remarks, he attempts to unpack his unease by providing specific reasons.

Step 6: Suggest that the burden of argument falls on those who do not share in the ‘ordinary’ person’s disgust.

The final step of the Appeal to Disgust is to assert that, once the Appeal to Disgust is made, one’s opponent bears the burden of proof. That is, dialectically, appealing to disgust allows one to go on the defensive. A presumptive suspicion arises that one’s opponents are not in touch with the ‘normal’ feelings of the rest of us. It is *they* who have some explaining to do. In the meantime, the tentative winner (or frontrunner) of the debate is the majority; their widely shared revulsion should be adequate to curb our appetite for an argument, at least temporarily, until opponents can offer a substantive, compelling reason to think the majority’s disgust might be misguided.

For example, Kass (1997) explicitly writes,

generalized horror and revulsion are *prima facie* evidence of foulness and violation. The burden of moral argument must fall entirely on those who want to declare the widespread repugnances of humankind to be mere timidity or superstition (p. 689).

He later adds,

Perhaps [a ban on human cloning] will prove ineffective; perhaps it will eventually be shown to have been a mistake. But it would at least place the burden of practical proof where it belongs: on the proponents of this horror, requiring them to show very clearly what great social or medical good can be had only by the cloning of human beings (p. 704).

Kekes (1992) concurs: for some actions, moral disgust is “the normal reaction” whose “absence requires further explanation” (p. 103). If some people are not disgusted, “there must be some very special explanation of their failure in terms of repression, brutalization, pathological inattention, or the like” (p. 433).

5.3. Two testable claims

As a recap, the Appeal to Disgust has six steps:

(Step 1) Critique a practice whose (alleged) wrongness can't easily be spelled out in language or concepts typical of liberal/libertarian analyses (e.g. harm, rights, fairness, consent, equality).

(Step 2) Appeal to the widespread repugnances of humankind.

(Step 3) Acknowledge that disgust is fallible.

(Step 4) Anticipate that the reasons underlying our disgust reaction will be difficult to unpack.

(Step 5) Offer secondary arguments supporting your repugnance, so as to partially articulate it.

(Step 6) Suggest that those who do not share the 'normal' person's disgust bear the burden of argument.

The first thing to note about the Appeal to Disgust is that it wouldn't be exactly right to call it an argument. Arguments consist of premises that are logically related in such a way that, if they are true, the conclusion they support cannot fail to be true. If I give you a *good* argument that (e.g.) cloning is wrong—one with true premises and a valid structure—you are *compelled* to agree with me (or else you are being irrational, forming beliefs whimsically, regardless of whether they are evidentially supported). But the Appeal to Disgust isn't like that. If we're disagreeing about the ethics of cloning—you think it's permissible, but I think it's wrong—and I appeal to disgust, my appeal will not convince you to change your mind. What, then, is the point of appealing to disgust?

Here's my attempt at a charitable interpretation: the Appeal to Disgust is making more of procedural point than a substantive one. The idea is that if a practice, X, is widely considered deeply disgusting, this should be seen as *prima facie* evidence that X is seriously wrong. That is, in the absence of competing evidence—such as a strong substantive argument that X is permissible or even good, or the introduction of a defeater that undercuts our trust in our initial disgust reaction—we have some amount of justification for believing that X is seriously wrong.

This doesn't mean we have a *high* degree of justification; just enough to make it practically rational to act as if X is seriously wrong. In other words, if someone asks "What should be our public policy towards X?" our answer will be "Our public policy should reflect our tentative presumption that X is seriously wrong." In short, the Appeal to Disgust attempts to shift the burden of argument to one's opponents.

But that's not all. Also important is that there is something about the "repugnances of humankind" that make them difficult to articulate. Even in cases where they are undergirded by good reasons or good arguments, if someone asks us to explain those reasons, we may struggle to find any words. Perhaps giving an explanation or defense is possible, but we need time to think and cannot be rushed. This aspect of our disgust makes it especially important to place the burden of argument where it belongs. We can and should carry out a public debate about (e.g.) cloning. But the public should not be "put on the spot." Not every ethical question is easily penetrated by an analysis of the potential or likely bodily or psychic harms to individuals; not everything that is wrong is wrong *in that particular way*; not every moral value that an overlapping consensus of citizens in a liberal democracy tacitly hold can be easily made explicit. Therefore, if we put the public on the spot and demand that they either give a clear reason for their repugnance or get out of the way, we are putting them in a losing position. This will not make debate go better. It will rig the odds in favor of anyone who can ask rhetorical questions—"Who will cloning hurt? Whose rights will be infringed? Whose dignity will be threatened and how?"—at the cost of potentially vital moral values and norms. So, for public debate and deliberation to really "track the truth," and not be skewed towards the steady erosion of important—albeit subtle, nuanced, and complex—values and norms, the presumption must be on the side of the public's repugnances.

So, contrary to suggestions that the appeal is “poetic” or “mystical,” note that, actually, a number of testable claims fall out of it—claims that (1) can be shown to be true or false and (2) whose truth-value can be assessed using standard, knowledge-producing methodologies, be they philosophical or empirical. Example:

Inaccessible Premise: Our moral disgust reactions are often rooted in reasons that we do not have direct introspective access to.

The Appeal to Disgust proposes a *psychological* explanation for the dumbfounding effect: our disgust is undergirded by reasons, but we become dumbfounded when asked to state those reasons because we do not have direct introspective access to them. This claim is testable, insofar as there may be *alternative* psychological explanations for the dumbfounding effect (e.g. are we sure people’s disgust reactions are undergirded by *reasons*, or something like reasons, as opposed to being rooted in status quo bias, or being triggered by mere unfamiliarity?), and it may turn out that the evidence better supports one of these alternatives.

A second example:

Wisdom Premise: Widespread feelings of moral disgust are sometimes directed at things that really are seriously wrong.

In other words, not only do we often lack introspective access to the reasons undergirding our moral disgust reactions (Inaccessible Premise), but also those reasons can be *good* ones. If we *could* access them, we *would* endorse them.

Finally, note that the Inaccessible and Wisdom Premises, taken together, entail a third claim:

Burden-Shifting Conclusion: Those who do not share in the widespread repugnances of humankind bear the burden of argument.

In other words, if the Inaccessible and Wisdom Premises are correct, it follows that those who do not share the widespread repugnances of humankind bear the burden of argument. They cannot

challenge the public's disgust by simply pointing out that their opponents (e.g., those who want to ban cloning) are unable to provide an argument. This, after all, is not evidence no such argument could be made. Presumably, there is such an argument, though it will take time to spell out. In the meantime, those who advocate (e.g.) a cloning ban should (1) come up with a strong reason to think that cloning would have very morally desirable consequences or (2) introduce a defeater that undercuts our initial faith in the public's (often reliable) disgust.

5.4. *Four misconceptions resolved*

So, that's my interpretation of the Appeal to Disgust. The remaining sections of this chapter will argue that there is some truth in the Appeal. For now, let me briefly return to the four misconceptions about it with which this section began and explain what they get wrong.

Misconception A complained that the Appeal to Disgust implausibly treats one's own disgust as an infallible moral guide. This is incorrect. First, Step 2 of the Appeal is to cite the *widespread* repugnances of humankind, not one's own, idiosyncratic feelings. Second, Step 3 of the appeal is to explicitly acknowledge that repugnance, even when widespread, is fallible; it can be, and often is, misguided.

Misconception B complained that the Appeal to Disgust is unsuitable for discourse in a pluralistic democracy because it is ultimately rooted in private religious commitments. This, too, is incorrect. Step 2 of the appeal, which is critical, is to invoke the *widespread* repugnances of humankind, or the repugnances of the 'ordinary' citizen. Recall, Kass appeals to moral sentiments that (he believes) are shared by both "the man or woman in the street, and...the intellectuals...believers and atheists...humanists and scientists" (p. 686), "the overwhelming majority of our fellow Americans...all the human race, and of most...practicing scientists" (p. 701). In sum, the sentiment he appeals to is, allegedly, a *public* one.

Misconception C complained that the Appeal to Disgust is anti-argument. This is not the case. First, the appeal does not aim to persuade; it aims to place the burden of argument where it (allegedly) belongs (Step 6). This is a merely procedural recommendation and is consistent with vigorous public debate being carried out. Second, Step 5 of the appeal is to offer a series of substantive, albeit preliminary, arguments that opponents can engage with. As we saw, authors who deploy the Appeal actually spend *most* of their time on this step.

Finally, Misconception D complained that the Appeal to Disgust is radical and illiberal, insofar as it disregards the harm principle. But this not the case. First, the appeal does invoke values that citizens in a liberal democracy should care about (see Step 5). Second, as per Step 1, it is not so much that the appeal ignores that harm principle; rather, the practices that are typically being debated when the Appeal appears are practices that evoke deep moral unease and disgust *despite* not being obviously harmful or easily objected to in liberal/libertarian terms. That's the whole reason the Appeal to Disgust, with all six steps, is made.

6. Why Dumbfounding?

The previous section considered several complaints about the Appeal to Disgust that appear in the literature. I argued that these complaints fundamentally misunderstand the appeal and, thus, are not legitimate objections. I then suggested that a better way to push back against the appeal would be to scrutinize the testable claims that fall out of it, such as the following:

- (1) *Inaccessible Premise*: Our moral disgust reactions are often rooted in reasons that we do not have direct introspective access to.
- (2) *Wisdom Premise*: Widespread feelings of moral disgust are sometimes directed at things that really are seriously wrong.

These theses are worth scrutinizing because they underpin the Burden-Shifting Conclusion, the establishment of which is the Appeal's main goal:

(3) *Burden-Shifting Conclusion*: Those who do not share in the majority's disgust bear the burden of argument.

So, if either of the Premises can be refuted, the inference would be blocked, and the Appeal's goal would be frustrated.

On that note, the Appeal to Disgust may not be vulnerable to the misconceptions in the preceding section, but it is vulnerable, I believe, to the following charge: *the authors behind the Appeal to Disgust do not actually offer reasons to think that the Inaccessible Premise or the Wisdom Premise are true*. It's helpful that they fill out the formal structure of their argument, and I agree that it is valid; the Inaccessible and Wisdom Premises Theses entail the Burden-Shifting Conclusion. But valid arguments are useless without true premises. So the Inaccessible and Wisdom Premises stand in need of defense. The next three sections (6-8) aim to fill this lacuna.

As for the rest of this section, I aim to motivate and unpack the Inaccessible Premise. Note that the Premise makes an empirical claim; it offers a psychological explanation for the widely documented phenomenon of moral dumbfounding. Below, I start by critiquing a prominent psychological explanation for the dumbfounding effect found in the literature, which is due to Jonathan Haidt. Critiquing Haidt's theory will clear the way for my own.

6.1. *The Dumbfounding Question*

Whether the Inaccessible Premise is true is a question; however, whether moral dumbfounding is a prevalent phenomenon is not. Ample research in moral psychology shows that people often express confident and unhesitating moral disapproval of certain actions or practices—typically taboo or disgusting ones (more on this momentarily)—despite failing to provide reasons supporting their judgment. To elicit dumbfounding, researchers typically have

subjects read a vignette containing a moral violation, then have subjects report their moral intuitions. The best known vignette in this space is the Incest Story, which comes from Haidt et al. (2000):

Julie and Mark, who are brother and sister, are traveling together in France. They are both on summer vacation from college. One night they are staying alone in a cabin near the beach. They decide that it would be interesting and fun if they tried making love. At very least it would be a new experience for each of them. Julie was already taking birth control pills, but Mark uses a condom too, just to be safe. They both enjoy it, but they decide not to do it again. They keep that night as a special secret between them, which makes them feel even closer to each other. So what do you think about this? Was it wrong for them to have sex? (p. 15)

Subjects roundly condemn Julie and Mark's escapade. When asked why, they respond by either admitting they are dumbfounded or by exhibiting telltale signs of internal dumbfounding. For example, they may repeat unsupported claims, like "That's terrible!" or "It's just wrong!"; repeatedly say "I don't know"; display confusion via non-verbal gestures, like excessive laughing, face-touching, or saying "um"; or start putting forward an argument, only to abruptly give up upon realizing it leads nowhere.

Psychologically, what is going on here? Why is it so intuitive that Julie and Mark should not have sex but so hard to explain why? Call this the *Dumbfounding Question*.

6.2. Social Intuitionism

One immediate takeaway, which the researchers behind these studies themselves glean, is that some moral judgments *aren't caused by reasoning*. When someone reaches a judgment through a reasoning process, they can look back and report on how that judgment came about. Meanwhile, if someone can't report on how a judgment came about, this suggests they didn't reach their judgment via reasoning. If some judgments don't arise from reasoning, this raises the question: where are they coming from?

Jonathan Haidt's (2001) preferred explanation is what he calls the Social Intuitionist model of moral reasoning. On this model, the process that produces moral judgment in individuals is driven by "quick, automatic evaluations"—intuitions—not "ratiocination or reflection" (p. 814). What we typically think of as "moral reasoning"—"conscious, language-based" thinking, analysis, and formulation of arguments (p. 816)—is "rarely the direct cause of moral judgment" (p. 815). Instead, "moral emotions come first and directly influence moral intuitions" (p. 814), and intuitions, in turn, produce a judgment (one's official stance on a moral question). In short, the role of reasoning in the production of moral judgment is a secondary, and primarily social, one: "reason is the press-secretary of the intuitions, and can pretend to no other office than that of ex-post facto spin doctor" (Haidt et al., 2000: 3).

Haidt originally proposed the Social Intuitionist model as an alternative way to understand the production of moral judgment to the Rationalist models that had been historically dominant in the field of moral psychology. According to a basic Rationalist model, moral judgments are primarily caused by a reasoning process. For example, the influential rationalist Lawrence Kohlberg saw moral reasoning as "the conscious process of using ordinary moral language" (quoted in Haidt, 2001: 816). On this view, emotion or affect may have a small influence on judgment, but reason "ultimately makes the decisions" (p. 816). Haidt explains that Kohlberg came to this view by presenting subjects with moral dilemmas and observing how they resolved the conflicts. Kohlberg's best known dilemma is the Heinz Story:

In Europe, a woman was near death from a very bad disease, a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium for which a druggist was charging ten times what the drug cost him to make. The sick woman's husband, Heinz, went to everyone he knew to borrow the money, but he could only get together about half of what it cost. He told the druggist that his wife was dying, and asked him to sell it cheaper or let him pay later. But the druggist said, "No, I discovered the drug and I'm going to make money from it." So, Heinz got desperate and

broke into the man's store to steal the drug for his wife. Was there anything wrong with what he did? (quoted in Haidt et al., 2000)

Broadly, Kohlberg's work studied subjects' reactions to such scenarios. Kohlberg is best known for developing a theory that moral reasoning develops from childhood to adulthood through a series of six stages, where children begin as egoists but gradually learn to perspective-take, transforming their moral reasoning (Haidt, 2001: 816).

According to Haidt, Kohlberg's mistake was falling prey to the "psychologist's fallacy": "The tendency for psychologists to confuse a psychological phenomenon with the way they have chosen to study the phenomenon" (Haidt et al., 2000: 11). That is, Kohlberg believed that reasoning drives moral judgment as a result of his use of scenarios like the Heinz Story, which always involve "questions of rights and harms" (p. 11) or require subjects to "balance the interest of two people" (e.g. the wife versus the drugstore owner) (p. 5). In Haidt's view, such cases are "easy for the 'press-secretary' [i.e. reason] to discuss" (p. 5), "unique in encouraging analytical reasoning" (p. 8), and provide "fodder for post-hoc 'reasoning-why'" (p. 11). If Kohlberg had used a broader sample of moral judgment tasks—such as the Incest Story—"he might have come up with a different theory, one that gave greater prominence to moral emotions and...intuitions" (p. 11).

So, for Haidt, the Social Intuitionist model of moral reasoning corrects Kohlberg's mistake. It identifies affect and intuition as the primary driver of moral judgment, with reasoning playing a minor role. But we might wonder, if affect and intuition drive moral judgment, what influences are shaping affect and intuition? Here, Haidt answers that the most important influences are social. While an individual's reasonings may not have a strong causal influence on their intuitions, *other people's* reasonings, intuitions, and judgments do:

moral judgment is not just a single act that occurs in a single person's mind but is an ongoing process, often spread out over time and over multiple people. Reasons and arguments can circulate and affect people, even if individuals rarely engage in private moral reasoning for themselves (p. 829).

In short, people's moral psychologies are built to be subject to social influence.

6.3. *Combining Social Intuitionism and Moral Foundations Theory*

You might be thinking “Okay, so social influence shapes intuition, and intuition, in turn, drives moral judgment. That doesn't yet tell us much about how human moral psychology works; it just pushes the question back a step. If social influence molds moral judgment—moral judgment is an “ongoing process, often spread out over time and over multiple people”—what drives this larger process? Are collective moral judgments random, or are there observable motifs? If so, what are they, and what's causing *them*?” Haidt must have been asking himself the same question in the early 2000s because quickly after developing the basic building blocks of the Social Intuitionist model, he developed Moral Foundations Theory (MFT), a broader theory of why, despite significant differences across cultures, morality often exhibits common themes and similarities across populations. Of course, he didn't forsake the Social Intuitionist model; it plays a central role in MFT.

Here's the gist of MFT⁵⁴: Humans have innate sensitivities—“moral tastebuds”—to discrete moral domains (“moral foundations”) that cultures use as building blocks to construct moral systems. As children grow and mature, the surrounding culture shapes the development of these sensitivities, amplifying some and downplaying others. Strong candidates for moral foundations include (1) harm/care, (2) fairness/reciprocity, (3) authority/respect, (4)

⁵⁴ The following is based on Haidt (2012), Haidt and Graham (2007), and Haidt and Joseph (2004). For a more detailed discussion of MFT, see Chapter 2, Section 6; Chapter 3, Section 3.1.

loyalty/reciprocity, and (5) purity/sanctity. Each of these foundations has roots in biological evolution. For example, the harm/care tastebud originally evolved to help mothers detect signs of suffering in their offspring. The purity/sanctity tastebud originally evolved to protect humans from pathogens and parasites. However, over time, cultures “elaborated” these systems, making them sensitive to new concerns. For example, the harm/care tastebud expanded to all kin, which is why people today are sensitive to the suffering of their in-laws, coworkers, and total strangers. Likewise, the purity/sanctity domain was expanded from purely biological concerns to regulate “a great many bodily functions and practices,” such as “menstruation, eating, bathing, sex, and the handling of corpses” (Haidt and Joseph, 2004: 60). This ultimately led to the “profound moralization of the body and bodily activities” in many cultures (Haidt and Joseph, 2004: 60).

Where does the Social Intuitionist model fit into MFT? Simple: recall, the model holds that intuition is the primary cause of moral judgment, and social influence shapes intuition. MFT simply plugs in for the “social influence” part. Ultimately, the social intuitions that give rise to our judgments are rooted in five foundations, which themselves represent cultural elaborations of more primal emotional systems.

Now, returning to the Dumbfounding Question: psychologically, what explains the moral dumbfounding effect, in cases like the Incest Story? Why is it so intuitive that Julie and Mark should not have sex but so hard to explain why? Haidt’s answer is that, in general, our moral judgments—such as our disapproval of incest—are caused by our intuitions, and these intuitions arise from primal emotional systems that cultures have elaborated to boost in-group coordination and cooperation. The reason you can’t explain your disapproval of incest is because you didn’t reason to your view; purity is a basic moral foundation in human psychology, and your culture has programmed you to instinctively view incest as impure. Of course, you do not have

introspective access to this process, so if you “look inside” and try to report on what happened, you’ll come up empty-handed.

6.4. *Haidt’s hiccup*

In my view, there is something about Haidt’s answer to the Dumbfounding Question, which simply combines MFT and Social Intuitionism, that is less than satisfying. It does not clearly explain why taboo, or disgust-arousing, moral violations elicit dumbfounding more frequently than other types of violation. For example, in the classic dumbfounding study by Haidt et al. (2000), subjects were given moral judgment tasks involving both the Incest Story and the Heinz Story. What the study found was that the Incest Story elicited dumbfounding, while the Heinz Story did not. In response to the Heinz story, subjects (1) took longer to offer an evaluation (8.8 seconds versus 2.7); reported themselves that their judgment about Heinz was more based on reasoning than “gut-feeling”; admitted to being dumbfounded less often and exhibited fewer non-verbal signs of dumbfounding; on average, gave their first argument in response to Heinz *before* their first evaluation (whereas, for the Incest Story, evaluation preceded argument); and were more likely to change their mind in response to the interviewer’s playing devil’s advocate. In short, at least at a glance, the profile of subjects’ responses to Heinz appears to support a rationalist interpretation: reasoning is driving their judgment and “making the decisions.”

How does Haidt explain this discrepancy between Incest and Heinz? As I already noted, in the original study, the researchers suggest that Heinz is simply “fodder for post-hoc ‘reasoning why’” (p. 11), “easy for the ‘press-secretary’ [i.e. reason] to discuss” (p. 5). In other words, subjects aren’t reasoning about Heinz while emoting about Incest. They’re emoting about both, followed by post-hoc rationalizing; it’s just that, in the Heinz case, it’s easier for subjects to

“pass” as reasoners. Of course, this raises the question, what is it about Heinz that makes it easier for subjects to simulate (insincere!) reasoning?

In the original study, Haidt’s answer is that Heinz involves “questions of rights and harms” (p. 11) and “requires participants to balance the interests of two people” (p. 5). Of course, we might object to this explanation in the following way. If subjects ‘reason more’ because they must balance interests and weigh rights and harms—and this balancing process may change their mind—isn’t that just *moral reasoning*?

In another place, Haidt (2001) suggests that Heinz is more amenable to simulated reasoning than Incest because the devil’s-advocate-playing interviewer, “who challenges your judgment at every turn” (p. 820), introduces a strong social demand for justification and thus strongly motivates simulated reasoning; after all, people are generally motivated to be seen as being able to justify their judgments. Therefore, the interview induces an “unnaturally reasoned form of moral judgment,” which, if we’re not careful, may give us the “erroneous” impression that moral judgment is generally reasoning-driven (p. 820).⁵⁵ However, we might challenge this explanation by noting that the presence of a devil’s-advocate-playing interviewer is not a difference-maker between the two stories. When subjects are interviewed on the Incest Story, presumably this too introduces a social demand for justification and, thus, a motivator to simulate reasoning. Yet, despite this added motivation, many subjects struggle to find the words to defend their disapproval.

⁵⁵ Does introducing a social demand for an account of one’s moral judgments really cause the judgment-producing process to become “unnatural”? Maybe my intuitions are spoiled as a philosopher, but it seems to me, facing a social demand for justification is quite natural. In fact, even when I am thinking about ethics all by myself, with no one around, I often automatically imagine that fictional characters who *might* someday challenge my view on a given question, and these imagined interlocutors provide motivation to make sure that I’m really looking at the question the right way and forming my judgment in a way that makes sense.

In a third place, Haidt suggests that dumbfounding happens because mainstream American political culture is dominated by liberalism—a “discourse that heavily favors the ethics of autonomy”⁵⁶ (Haidt and Hersh, 2001: 195)—so those who morally disapprove of incest “will face pressure to justify their condemnation in the language of the ethics of autonomy” (p. 196). Of course, incest is a purity violation—it’s icky, not harmful or unfair, etc.—so justification “in the language of the ethics of autonomy” is, of course, not forthcoming. Haidt and Hersh (2001) tested this hypothesis by comparing rates of moral dumbfounding between self-identified liberal college students and self-identified conservative students in response to a vignette portraying homosexuality (a minority sexual practice disapproved of by conservatives at higher rates than liberals). The findings seemingly confirmed their hypothesis: “conservatives showed higher levels of ambivalence and dumbfounding because they were trying to make tradeoffs among” different moral ethics (p. 215), such as harm/care (“Homosexuality is harmless”), fairness/reciprocity (“Not allowing gay marriage is unfair”), and purity/sanctity (“Homosexuality is impure”). Meanwhile, liberal students effortlessly translated their intuitive (i.e. unreasoned) tolerance of homosexuality and gay marriage into terms compatible with the moral *lingua franca* of liberal societies, such as arguments based in harm/care or fairness/reciprocity language.

What the above explanation would seem to predict is that, in a non-liberal (or less liberal) culture than America, where moral rationalizations cast in the language of purity presumably remain socially acceptable, dumbfounding would be a relatively less prevalent phenomenon.

⁵⁶ The “ethics of autonomy” is a proto-moral foundation. This particular paper was written before Haidt introduced his Moral Foundations Theory (MFT). MFT was originally a development on Richard Shweder’s theory of the Three Ethics of Morality, the “three ethics” of which are autonomy, community, and divinity. In MFT, the “ethic of divinity” became the purity foundation, and the “ethic of autonomy” influenced the harm/care and fairness/reciprocity foundations.

Unfortunately, dumbfounding research on non-Western or non-liberal populations is scarce. However, at least one recent study, McHugh et al. (2023), is available. This study presented Heinz and Incest to samples from China, India, North Africa, and the Middle East and examined their responses. Contrary to Haidt's prediction, the study found that "moral dumbfounding could be evoked in non-Western samples" (p. 1056). Specifically,

For Heinz, people were significantly better at providing reasons, and significantly less likely to present as dumbfounded. . . for Incest, people were significantly less likely to provide reasons (p. 1052).

Ultimately, I think, the Dumbfounding Question remains unanswered, or not fully adequately answered, by Haidt's combination of MFT and Social Intuitionism. Specifically, Haidt fails to account for the difference in rates of dumbfounding between Heinz and Incest. Why is it so hard to elaborate reasons for our judgment of Julie and Mark, but not Heinz?

7. Vindicating the Inaccessible Premise

I think the key to answering the Dumbfounding Question is to go back to the primal emotional system in which our disapproval of Julie and Mark is rooted: *disgust*, the poison/parasite police. Haidt himself notes that insofar as our moral psychologies are fitted to be receptive to enculturation, this is because cultures found ways to "elaborate" existing psychological systems, attuning them to be sensitive to new, and socio-moral, concerns. A possibility that Haidt does not consider is that, when cultures began meddling with these pre-existing systems, they already had more-or-less flexible (i.e. more-or-less dumbfounding-prone) features, which were largely retained as byproducts when cultures co-opted them to regulate behavior or enforce norms. For example, moral anger and moral disgust are more and less flexible. Plausibly, this is because their non-moral analogs followed different evolutionary paths.

Several authors have suggested that anger is a more “flexible” emotion than disgust in the following senses: (1) it is relatively more sensitive to intentions (Russell and Giner-Sorolla, 2011a); (2) more likely to respond to mitigating circumstances (Russell and Giner-Sorolla, 2011b); (3) and more open to emotional reappraisal (Feinberg et al., 2014). Meanwhile, disgust is relatively inflexible along these dimensions. This contrast makes sense in the light of these systems’ original adaptive functions. Plausibly, anger has long been, or perhaps was always, a social emotion. Anger arouses the sympathetic nervous system, preparing the body for exertion, which is useful in fistcuffs (e.g. a mother protecting offspring or two males competing for sexual access to a female). But anger can also elicit counter-aggression. That is, A becoming angry at B is liable to elicit anger in B (Tybur, 2013). So, it’s important for anger to fire at the right moments. A *misfire* could be potentially costly (fighting always risks personal injury). When should anger fire? Knowing this would require a healthy amount of reading others’ minds, evaluating their likely intentions. For example, did the beta male who just bumped your shoulder while walking by do so accidentally or intentionally? If intentionally, you may need to fight him to make sure he knows who’s alpha (lest he challenge your reproductive supremacy). If accidentally, then presumably he still knows his place and there’s no need to fight and risk personal injury (i.e. in this case, anger would be a potentially costly misfire). In short, anger needs to be flexible.

Meanwhile, disgust is most effective when *inflexible*. The evolutionary logic of disgust starkly contrasts with that of anger. The costs of misfiring are low (e.g. you miss out on eating a safe food), while the costs of not firing at the right moment are high (e.g. you accidentally eat an unsafe food). If you get a little peckish and happen to be within arm’s reach of a piece of meat that looks fine but smells funky, the cost-benefit analysis is pretty clear: *better safe than sorry*

(Kelly, 2011; Bloom, 2004). Risking death isn't worth satiating minor hunger. Similar logic is observed in the Garcia effect mentioned in Chapter 2, Section 3.2: humans (and other animals) can automatically and unconsciously acquire new, enduring food aversions/disgust elicitors after a *single* nausea-inducing experience with a food item (e.g. shellfish). Of course, it could be that your current spell of nausea was caused by that rollercoaster ride you recently took, not the shellfish you ate. Maybe you could give shellfish another go at a future time? No: better safe than sorry. There are countless alternative nutritive foods available; ten false positives is better than one false negative.

7.1. *Disgust goes social*

So, disgust was inflexible before it ever became a moral emotion. But that's not all; it was likely an inflexible *social* emotion before it was a moral emotion. Recall that the Garcia effect/conditioned taste aversion is an individual-learning system that many animals have, not just humans (the effect, in fact, was first discovered in rats in the 1950s) (see Garcia et al., 1955). As discussed elsewhere in this dissertation, the development of *human* disgust (qua food-rejection system) was shaped by a unique set of pressures that ancestral humans faced upon becoming, at some point, meat-eating omnivores; experimenting with meat introduced a new wave of deadly pathogens, so it was critical to balance curiosity about new foods with caution (Chapter 2, Sections 3.2).

An additional, unique pressure on human disgust was that ancestral human groups, enabled by cumulative culture and cultural learning (more on this momentarily), expanded to non-native and highly varied habitats (e.g. from the tundra to the grassland to the tropical rainforest). What this entailed was that disgust-relevant information—"What should I eat/not eat? Touch/not touch?"—became localized (Tybur et al., 2013; see also Kelly, 2011: Chapter 3).

As a result, individual trial-and-error learning methods like the Garcia effect, and especially the role of genes in disgust acquisition, were downplayed (see Kelly, 2011: Chapter 3); far more important was *cultural learning*: share information, imitate the successful, accumulate knowledge over time, and transmit bodies of cultural information to upcoming generations. Empirical evidence corroborates this hypothesis. As discussed in Chapter 2, Section 5, the main influence on your disgust database (the elicitors you acquire) is the culture you grow up in. When you're born, you have no disgust triggers (not even feces); around age 4, you suddenly become a "neophobe," generally averse to unfamiliar foods. This developmental trajectory, disgust researchers propose, is designed to allow a crucial learning window for children to soak up local knowledge and eating customs (Tybur et al., 2013).

Additional support for the "human disgust became a social emotion early on" hypothesis is that experimental evidence shows disgust is associated with memory and attentional biases. Disgusting stimuli are far more memorable, and more attention-grabbing, than (e.g.) neutral or fear-inducing stimuli, even for infants (Schienle et al., 2021; Chapman, 2018; Chapman et al., 2013; Carver and Vaccero, 2007; Heath et al., 2001). Why is this? The most plausible explanation, I think, comes from gene-culture co-evolutionary theory (GCCE). According to GCCE, when humans became a cultural (i.e. knowledge-sharing and -transmitting) species, the growing size of bodies of cultural knowledge, combined with individuals' survival depending on their ability to avail themselves of this knowledge, led to a "ratchet effect" whereby a new set of selective pressures shaped human genes. Specifically, there was pressure for Mother Nature to select for humans who were *good at social learning*: when you encounter important information or have the chance to imitate a successful role model, you better (1) pay attention and (2) remember what you learned. The attentional and memory biases associated with disgust in

modern humans are most likely holdovers from this more primal condition: if your community is teaching you what to eat/not eat or touch/not touch, you need to pay attention and remember. If you don't, neither genes nor individual learning will compensate.

7.2. *Disgust goes dumb*

At this point, we have encountered two bits of disgust's evolutionary logic: *better safe than sorry* and *learn from others what's disgusting*. These two principles combined, quite unsurprisingly, produce a third: *don't question culture, just imitate others and adopt their disgust elicitors*. Surviving as an ancestral human would not have required *understanding* why a given item should not be eaten or touched; the important information is *that* it is disgusting. In other words, for cultural transmission of disgust elicitors to successfully keep everyone safe, it is necessary for there to be high-fidelity transmission. If ancient human, Hunter, decides to exercise a little American Individualism and test the safety of various foods for himself, he won't survive long (his genes will offer no guidance, and trial-and-error learning is futile, given the high stakes) (the story of Alexander Supertramp comes to mind). Meanwhile, if Fisher faithfully adopts the tastes, disgust elicitors, and food taboos of his community, and doesn't fret over silly questions like "What if one of every ten foods we've tabooed is actually safe and quite tasty?" he boosts his survival odds.

Actually, GCCE theorists note that an instinctive tendency to faithfully imitate, and engage in high-fidelity intergenerational transmission of, cultural practices is a distinctive feature of human nature, as a result of thousands of years of gene-culture co-evolution. For example, despite the irony that "to ape" is a synonym⁵⁷ for "to imitate," the ability of adult apes to imitate,

⁵⁷ The following passage from Mill's *On Liberty* comes to mind: "He who lets the world, or his own portion of it, choose his plan of life for him, has no need of any other faculty than the ape-like one of imitation. He who chooses his plan for himself, employs all his faculties" (p. 56).

for social learning purposes, is poor, paling in comparison to human children. For example, Henrich (2016) points to studies that measure the performance of adult chimpanzees versus human toddlers on a series of social learning tests. In the tests, subjects “observe a demonstrator use a hard-to-discover technique to obtain a desirable object, such as extracting some food out of a narrow tube,” then “use what they just saw demonstrated to help them obtain the desired object” (p. 14). These studies consistently find that human subjects, on average, outperform the chimps by a mile. For example, in one study, almost 90% of human toddlers passed the test and got the snack, compared to around 10% of adult chimps. Henrich emphasizes, “the averages...conceal the fact that most of the two-and-a-half-year-olds scored 100% on the test, while most of the apes scored 0%” (p. 14). These differences in ability reflect the broader contrast that apes don’t need social learning; meanwhile, for humans, it’s an essential survival skill.

Not only do humans imitate. Crucially, we “over-imitate.” Over-imitation refers to a phenomenon where people show high willingness to copy seemingly irrelevant actions in a process that a demonstrator used to successfully achieve a given goal (e.g. using simple tools to push, tap, pull, or poke an opaque box with doors and holes to obtain a snack). Henrich (2016) describes a specific experiment that has been “replicated with children, adults, and chimps” here:

In the experiment, participants first observe a model engage in a series of steps using a slender rod to access a reward in an “artificial fruit” [box containing toy or snack]. The fruit is a large opaque box with two entry points. The first entry point is sealed by bolts, which can be (a) *pushed* or (b) *dragged* out of the way—using the rod—to provide access to the tube. This tube, however, merely dead ends—it’s a decoy and is irrelevant to obtaining the reward. The second entry point is concealed by a doorway, which can be (a) *slid* or (b) *lifted*. The rod, which has a Velcro tip, can then be maneuvered down the tube to obtain the reward... (p. 108).

The results of such experiments consistently find that children and adults are more inclined than chimps to copy whatever the demonstrator did—even the irrelevant steps, even after they think the experiment is over, and even when they have been instructed *not* to copy irrelevant steps. You might think children over-imitate more than adults, given their increased daily dependence on adults, but it's the opposite; rates of over-imitation increase with age.

However, the really important finding comes not from this version of the experiment but a modified version that swaps the opaque box for a *transparent* one. This (hypothetically) allows subjects to “easily see” which steps in the procedure are relevant (e.g. use a tool to poke open a door that leads to a reward) versus irrelevant (e.g. use a tool to tap a tube unconnected to the reward) (p. 109). Interestingly, in this version, the chimps dropped the irrelevant steps once they had visual evidence of their uselessness. Meanwhile, children continued copying all the irrelevant steps, at the same rates as the previous experiment. In short, apes don't mindlessly ape; that's more of a human thing.

You might think, “Isn't over-imitation *irrational*—a bug in our psychology that even lowly chimps are not subject to?” Actually, viewed through the lens of cultural evolution, over-imitation—faithful imitation of successful models—is adaptive (see especially, Henrich, 2016: Chapter 7). As Henrich explains, as cultural bodies of knowledge (culinary practices, hunting techniques, group customs and norms, etc.) grew, confidence in one's individual learning became an increasingly unhelpful, if not risky, predisposition. This is because many cultural practices, through selection processes, advanced beyond the understanding of their adherents. In Henrich's verbiage, they became *causally opaque*.

An especially salient example in this context, discussed by Henrich (2016), involves the cooking of corn among indigenous South Americans (pp. 102-104). Corn naturally contains

vitamin B3, a vitamin you desperately need; vitamin B3 deficiency causes a horrible disease called pellagra, which causes “diarrhea, lesions, hair loss, tongue inflammation, insomnia, dementia, and then death” (p. 103). Unfortunately, this vitamin, though in corn, remains chemically bound in normal cooking processes (i.e. you can boil corn and eat it, but you won’t get any B3). To extract this precious vitamin, you must introduce an alkali to the cooking process, such as calcium hydroxide or potassium hydroxide—which is very unintuitive (Western scientists didn’t discover this until 1915). This isn’t a big deal if you have other sources of B3. But ancient South Americans didn’t; corn was their staple crop. Interestingly, they didn’t suffer from epidemics of pellagra; they learned—through a cultural selection process—to burn seashells or some wood types (creating necessary alkalis) and to mix them with corn in the cooking process, to release B3. Meanwhile, when corn was introduced to Europe after the discovery of the New World, pellagra spread in epidemic fashion (especially among poor populations whose diets relied heavily on cheap cornmeal and molasses), killing millions. Pellagra was believed (erroneously) by Western doctors to be an infectious disease until the early 1900s.

The upshot of this story is that cultural evolution, through a selection process, produced a solution to pellagra that took Western scientists centuries to replicate via conscious, causal reasoning. Of course, the indigenous South Americans availing themselves of this solution *had no idea what it was doing*⁵⁸; pellagra can take months or years to develop, so it would not have been easily observable that [eating non-alkalized corn] → [pellagra]. That is, the solution was causally opaque. Henrich (2016) is teeming with other fascinating cases of causal opacity in

⁵⁸ Henrich (2016) reports qualitative evidence that to this day, some indigenous tribes that alkalize corn via adding burnt seashells to the cooking process don’t know why they do this. When asked, they respond “It’s our custom” (see Chapter 7).

cultural practices (see especially Chapter 7). The lesson Henrich draws from such cases: *culture is smarter than you*. Ancestral humans who blindly trusted cultural teachings over their own wits enjoyed a competitive advantage in natural selection (i.e. Mother Nature selected for faithful cultural students). The “over” in over-imitation is a misnomer. If you are an ancestral human, you *should* be copying all the steps that a successful model uses as closely as possible, without trying to guess which ones are relevant versus irrelevant. Like the Western doctors who were stumped about pellagra for centuries, you are unlikely to “see” the wisdom inherent in what cultural models are doing. Faith in culture is adaptive; confidence in your individual learning abilities is not (just ask Alexander Supertramp).

Similar faithfulness to food taboos—adherence to them without understanding them or asking why—is also a common phenomenon in indigenous tribes, Henrich (2016) notes. For example, in indigenous Fijian groups, women avoid eating several large marine species (shark, eel, barracuda) while pregnant and breastfeeding. It turns out, scientifically, this is necessary to avoid leaving fetuses vulnerable to a toxin called ciguatera, and women, while pregnant, are generally less resistant to toxins. So, this taboo serves an adaptive function. Of course, the women who abide by it (and the surrounding culture that enforces it) don’t know this. Here’s what happened when Henrich and his wife, doing fieldwork, asked the locals why they enforced this particular taboo: (1) Many women “simply said they did not know and clearly thought it was an odd question” (p. 102); (2) others said it was “custom” (p. 102); (3) some did say violating the taboo would harm the fetus, but the alleged harm “greatly varied” (e.g. “The baby will grow shark scales,” “The baby will grow smelly joints”) (p. 102-103). In response to this, Henrich writes, “the answers here had the flavor of post-hoc rationalization: ‘Since I’m being asked for a

reason, there must be a reason, so I'll think one up now'" (p. 103). In other words, we could say that these women were *dumbfounded*; they endorsed the taboo but didn't know why.

Note: Henrich is not a "functionalist" about food taboos. He doesn't think that every taboo serves an obvious, biologically valuable function, such as avoiding a toxin. Indeed, most societies (including modern ones) have culturally specific food taboos/disgust elicitors that are (as far as anyone can tell) baseless, simply constituting rejection of nutritive foods. For example, many authors point out that most mammals are coded as disgusting (e.g. coyote, manatee, bear), minus a few (e.g. cow, pig, lamb), as well as most mammal parts (e.g. lips and eyeballs), despite these rejected categories being nutritive.

However, to suggest that individually functionless food taboos are "irrational" is probably the wrong way to look at things. Food taboos—even the unnecessary, functionless ones—arise from the deeper logic encoded in disgust's structure: (1) *better safe than sorry*; (2) *learn from others*; and (3) *have faith in culture, even if you are dumbfounded by its prescriptions and prohibitions*. In other words, seemingly "irrational" taboos are misfires caused by overly-cautious faith in culture (an adaptive trait).⁵⁹

7.3. *Disgust goes moral*

Disgust became "moral" when it was co-opted by culture to regulate behavior and enforce group norms. Tybur's (2009) theory that disgust was extended by culture to police anti-

⁵⁹ I felt compelled to make this point after reading this from Sripada and Stich (2004):

While there are few decisive arguments in the social sciences, the weight of evidence strongly suggests that, at least in the domain of food taboos, functionalist justification is often absent, and indeed many food taboos appear to be clearly maladaptive... Particularly difficult for a functionalist hypothesis to explain are cases in which closely related groups living in the same environment, and otherwise sharing similar social structures, exhibit widely different food taboos (p. 20).

Sripada and Stich's fail to acknowledge that the mechanisms that originate functionless food taboos, and sustain their existence over time, is faith in culture—an adaptive (or "rational") individual trait. So, it is an oversimplification to say "food taboos are often irrational." It is more accurate to say "a rational heuristic ('trust culture') sometimes causes irrational outcomes."

social violations (wrongs that impose *indirect* costs on the group) is a good one. Once disgust became moral, there was no limit on how many customs or norms culture could use disgust to regulate, as long as the proposed regulation was *psychologically tethered*. That is, emerging norms would have needed to “anchor on and extend [existing] instincts” or emotional systems (Henrich, 2020: 71), and potential anchors would be constrained by the prepared character of the relevant system.

For example, Henrich (2020) explains the emergence of incest taboos in the following way. First, “incest aversion” (biologically innate aversion to sex with close biological relatives, like a parent or sibling) was an available anchor. What adaptive norms (norms that give groups additional survival advantages) could culture tether to this anchor? Henrich (2020) proposes a plausible answer: expand members’ kinship networks (people who will show you altruism) by creating affines (non-blood relatives). Like other primates, humans exhibit kin altruism: innate altruistic motivations toward genetic relatives. Kin altruism can also be extended to affines (e.g. in-laws). The bigger your family, or kin network, the more allies and support you have (survival advantage). The problem is, if Hunter marries a first or second cousin, he expands his kinship network by fewer units than Fisher, who marries a third or fourth cousin. Hunter’s new in-laws will be his aunt and uncle, who already showed a degree of altruism to him, being close blood relatives. Meanwhile, Fisher’s new in-laws were more distant blood relatives who may not have reciprocated much altruism, but now will.

From culture’s perspective, there’s a clear lesson: harness humans’ innate incest aversion system to construct incest *taboos*, culturally designated, arbitrary rules about who is “related” to whom. This will stretch the innate system to include more cousins and relatives who are not actually blood-related (e.g. step-siblings), and *voilà*: you have a powerful, psychologically

tethered method to motivate members of your group to favor marriage partners/mates that will greatly expand their kinship networks. This gives them individually, and your group collectively, a distinct survival advantage. This, Henrich explains, is the ultimate explanation for why people are disgusted by sex between step-siblings. The genetic proximity detector (disgust) associated with innate incest aversion is going off. You could call this a “misfire”; there’s no actual risk of inbreeding, so the innate biological system is delivering a false positive. But that would be to privilege a biological perspective. From culture’s perspective, sex with a step-sibling should be tabooed, since step-siblings who marry will gain *precisely zero in-laws*. Your disgust is firing exactly when culture wants it to.

While there is a cultural logic behind incest taboos, individual members aren’t necessarily privy to this logic. Mother Nature selected for faithful cultural learners, not critical thinkers. If your culture signals that sex with a step-sibling is disgusting, you listen. As a result of this faithful adoption and transmission of sex and food taboos and disgust elicitors, dumbfounding follows. When ethnographers ask the !Kung people (indigenous nomadic hunter-gathers of South Africa) why they will not have sex with or marry first, second, or third cousins, nor anyone closer (e.g. a niece), *they don’t know*; they just say it is horrible and disgusting, “like” having sex with a sibling (some women refused to even discuss the idea) (Henrich, 2020: 75).

At this point, I think we have our answer to the Dumbfounding Question: Why do social scientists observe a close link between moral dumbfounding and moral disgust? Why is it so hard to explain our judgment of Julie and Mark, but not Heinz? Answer: by the time disgust was culturally co-opted to the moral domain, it already had an inflexible, false positive-prone, faithful/deferential character, and these features were retained as byproducts. This is because co-

optation involves no *modification* (significant changes to the basic structure of the system being co-opted), certainly not on so small a geologic time scale (the cultural co-optation of disgust happened relatively rapidly, compared to e.g. the two million years it took biological evolution to speciate Darwin's finches). So, moral dumbfounding in response to moral disgust elicitors happens because disgust has long been, perhaps always was, a dumbfounding emotion. Culture has its reasons for coding things as disgusting, but these are rarely (if ever) transmitted to you. So, if you introspect, you won't find them.

In conclusion, this section has vindicated the Inaccessible Premise:

Inaccessible Premise: Our moral disgust reactions are often rooted in reasons that we do not have direct introspective access to.

The "reasons" here are our cultures', not our own. Note that the Inaccessible Premise does *not* attribute moral wisdom to culture, or anything like that; perhaps cultures' reasons are often morally bad ones. The Premise is a purely descriptive or psychological explanation of moral dumbfounding.

8. Vindicating the Wisdom Premise

The preceding sections have, I believe, vindicated the Inaccessible Premise. That means we're halfway to establishing the Appeal to Disgust's Burden-Shifting Conclusion. All that's left is to vindicate the Wisdom Premise, which holds the following:

Wisdom Premise: Widespread feelings of moral disgust are sometimes undergirded by *good* reasons—i.e. reasons which, if we *could* access them introspectively, we *would* endorse.

Notice that this Premise makes a modest claim. It does not say that widespread feelings of moral disgust are always (or usually, or frequently) undergirded by good reasons, only that they sometimes are. This Premise is consistent with widespread moral disgust delivering false

positives, or false negatives, *most* of the time. The important point is, tools that are only sometimes correct can still be helpful. For example, suppose you have a metal detector that detects metal at a rate better than chance. It correctly detects metal 40% of the time, while the rest of the time it delivers a mix of false positives (beeps when it should not) and false negatives (fails to beep when it should). If you take this metal detector to the beach for an afternoon to hunt lost watches and old coins, you will miss lots of items—but you will still leave with a trunk full of successful finds. Upshot: a bad metal detector is more useful than no metal detector. Analogously, if widespread feelings of moral disgust are undergirded by good reasons at least sometimes, then, arguably, this gives them some evidential weight. Exactly how much weight is a separate question. The point is, if moral disgust is sometimes right, giving it no weight at all in our moral deliberations is a missed opportunity.

I develop an argument for the Wisdom Premise below. The gist of my argument goes like this. Cultural evolution often produces adaptive cultural artifacts—taboos, customs, norms, and collective attitudes that make societies more stable, successful, and better at coordinating or cooperating. While “adaptive” is not the same as “morally good” (that goes without saying), these things can sometimes overlap. If we had a list of every artifact that is distinctive of our culture, along with an explanation for why each particular artifact evolved, there would be many items on this list that would make us say “Huh, that’s actually a good idea—we should keep doing that.” Likewise, for many of the actions and practices that widely elicit moral disgust/have been tabooed, presumably they were tabooed because it was culturally adaptive, for whatever reason. If we had access to all of these reasons, surely we would endorse at least some of them. That is, we would see them and say “Huh, that practice really is bad—we should all keep being disgusted by it.”

8.1. What makes monogamy adaptive?

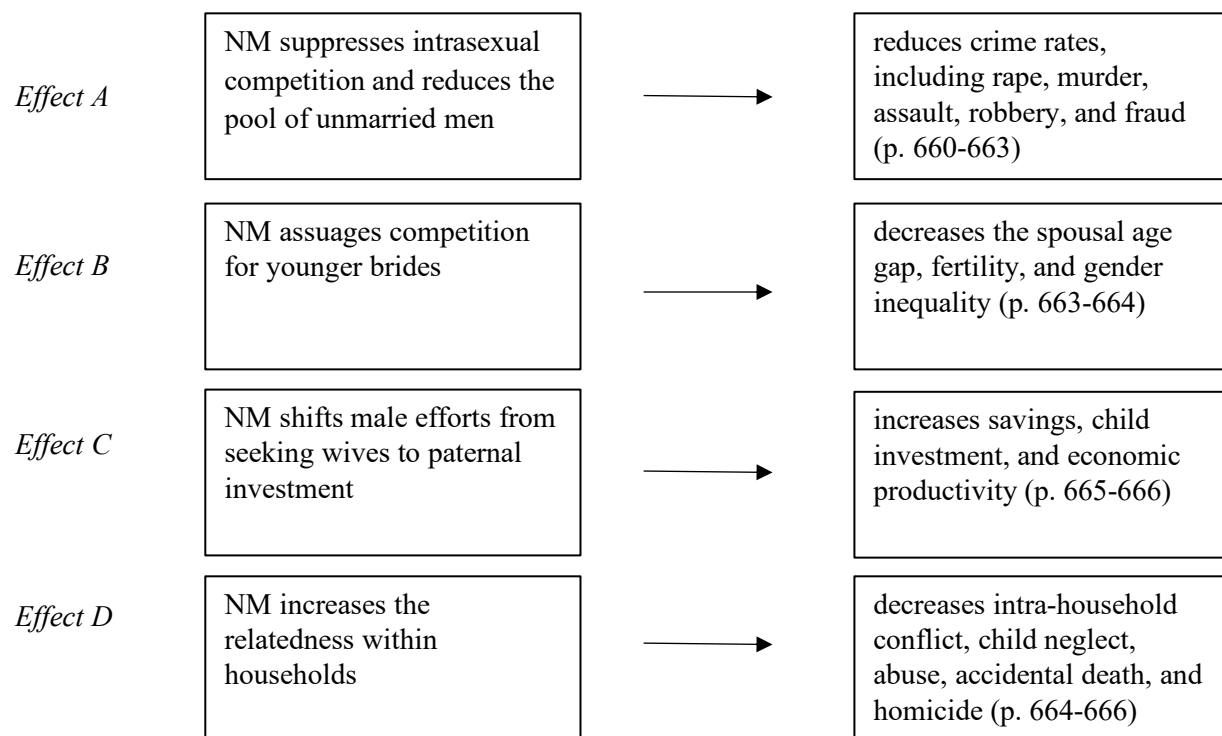
To illustrate my point that “adaptive” and “morally good” sometimes overlap, it will help to use an example: polygyny. By “polygyny” I mean “when one husband has more than one wife at the same time.” Polygyny is taboo in most highly developed societies (Henrich et al., 2012). For example, in the U.S., most Americans morally disapprove of it; the idea of men having multiple wives is icky. Moreover, the practice is illegal in every state. It is also illegal in Canada, Mexico, China, Japan, India, Nepal, and throughout Europe (Henrich et al., 2012). Most people alive today live in a society where monogamy is the norm, and those who deviate can face significant social and legal sanctions.

The fact that monogamy is increasingly the global norm, while polygyny is increasingly rare, is a bit of an anthropological puzzle. The vast majority of human societies in the anthropological record—one estimate is 85% (White et al., 1988)—have been polygynous. In such societies, taking multiple wives is almost always associated with status (high in wealth, nobility, or power) (Henrich et al., 2012: 657). Therefore, the people who are most likely to benefit from polygynous marriage norms are the same people in any given society who are most likely to influence law: high-status men. Polygyny may also be a better fit for humans’ evolved psychology; in ancestral contexts, the biologically fittest mating strategies for both sexes might have converged on polygyny (p. 658). So, why have lawmakers and law-influencers (mostly high-status men) in recent centuries started discouraging a practice that benefits themselves, promoting instead an alternative—monogamy—that is, arguably, less congruous with humans’ evolved psychology?

One plausible explanation—which has nothing to do with high-status men having undergone a sudden change of heart or egalitarian awakening—is that monogamous marriage is

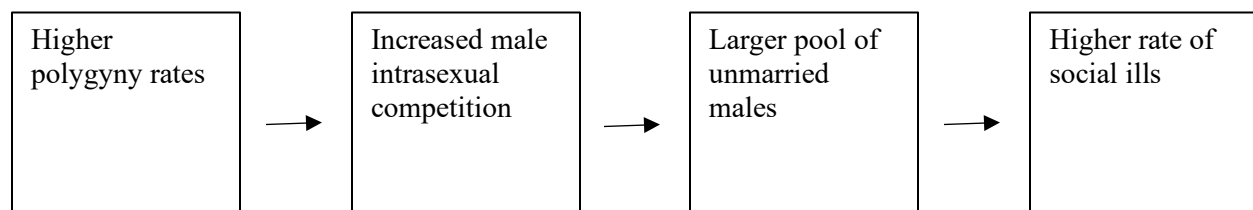
an adaptive cultural norm. Monogamous cultures tend to be more stable, more successful, and better equipped for inter-group competition, whether that competition takes the form of a violent conflict or a non-violent episode of cultural exchange. In short, cultural evolution—a selection process analogous in many ways to biological evolution—has favored monogamous societies.

What exactly makes monogamy more adaptive than polygyny? In a recent review, three leading figures in the field of gene-culture co-evolution—Joseph Henrich, Robert Boyd, and Peter Richerson (2012)—discuss and empirically test several hypotheses. They identify four main effects of Normative Monogamy (NM) (a condition in which monogamy is the cultural norm, while polygyny is taboo and subject to social and legal sanction):



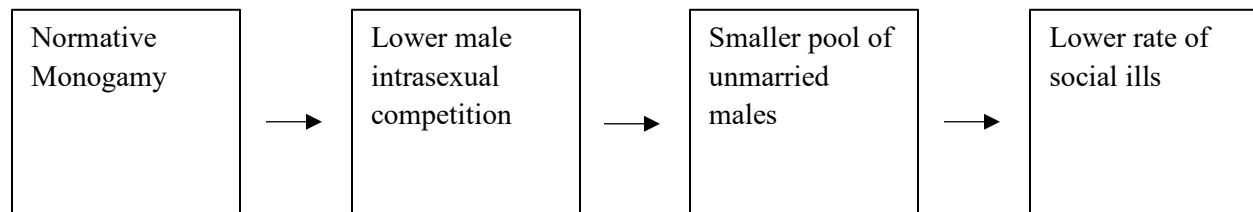
I don't want to take up valuable space going over each effect in detail, but it will be illustrative to give the flavor of at least one. Consider Effect A. This Effect holds that Normative Monogamy reduces crime rates by suppressing intrasexual competition. "Intrasexual competition" refers to competition between individuals of the same sex (male versus male,

female versus female) for access to mates or mating opportunities. High male competition means not every male is likely to find at least one mate; low competition means nearly all males can find a mate. Male competition typically centers around status (wealth, nobility, power). Thus, high competition increases pressure to gain status, and low competition decreases this pressure. When competition is high, this sparks “competitive motivations, steep temporal discounting and risk proneness” (p. 658). After all, the fitness difference between having one mate versus zero mates is large (a low-status male’s future isn’t so important to him when he’s on the brink of “fitness oblivion”) (p. 658). If you don’t have at least one mate, you should pursue “risky, high-stakes strategies” to change this (p. 658). Meanwhile, low male intrasexual competition reduces pressure to pursue risky, high-stakes strategies; in fact it incentivizes risk-aversion and long-term strategies like pair-bonding. Essentially, polygyny—even at quite low rates⁶⁰—increases male competition (i.e. decreases the chances of every male getting at least one mate) by draining the pool of available, unmarried women. This draining of the pool increases the need for males to vie for status (to avoid fitness oblivion). This, in turn, increases “men’s discounting of the future and risk-taking, resulting in more socially undesirable behaviours” (p. 660). In short, the causal chain looks like this:



⁶⁰ “To see why even a small increase in polygyny leads to a substantial increase in men without mates, imagine a society of 40 adults consisting of 20 males and 20 females. Suppose these 20 males vary from the unemployed high-school drop outs to CEOs. Let us assume that the 12 men with the highest status marry 12 of the 20 women in monogamous marriages. Then, the top five men (25% of the population) all take a second wife, and the top two (10%) take a third wife. Finally, the top guy takes a fourth wife. This means that of all marriages, 58 per cent are monogamous. Only men in the top 10 per cent of status married more than two women. The most wives that anyone has is four. While this degree of polygyny is not extreme in cross-cultural perspective, it creates a pool of 40 per cent of the male population who are shut out of the marriage” (Henrich et al., 2012: 661).

Conversely, monogamy reduces competition (increases the chances of nearly all males getting a mate); it “drains the pool of low-status unmarried men resulting in lower rates of social ills” (p. 661):



Henrich et al. (2012) confirm these theory-based predictions with ample empirical data. Polygyny is indeed associated with higher intrasexual competition (e.g. male competition among American Mormon communities between 1830 and 1890 declined dramatically as governmental forces suppressed polygyny) (p. 660). Longitudinal data shows that married men are less likely to commit crimes; crime rates for divorced *and widowed* men increase (p. 661). Marriage inversely correlates with “personal abuses” (e.g. binge-drinking or taking drugs) (p. 661). These individual impacts aggregate up to affect the whole society. Cross-national data show that percentage of unmarried men predicts crime rate. Intra-nation data show that more unmarried males means more crime (e.g. crime doubled in China between 1988 and 2004 during the implementation of the one-child policy, which skewed the male-female ratio toward males, producing more unmarried men) (p. 662). Regression analyses show that a 0.01 increase in sex ratio is associated with a 3% increase in crime, even after controlling for a number of potential confounds (e.g. unemployment, urbanization, inequality) (p. 662).

So, crime is increased by polygyny and decreased by monogamy. But why is reducing crime culturally adaptive, such that cultural evolution will favor crime-reducing norms (e.g. monogamy) over alternatives? Answer: among the most complex societies, “competitive success” is influenced by factors like “economic output, standing armies, innovation rates, trade,

division of labour and offspring quality” (p. 659). Reduced crime is a boon for such societies because it “spurs commerce, travel, and the free flow of ideas and innovations” and many more people can make long-term investments in “businesses, apprenticeships, and education” (p. 666).

8.2. *A tale of two supreme court cases*

So, monogamy is more adaptive than polygyny. Obviously, “adaptive” and “morally good” are distinct notions, and to say “Monogamy is adaptive, therefore it is good” would be to commit the naturalistic fallacy. Nevertheless, many of the features that make monogamy adaptive also happen to be morally praiseworthy: lower crime, lower gender inequality, higher child investment, and less intra-household conflict. These things promote, in a very basic sense, human flourishing, and flourishing is a valuable end for itself.

Note also that the goods associated with monogamy are *public* in nature; much like clean air, public parks, and a public education system, they benefit all equally. Meanwhile, the harms associated with polygyny are *public harms*; they reduce the welfare of everyone equally (if anything, the harms of polygyny disproportionately accrue to the least well-off). In short, if citizens in a liberal democracy were debating whether to legalize polygyny or not, highlighting the public harms associated with the practice, alongside the public benefits of monogamy, would be a legitimate argument; it would be understood by, and probably would be forceful to, an overlapping consensus of citizens who hold different comprehensive normative doctrines.

You might ask: “What does this have to do with the Appeal to Disgust? The argument against polygyny, which places the spotlight on its public harms, is a substantive argument that we can evaluate on its own merits. We don’t need to bolster this argument by invoking the ‘wisdom’ inherent in the ‘widespread repugnances of humankind’. Nor do we need to shift the

‘burden of argument’ to polygyny’s proponents in order to ‘improve the procedure’ of public debate.”

Well, maybe it would be weird to cite widespread repugnance towards polygyny as evidence that the practice is wrong *now*. But the social science on the public harms of polygyny has only become available in recent decades. Meanwhile, polygyny has been taboo in the West for centuries. I might add, the science is *surprising*. Who would have guessed that a 0.01 increase in sex ratio is associated with a 3% increase in crime? If you had asked me a few years ago, before I read Henrich et al. (2012), to explain why anti-polygyny statutes are a good idea, I likely would have exhibited signs of dumbfounding. Maybe I would have said “Polygyny is archaic!” But a practice’s being archaic is not a wrong-maker. Or maybe I would have said “Polygyny is sexist and exploitative!” But why is that the state’s business, provided those in polygynous marriages give their explicit and untainted consent to be in them? We can’t stop people from entering sexist relationships, especially if they feel a deep sense of religious obligation to do so. Proponents of decriminalizing polygyny also point out that sexual abuse, exploitation, and subjugation are hardly unique to plural marriages; monogamous marriages can involve those things (Turley, 2015).

I doubt my (hypothetical) dumbfounding is anomalous. Presumably, many people (who don’t know the social science) would not impress if asked to debate a proponent of decriminalizing polygyny. As a test case, consider the 1878 U.S. Supreme Court case *Reynolds v. United States*.⁶¹ This case considered the question of whether anti-bigamy statutes violate the Free Exercise clause of the First Amendment. The specific law being challenged was the Morrill Anti-Bigamy Act of 1862, signed by Abraham Lincoln, which banned bigamy (marrying more

⁶¹ <https://tile.loc.gov/storage-services/service/ll/usrep/usrep098/usrep098145/usrep098145.pdf>

than one person at a time) in U.S. territories. The petitioner in the case was George Reynolds, who was charged with bigamy in Utah, a territory at the time. Reynolds was a practicing Mormon (a secretary to Brigham Young, in fact) and argued that it was his religious duty to marry multiple times—a duty he had a First Amendment right to pursue. The Court unanimously rejected Reynold’s petition and upheld anti-bigamy statutes as constitutional.

However, if you look at *why* the Court reached its decision, you immediately notice that their arguments have the flavor of post-hoc rationalization. For example, one of their arguments is that polygamy has “always been odious” and “treated as an offense to society” in northwestern Europe and the U.S. (p. 164). But the fact that something has always been odious in a given society doesn’t necessarily mean that it is worthy of odium. A second argument is that Congress saw fit to make bigamy a crime “because of the evil consequences that were supposed to flow from plural marriages” (p. 168). Specifically, polygamy “feters the people in stationary despotism, while that principle cannot long exist in connection with monogamy” (p. 166). In other words, monogamy breeds committed democrats, says the Court, while polygamy makes people more susceptible to being dominated by tyrants. But what actual evidence is there for this claim? Their evidence is that this point was proposed by a relatively obscure, German-American jurist named Francis Lieber in his 1838 treatise *Manual of Political Ethics*. The Court adds, James Kent, a prominent law professor at the time, thinks that Lieber’s remark is “striking and profound” (p. 166). Huh? Since when does a law professor thinking a remark is striking and profound count as evidence that the remark is true? A third argument is that Congress “cannot interfere” with religious belief, but it can interfere with “practice” (p. 166). After all, Congress is obviously allowed to ban “human sacrifices” and widow-burning, even if someone *believes* these

practices to be a religious duty (p. 166). However, human sacrifice and widow-burning are quite different cases. They involve direct and severe *harm*. Meanwhile, who does polygamy harm?

All of the Court's arguments are weird and unsupported. I think the best explanation for this fact is that the Court was simply dumbfounded: their mind was made up that polygamy is bad, probably because they were instinctively repulsed by it (the practice is "odious"), but they didn't have justification for their judgment readily available to them because they didn't *reason* to their judgment. Their judgment was the result of enculturation. Culture trains us to be morally disgusted by certain things without bothering to explain *why* disgust is warranted (as long as you're a faithful cultural learner and replicator, culture doesn't care whether you understand what you're doing).

Compare the *Reynolds* case to a similar case out of Canada over a century later: *Reference re: Section 293 of the Criminal Code of Canada*.⁶² The Criminal Code of Canada has treated polygamy as a criminal offense since 1890. In 2011, the Code was challenged on the grounds that it violates the religious freedoms of Mormon fundamentalist groups guaranteed to them by the Canadian Charter of Rights and Freedoms. The challenge went up all the way to the Supreme Court of British Columbia. After all the relevant proceedings, the Court decided that criminal prohibitions on polygamy are constitutional. The Chief Justice wrote that "this case is essentially about harm," and the harm of polygamy can be "demonstrably justified in a free and democratic society." One of the harms of polygamy identified by the court was that "It generates a class of largely poor, unmarried men who are statistically predisposed to violence and other anti-social behaviour." If this argument sounds a lot like the argument presented in Henrich et al. (2012), that's because it is the same argument. Joseph Henrich was called as an expert witness to

⁶² <https://www.canlii.org/en/bc/bcsc/doc/2011/2011bcsc1588/2011bcsc1588.html>

the Court during its proceedings. When asked about the potential harms of polygamy, he noted that it creates a pool of low-status men, reduces male parental investment, and increases gender inequality.

The most fascinating part of the court proceedings, I think, was a more speculative portion of Henrich's testimony when he was asked to address whether legalizing polygamy in Canada would cause it to spread, and if so how quickly. Henrich's answer: "a non-trivial increase in the incidence of polygyny...is quite plausible if polygyny is legalized given what we know about male and female mating preferences." For starters, he explains that fertility is higher among polygynous families: "It's just robust. So these communities are going to...expand faster than monogamous communities." As for the majority population in Canada, it could be that "gender norms" are sufficiently entrenched that they would form a "shield or wall against the spreading of polygamy." However, "social changes occur quickly"; for example, most Americans in the 1950s would "probably be surprised" to learn that the U.S. would have a black president by 2008, Henrich says. Moreover, polygyny is a "ready response of our evolved psychology," an easy norm "for both males and females" to adjust to.

Combine this prepared psychology with the fact that (i) group norms are more likely to be influenced by high status people (e.g. celebrities), and (ii) high status men already exhibit a psychological tendency to polygyny, as evinced by their engaging in "serial monogamy" (I'm looking at you, Trump and Elon!), and a recipe for the spread of polygyny is there. Here's Henrich:

when high status people do things it's likely to transmit and spread through the social fabric... I can imagine this starting by actors and people of very high social status...taking a second wife or whatever, and then it would become legitimized and could potentially spread... I see serial monogamy amongst high status males as showing us that the psychology of polygyny is really there; right? So these are high status men who divorce the older wife in order to marry a younger wife, and in a polygynous society

they would just add a younger wife. It's a lot more convenient; you can still live with your children. So you could see where this thing might begin to ebb into mainstream culture... it's not as crazy as some people think.

In short, not only is polygamy bad for society; it also *fits the mold of an anti-social violation*: (1) it is liable to spread if not contained and (2) it imposes indirect costs on the group. Therefore, disgust is, arguably, a morally fitting response to polygamy, insofar as disgust is a fitting emotional response to contaminating social behaviors.

8.3. Conclusion

In conclusion, both the Supreme Court of the United States and the Supreme Court of British Columbia made the right decision on polygamy: it should be illegal, in light of the public benefits of normative monogamy and the public harms of polygamy. Interestingly, however, the U.S. Court got to the right decision *despite* not possessing the social science. Their arguments that polygamy represents a threat to the public were not at all persuasive. If you ask me, this was because they were morally dumbfounded as to why polygamy is wrong. Their moral disgust at polygamy was enculturated, not the result of a conscious reasoning process. I'm not defending poor legal procedure, but it is tempting to say that Americans got *lucky* that the Court was influenced by their intuitive disgust. In light of Henrich's ominous but plausible remarks in the Canadian case about the tendency of polygamy to spread if not suppressed via social and legal sanction, the legalization of polygamy in America could have had momentous consequences.

I think the preceding discussion is sufficient to demonstrate the Wisdom Premise. It is the case that widespread moral disgust among the public is sometimes undergirded by good reasons. That is, we may not have introspective access to those reasons, but if we *did*, we *would* endorse them. The two Supreme Court cases discussed above demonstrate this. If the *Reynolds* court had been able to access the reasons undergirding their disgust towards polygamy, they would have

endorsed them. We know this because the Canadian court *did* endorse them, once they were made available by modern social science—centuries after culture discovered them. This doesn't mean moral disgust is always "wise." It simply means that cultural evolution produces adaptive cultural artifacts. The reasons why it selected for those artifacts is often beyond our ken. But if we knew what culture's reasons were, there would be many times when we would *endorse* its reasons.

9. Conclusion

At this point, I believe the Appeal to Disgust is vindicated. Specifically, as I explained in Section 5.3, the Appeal encompasses two testable claims:

- (1) *Inaccessible Premise*: Our moral disgust reactions are often rooted in reasons that we do not have direct introspective access to.
- (2) *Wisdom Premise*: Widespread feelings of moral disgust are sometimes directed at things that really are seriously wrong.

Taken together, these Premises entail the following conclusion:

- (3) *Burden-Shifting Conclusion*: Those who do not share in the disgust of the majority bear the burden of argument.

The Inaccessible Premise was defended in Sections 6 and 7, and The Wisdom Premise was defended in Section 8. Therefore, on the basis of the arguments so far, we have reason to accept the Burden-Shifting Conclusion.

In informal terms, the reason the Inaccessible and Wisdom Premises entail the Burden-Shifting Conclusion is as follows. First, given the Wisdom Premise, when the public is widely morally disgusted by something, X, it might be the case that their disgust is properly directed, i.e. it might be the case that X is seriously wrong. This is because (i) our disgusts are shaped by enculturation, not reasoning; (ii) cultural evolution often has reasons for producing the taboos

that it does; and (iii) if we were privy to those reasons, we often would endorse them *qua* moral reasoners.

The problem is, moral disgust is, as a result of its evolutionary history, a dumbfounding emotion. Originally a biological safeguard, its inflexible nature—guided by the evolutionary dictum *better safe than sorry*—served ancient humans well. As it became a social emotion, its inflexibility was not only retained but likely accentuated under pressure from additional evolutionary dictums, such as *learn from others what is disgusting* and *don't question culture*. When disgust finally became moral, it was already dumbfounding-prone. Even today, we often lack direct introspective access to the ways our moral disgusts were shaped by culture, or why. Because of this, we may respond clumsily, at least initially, when challenged to provide justification for our disgusts. However, this is not evidence that our disgust is unfounded or rooted in mere prejudice. Our disgust reactions are rooted in culture, whose prescriptions and proscriptions are often wise but also causally opaque; if we examine them, we are unlikely to “see” the wisdom inherent in them (if there is any).

None of this means that we should be unwilling to question or debate the moral merits of longstanding customs and norms. Of course we should be willing. However, the widespread moral disgust underlying many customs will be difficult for ordinary people to articulate at the drop of a hat. Thus, it would be a mistake to frame the debate as if proponents of disgust-based customs and norms bear the burden of argument, and if they can't shoulder it, it follows that the relevant norms are outdated prejudices awaiting our enlightened erasure. This distribution of the argumentative burden would unfairly tip the debate in favor of opponents and quicken the destruction of potentially valuable norms—norms which, once the sentiments undergirding them are uprooted and demystified, there's no going back. Thus, opponents of disgust-based norms

should take up the burden of explaining why those norms should be changed. Either give a substantive reason to think that change would be in the public's interest, or provide an undercutting defeater that undermines our trust in our disgust. Just asking rhetorical questions like "Who would be harmed, or whose rights would be violated?" are too easy, and not helpful.

9.1. The upshot

What is the upshot of all this? One upshot, I think, relates to the story about the Scottish public petitions committee with which this dissertation opened. Recall, the committee swiftly rejected a petition to legalize incest. Their reason: considering the petition any further would serve "no public interest." The committee did not consider any of the substantive arguments in favor of the petition, such as that anti-incest laws (i) are discriminatory, (ii) violate individual rights, or (iii) are incompatible with the European Convention on Human Rights.

I think the committee's actions are justified—or at least, they were framing the issue in the correct way. Incest, even with consent and even with non-genetically related family members, is a longstanding (indeed, an ancient) taboo in most (if not all) societies, including Scotland. Presumably, most Scottish citizens would, like the committee, be instinctively repulsed by the proposition of legalizing incest. Perhaps their intuitive repulsion is misguided; perhaps legalizing incest would be merely distasteful to the majority population, but would not actually entail any public harms. But it is not the sole burden of proponents of the existing taboo to demonstrate that it is wise. The arguments presented to the committee by the petitioners had the flavor of unhelpful rhetorical questions, such as "Who would legalizing incest hurt?" or "Isn't it discriminatory to treat incestuous relationships differently than other types of relationships?"

Actually, these are *good questions*. What, indeed, are the potential public goods secured by treating incestuous relationships differently than non-incestuous ones in the law, and what

potential public harms may be prevented? It is difficult to analyze this question at the drop of a hat, but an analysis may turn up public reasons to criminalize incest, reasons that an overlapping consensus of citizens would endorse. Given the frequent wisdom of widespread repugnance, the advocates of legalizing incest behind the petition should themselves take up the burden of analyzing incest taboos—why they exist, what purposes they serve, what would happen if they were eroded. This is hard intellectual work—as it should be.

References

- Angyal A (1941) Disgust and related aversions. *Journal of Abnormal and Social Psychology* 36: 393-412.
- Aznar A, Tenenbaum HR, and Russell RS (2023) Is moral disgust socially learned? *Emotion* 23 (1): 289-301.
- Barnett R (2003) *Restoring the Lost Constitution: The Presumption of Liberty*. Princeton University Press.
- Bedke MS (2010) Intuitional epistemology in ethics. *Philosophy Compass* 5 (12): 1069-1083.
- Benokraitis N (2005) *Marriages and Families: Changes, Choices, and Constraints*. Pearson.
- Bensaï M, Rinck F, Schaal B, and Rouby C (2007) Verbal cues modulate hedonic perception of odors in 5-year-old children as well as in adults. *Chemical Senses* 32 (9): 855-862.
- Bergelson V (2013) Vice is nice but incest is best: The problem of a moral taboo. *Criminal Law and Philosophy* 7: 43-59.
- Boillat M, Challet L, Rossier D, Kan C, Carleton A, and Rodriguez I (2015) The vomeronasal system mediates sick conspecific avoidance. *Current Biology* 25 (2): 251-255.
- Blackburn S (2011) Review of *Yuck! The Nature and Moral Significance of Disgust*. *Times Higher Education* (August 4): 52.
- Bloom P (2004) *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human*. Basic Books.
- Bollard M (2022) Is there such a thing as genuinely moral disgust? *Review of Philosophy and Psychology* 13 (2): 501-522.
- Bostrom N (2005) In defense of posthuman dignity. *Bioethics* 19 (3): 202-214.
- Bostrom N and Roache R (2008) Ethical issues in human enhancement. In J Ryberg, T Petersen, and C Wolf (eds.) *New Waves in Applied Ethics*. Palgrave Macmillan.
- Brennan J and P (2015a) Markets without symbolic limits. *Ethics* 125 (4): 1053-1077.
- (2015b) If you may do it for free, you may do it for money. *Cato Unbound: A Journal of Debate* (November): <https://www.cato-unbound.org/issues/november-2015/what-are-limits-markets/>
- Brothánková P, Vodicka J, and Bazant J (2021) The influence of a child's age on the evaluation of smells and their hedonic assessment. *European Journal of Pediatrics* 180 (3): 745-749.
- Cahill M (2013) Abortion and disgust. *Harvard Civil Rights-Civil Liberties Law Review* 410: 409-456.
- Carmichael C (1995) Incest in the Bible. *Chicago-Kent Law Review* 71 (1): 123-148.
- Carroll N (1990) *The Philosophy of Horror: Or, Paradoxes of the Heart*. Routledge.
- Carver LJ and Vaccero BG (2007) 12-month-old infants allocate increased neural resources to stimuli associated with negative adult emotion. *Developmental Psychology* 43 (1): 54-69.
- Chapman HA (2018) Enhanced recall of disgusting relative to frightening photographs is not due to organisation. *Cogn Emot* 32 (6): 1220-1230.
- Chapman HA and Anderson KA (2013) Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological Bulletin* 139 (2): 300-327.

- Chapman HA, Johannes K, Poppenk JL, Moscovitch M, and Anderson AK (2013) Evidence for the differential salience of disgust and fear in episodic memory. *Journal of Experimental Social Psychology: General* 142 (4): 1100-1112.
- Chapman HA, Kim DA, Susskind JM, Anderson AK (2009) In bad taste: Evidence for the oral origins of moral disgust. *Science* 323 (5982): 1222-1226.
- Clark JA and Fessler DMT (2015) The role of disgust in norms, and of norms in disgust research: Why liberals shouldn't be morally disgusted by moral disgust. *Topoi* 34 (2): 483-498.
- Cohen CB (2002) Public policy and the sale of human organs. *Kennedy Institute of Ethics Journal* 12 (1): 47-64.
- Cosio MQ (1988) Soda pop vending machine injuries. *JAMA* 260 (18): 2697-2699.
- Courtland SD, Schmidtz D (2002) Liberalism. In EN Zalta (Ed.) *Stanford Encyclopedia of Philosophy*: <https://plato.stanford.edu/archives/spr2022/entries/liberalism/>.
- Curtis V, Aunger R, and Rabie T (2004) Evidence that disgust evolved to protect from risk of disease. *Proceedings: Biological Sciences* 271 (4): S131-S133.
- Curtis V and Biran A (2001) Dirt, disgust, and disease: Is hygiene in our genes? *Perspectives in Biology and Medicine* 44 (1): 17-31.
- Curtis V, de Barra M, and Aunger R (2011) Disgust as an adaptive system for disease avoidance behavior. *Philosophical Transactions: Biological Sciences* 366 (1563): 389-401.
- Danovitch J and Bloom P (2009) Children's extension of disgust to physical and moral events. *Emotion* 9 (1): 107-112.
- Danziger S, Levav J, and Avnaim-Pesso L (2011) Extraneous factors in judicial decisions. *PNAS* 108 (17): 6889-6892.
- Demetriou D (2013) There's some fetish in your ethics: A limited defense of purity reasoning in moral discourse. *Journal of Philosophical Research* 38: 377-404.
- Devlin P (1959) *The Enforcement of Morals*. Oxford University Press.
- Duff RA (2014) Towards a modest Legal Moralism. *Criminal Law and Philosophy* 8 (1): 217-235.
- Ekman P (1999) Basic emotions. In T Dalgeish and M Power (eds.) *Handbook of Cognition and Emotion* (pp. 45-60). John Wiley & Sons Ltd.
- (1972) Universals and cultural differences in facial expression of emotion. In J Cole (Ed.) *Nebraska Symposium on Motivation* 9: 207-282. University of Nebraska Press.
- Ekman P and Cordaro D (2011) What is meant by calling emotions basic. *Emotion Review* 3 (4): 364-370.
- Ekman P and Friesen WV (1971) Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology* 17 (2): 124-129.
- Fallon AE, Rozin P, and Pliner P (1984) The child's conception of food: The development of food rejections with special reference to disgust and contamination sensitivity. *Child Development* 55 (2): 566-575.

- Feinberg J (1985) *The Moral Limits of the Criminal Law, Volume 2: Offense to Others*. New York: Oxford University Press.
- (1984) *The Moral Limits of the Criminal Law, Volume 1: Harm to Others*. Oxford University Press.
- Feinberg M, Antonenko O, Willer R, Horberg EJ, and John OP (2014). Gut check: Reappraisal of disgust helps explain liberal–conservative differences on issues of purity. *Emotion* 14 (3): 513-521.
- Fessler D and Navarrete CD (2003) Meat is good to taboo: Dietary proscriptions as a product of the interaction of psychological mechanisms and social processes. *Journal of Cognition and Culture* 3 (1): 1-40.
- Fileva I (2021) You disgust me. Or do you? On the very idea of moral disgust. *Australasian Journal of Philosophy* 99 (1): 19-33.
- Fischer RW (2016) Disgust as heuristic. *Ethical Theory and Moral Practice* 19 (3): 679-693.
- Fiske AP (2020) The lexical fallacy in emotion research: Mistaking vernacular words for psychological entities. *Psychological Review* 127 (1): 95-113.
- Garcia J, Kimeldorf DJ, and Koelling RA (1955) Conditioned aversion to saccharin resulting from exposure to gamma radiation. *Science* 122 (3160): 157-158.
- George RP (1993) *Making Men Moral: Civil Liberties and Public Morality*. Clarendon Press.
- Gert J (2015) Disgust, moral disgust, and morality. *Journal of Moral Philosophy* 12 (1): 33-54.
- Giner-Sorolla R, Kupfer T, Sabo J (2018) What makes moral disgust special? An integrative functional review. *Advances in Experimental Social Psychology* 57: 223-289.
- Giubilini A (2016) What in the world is moral disgust? *Australasian Journal of Philosophy* 94 (2): 227-242.
- Green SP (2020) *Criminalizing Sex: A Unified Theory*. Oxford University Press.
- Haidt J (2012) *The Righteous Mind: Why Good People are Divided by Politics and Religion*. Pantheon.
- (2007) The new synthesis in moral psychology. *Science* 316: 998-1002.
- (2001) The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review* 108 (4): 814-834.
- Haidt J, Björklund F, Murphy S (2000) Moral dumbfounding: When intuition finds no reason. *Unpublished manuscript*, University of Virginia.
- Haidt J and Graham J (2007) When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research* 20 (1): 98-116.
- Haidt J, Hersh MA (2001) Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology* 31 (1): 191-221.
- Haidt J and Joseph C (2004) How innately prepared intuitions generate culturally variable virtues. *Daedalus* 133 (4): 55-66.
- Haidt J, Koller SH, Dias MG (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 CR, and Rozin P (1994) Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual Differences* 16 (5): 701-713.
- Haidt J, Rozin P, McCauley CR, and Imada S (1997) Body, psyche, and culture: The relationship between disgust and morality. *Psychology and Developing Societies* 9 (1): 107-131.
- Han BA, Kramer AM, and Drake JM (2016) Global patterns of zoonotic disease in mammals. *Trends in Parasitology* 32 (7): 565-577.
- Harris LT and Fiske S (2007) Social groups that elicit disgust are differentially processed in mPFC. *Soc Cogn Affect Neurosci* 2 (1): 45-51.
- (2006) Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science* 17 (10): 847-853.
- Hart HLA (1963) *Law, Liberty, and Morality*. Stanford University Press.
- Hauskeller M (2006) Moral disgust. *Ethical Perspectives* 13 (4): 571-602.
- Hayes C, Stevenson R, and Coltheart M (2007) Disgust and Huntington's disease. *Neuropsychologia* 45: 1135-1151.
- Heath C, Bell C, and Sternberg E (2001) Emotional selection in memes: The case of urban legends. *Journal of Personality and Social Psychology* 81 (6): 1028-1041.
- Henrich J (2020) *The WEIRD People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. Farrar, Straus and Giroux.
- (2016) *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton University Press.
- Henrich J and Henrich N (2010) The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proc Biol Sci* 277 (1701): 3715-3724.
- Henrich J, Boyd R, and Richerson P (2012) The puzzle of monogamous marriage. *Philos Trans R Soc Lond B Biol Sci* 367 (1589): 657-669.
- Hoefling A, Likowski KU, Deutsch R, Hafner M, Seibt B, Muhlberger A, Wyers P, and Strack F (2009) When hunger finds no fault with moldy corn: Food deprivation reduces food-related disgust. *Emotion* 9 (1): 50-58.
- Huemer M (2013) An ontological proof of moral realism. *Social Philosophy & Policy* 30: 259-279.
- (2007) Revisionist intuitionism. *Social Philosophy & Policy* 25 (1): 368-392.
- (2005). *Ethical Intuitionism*. Palgrave Macmillan.
- Huis A and Rumpold B (2023) Strategies to convince consumers to eat insects? A review. *Food Quality and Preference* 110: 1-14.
- Imada S, Yamada Y, and Haidt J (1993) The differences of Ken'o (disgust) experiences for Japanese and American students. *Studies in the Humanities and Sciences* 34 (1): 155-173.
- Jacobson D (2012) Moral dumbfounding and moral stupefaction. In M Timmons (Ed.) *Oxford Studies in Normative Ethics: Volume 2*: 289-316. Oxford University Press.
- Jarvis RM (2024) Idaho's anti-cannibalism statute: An assessment of a unique law (with some suggestions for its improvement). *Idaho Law Review* 60: 31-62.

- Jylkkä J, Härkönen J, and Jukka H (2021) Incidental disgust does not cause moral condemnation of neutral actions. *Cognition and Emotion* 35 (1): 96-109.
- Kahan D (2000) The progressive appropriation of disgust. In SA Brandes (Ed.) *The Passions of Law*. New York University Press.
- (1998) “The anatomy of disgust” in criminal law. *Michigan Law Review* 96 (6): 1621-1657.
- Kass L (1997) The wisdom of repugnance. *The New Republic* 224 (21): 30-39.
- (2003) Ageless bodies, happy souls: Biotechnology and the pursuit of perfection. *The New Atlantis* 1 (Spring 2003): 9-28.
- Kekes J (2005) Review: *Hiding From Humanity: Disgust, Shame, and the Law* by Martha Nussbaum. *Mind* 114 (454): 439-444.
- (2000) The enforcement of morality. *American Philosophical Quarterly* 37 (1): 23-35.
- (1992) Disgust and moral taboos. *Philosophy* 67 (262): 431-446.
- (1985) Moral conventionalism. *American Philosophical Quarterly* 22 (1): 37-46.
- Kelly D (2011) *Yuck! The Nature and Moral Significance of Disgust*. MIT Press.
- Kelly D and Morar N (2014) Against the yuck factor: On the ideal role of disgust in society. *Utilitas* 26 (2): 155-177.
- Knapp C (2003) De-moralizing disgustingness. *Philosophy and Phenomenological Research* 66 (2): 253-278.
- Koster JM, Hodgen JJ, Venegas MD, and Copeland TJ (2010) Is meat flavor a factor in hunters’ prey choice decisions? *Human Nature* 21 (3): 219-242.
- Kumar V (2017) Foul behavior. *Philosophers’ Imprint* 17: 1-17.
- Kupfer TR and Giner-Sorolla R (2017) Communicating moral motives: The social signaling function of disgust. *Social Psychological and Personality Science* 8 (6): 632-640.
- Kurth C (2021) Cultivating disgust: Prospects and moral implications. *Emotion Review* 13 (2): 101-112.
- Laing PAF, Davey CG, and Harrison BJ (2022) Influence of negative mood states on moral decision-making. *Psychiatry Research Communications* 2 (2): 1-8.
- Landmann H and Hess U (2017) Testing moral foundation theory: Are specific moral emotions elicited by specific moral transgressions? *Journal of Moral Education* 47 (1): 34-47.
- Landy JF and Goodwin GP (2015) Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspect Psychol Sci* 10 (4): 518-536.
- Levin Y (2007) The paradox of conservative bioethics. *The New Atlantis* 1 (Spring 2003): 53-65.
- Levy N (2003) What (if anything) is wrong with bestiality? *Journal of Social Philosophy* 34 (3): 444-456.
- Lockhart C, Lee CHJ, Sibley CG, and Osborne D (2023) The sanctity of life: The role of purity in attitudes towards abortion and euthanasia. *Int J Psychol* 58 (1): 16-29.
- Lu M (2013) Explaining the wrongness of cannibalism. *American Catholic Quarterly* 87 (3): 433-458.

- Machery & Stich (2022) The moral/conventional distinction. In EN Zalta (Ed.) *Stanford Encyclopedia of Philosophy*: <https://plato.stanford.edu/entries/moral-conventional/>
- Macklin R (2006) The new conservatives in bioethics: Who are they and what do they seek? *The Hastings Center Report* 36 (1): 34-43.
- May J (2018) The limits of appealing to disgust. In N Strohminger and V Kumar (eds.) *The Moral Psychology of Disgust*: 151-170. Rowman & Littlefield.
- McHugh C, McGann M, Igou ER, and Kinsella EL (2020) Reasons or rationalizations: The role of principles in the moral dumbfounding paradigm. *Journal of Behavioral Decision Making* 33 (3): 376-392.
- (2017) Searching for moral dumbfounding: Identifying measurable indicators of moral dumbfounding. *Collabra: Psychology* 3 (1): 23.
- McHugh C, Zhang R, Karnatak T, Lamba N, and Khokhlova O (2023) Just wrong? Or just WEIRD? Investigating the prevalence of moral dumbfounding in non-Western samples. *Memory & Cognition* 51 (1): 1043-1060.
- Midgley M (2000) Biotechnology and monstrosity: Why we should pay attention to the “yuk factor.” *The Hastings Center Report* 30 (5): 7-15.
- Mill JS (1963) *Collected Works of John Stuart Mill*. JM Robson (Ed.). University of Toronto Press.
- Mill JS ([1859] 1978) *On Liberty*. Hackett Publishing Co. Inc.
- Miller WI (1998) Sheep, joking, cloning and the uncanny. In M Nussbaum and CR Sunstein (Eds.) *Facts and Fantasies about Human Cloning*. WW Norton.
- (1997) *The Anatomy of Disgust*. Harvard University Press.
- Mills C (2014) A defence of repugnance. In A Akabayashi (Ed.) *The Future of Bioethics: International Dialogues*: 366-370. Oxford University Press.
- Moore MS (2017) Legal moralism revisited. *San Diego Law Review* 54: 441-463.
- Nabi R (2002) The theoretical versus the lay meaning of disgust: Implication for emotion research. *Cognition and Emotion* 16 (5): 695-703.
- Nagel T (1969) Sexual perversion. *The Journal of Philosophy* 66 (1): 5-17.
- Nemeroff C and Rozin P (1994) The contagion concept in adult thinking in the United States: Transmission of germs and of interpersonal influence. *Ethos* 22 (2): 158-186.
- (1989) “You are what you eat”: Applying the demand-free “impressions” technique to an unacknowledged belief. *Ethos* 17 (1): 50-69.
- Nussbaum M (2004a) Danger to human dignity: The revival of disgust and shame in the law. *The Chronical Review* 50 (48): B6.
- (2004b) *Hiding from Humanity: Disgust, Shame, and the Law*. Princeton University Press.
- (1999) Secret sewers of vice: Disgust, bodies, and the law. In SA Bandes (Ed.) *The Passions of Law*: 19-62. New York University Press.
- Oaten M, Stevenson RJ, and Case TI (2009) Disgust as a disease-avoidance mechanism. *Psychological Bulletin* 135 (2): 303-321.

- Oaten M, Stevenson RJ, Wagland P, Case TI, Repacholi BM (2014) Parent-child transmission of disgust and hand hygiene: The role of vocalizations, gestures and other parental responses. *The Psychological Record* 64 (4): 803-811.
- Olatunji BO and Sawchuk CN (2005) Disgust: Characteristic features, social manifestations, and clinical implications. *Journal of Social and Clinical Psychology* 24 (7): 932-962.
- Olsen OK, Pallesen S, Eid J (2010) The impact of partial sleep deprivation on moral reasoning in military officers. *Sleep* 33 (8): 1086-1090.
- Patrick D (2018) When souls shudder: A brief history of disgust and the law. In S Bandes, JL Madeira, K Temple, and EK White (Eds.) *Research Handbook on Law and Emotions*. Edward Elgar Publishing.
- Patrick D and Lieberman C (2018a) *Objection: Disgust, Morality, and the Law*. Oxford University Press.
- (2018b) How disgust becomes law. In N Strohminger and V Kumar (eds.) *The Moral Psychology of Disgust*. Rowman & Littlefield.
- Petersen TS (2010) New legal moralism: Some strengths and challenges. *Criminal Law and Philosophy* 4 (2): 215-232.
- Plakias A (2025) Adjust your disgust. *Aeon*. <https://aeon.co/essays/would-you-eat-a-cockroach-what-if-it-was-sterilised-why-not?src=longreads>
- (2018) The response model of disgust. *Synthese* 195 (12): 5453-5472.
- (2013) The good and the gross. *Ethical Theory and Moral Practice* 16 (2): 261-278.
- Rawls J (2001) *Justice as Fairness: A Restatement*. Columbia University Press.
- (1993) *Political Liberalism*. Columbia University Press.
- Repacholi BM (1998) Infants' use of attentional cues to identify the referent of another person's emotional expression. *Developmental Psychology* 34 (5): 1017-1025.
- Richerson P and Boyd R (1985) *Culture and the Evolutionary Process*. University of Chicago Press.
- Roache R and Clark S (2009) Bioconservatism, bioliberalism, and repugnance. *Monash Bioethics Review* 28 (1): 1-21.
- Roth A (2007) Repugnance as a constraint on markets. *Journal of Economic Perspectives* 21 (3): 37-58.
- Royzman E and Kurzban R (2011) Minding the metaphor: The elusive character of moral disgust. *Emotion Review* 3 (3): 269-271.
- Royzman E, Leeman RF, and Sabini J (2008) "You make me sick": Moral dyspepsia as a reaction to third-party incest. *Motivation and Emotion* 32 (2): 100-108.
- Rozin P and Fallon A (1987) A perspective on disgust. *Psychological Review* 94 (1): 23-41.
- Rozin P, Fallon A, and Augustoni-Ziskind (1985) The child's conception of food: The development of contamination sensitivity to "disgusting" substances. *Developmental Psychology* 21 (6): 1075-1079.
- Rozin P and Haidt J (2013) The domains of disgust and their origins: Contrasting biological and cultural evolutionary accounts. *Trends in Cognitive Sciences* 17 (8): 367-368.

- Rozin P, Haidt J, and McCauley CR (2008) Disgust. In M Lewis, JM Haviland-Jones, and LF Barrett (eds.) *Handbook of Emotions*, 3rd Ed: 757-776. Guilford Press.
- (1999) Disgust: The body and soul emotion. In T Dalgeish and M Power (eds.) *Handbook of Cognition and Emotion* (pp. 429-445). John Wiley & Sons Ltd.
- Rozin P, Lowery L, and 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, 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.
- Rozin P, Markwith M, and McCauley CR (1994) Sensitivity to indirect contacts with other persons: AIDS aversion as a composite of aversion to strangers, infection, moral taint, and misfortune. *Journal of Abnormal Psychology* 103 (3) 495-504.
- Rozin P, Millman L, and Nemeroff C (1986) Operation of the laws of sympathetic magic in disgust and other domains. *Journal of Personality and Social Psychology* 50 (4): 703-712.
- Russell PS and Giner-Sorolla R (2011a) Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science* 2 (1): 360-364.
- (2011b) Bodily moral disgust: what it is, how it is different from anger, and why it is an unreasoned emotion. *Psychol Bull* 139 (2): 328-51.
- (2011c) The dangers of disgust in the courtroom. *The Jury Expert* 23 (4): 1-14.
- Sandel M (2004) The case against perfection: What's wrong with designer children, bionic athletes, and genetic engineering. *The Atlantic* (April): <https://www.theatlantic.com/magazine/archive/2004/04/the-case-against-perfection/302927/>
- Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, and Cohen JD (2003) The neural basis of economic decision-making in the Ultimatum Game. *Science* 300 (5626): 1755-1758.
- Schienze A, Potthoff J, Schöenthaler E, and Schlintl C (2021) Disgust-related memory bias in children and adults. *Evo Psych* 19 (2): 1-10.
- Schmidt HJ and Beauchamp GK (1988) Adult-like odor preferences and aversions in three-year-old children. *Child Development* 59 (4): 1136-1143.
- Schnaich B, Lieberman D, and Kiehl K (2008) Infection, incest, and iniquity Investigating the neural correlates of disgust and morality. *Journal of Cognitive and Neuroscience* 20: 1529-1546.
- Schnall S, Haidt J, Clore GL, and Jordan AH (2008) Disgust as embodied moral judgment. *Pers Soc Psychol Bull* 34 (8): 1096–1109.
- Sebo J (2006) The ethics of incest. *Philosophy in the Contemporary World* 13 (1): 48-55.
- Shand J (2010) Abhorrence and justification. *Ethical Perspectives* 17 (4) 516-530.
- Shoemaker DW (2000) “Dirty words” and the Offense Principle. *Law and Philosophy* 19 (5): 545-584.

- Shweder RA, Much NC, Mahapatra M, and Park L (1997) The “big three” of morality (autonomy, community, divinity) and the “big three” explanations of suffering. In AM Brandt and P Rozin (eds.) *Morality and health*: 119–169. Taylor & Frances/Routledge.
- Siegal M and Share D (1990) Contamination sensitivity in young children. *Developmental Psychology* 26 (3): 455-458.
- Silbaugh KB, Posner RA (1996) *A Guide to America’s Sex Laws*. Chicago: University of Chicago Press.
- Singer P (2001) Heavy petting. *Prospect Magazine*.
- Skitka LJ, Wisneski DC, and Brandt MJ (2018) Attitude moralization: Probably not intuitive or rooted in perceptions of harm. *Current Directions in Psychological Science* 27 (1): 9-13.
- Solomon R (1974) Sexual paradigms. *The Journal of Philosophy* 71 (11): 336-345.
- Sripada CS and Stich SP (2004) Evolution, culture, and the irrationality of the emotions. In Evans D and Cruse P (eds.) *Emotion, Evolution, and Rationality*. Oxford University Press.
- Steiner JE, Glaser D, Hawilo ME, and Berridge KC (2001) Comparative expression of hedonic impact: affective reactions to taste by human infants and other primates. *Neurosci Biobehav Rev.* 25 (1): 53-74.
- Sternberg RJ (2003) A duplex theory of hate: Development and application to terrorism, massacres, and genocide. *Review of General Psychology* 7 (3): 299-328.
- Stevenson RJ, Oaten M, Case TI, Repacholi BM, and Wagland P (2010) Children’s response to adult disgust elicitors: development and acquisition. *Developmental Psychology* 46 (1): 165-177.
- Strohming N (2014) Disgust talked about. *Philosophy Compass* 9 (7): 478-493.
- Tetlock PE (2000) Coping with trade-offs: Psychological constraints and political implications. In A Lupia, MD McCubbins, SL Popkin (Eds.) *Elements of Reason: Cognition, Choice, and Bounds of Rationality*: 239-263.
- Tomasello M (1999) *The Cultural Origins of Human Cognition*. Harvard University Press.
- Tralau J (2013) Incest and liberal neutrality. *The Journal of Political Philosophy* 21 (1): 87-105.
- Truly J (2015) The loadstone rock: The role of harm in the criminalization of plural unions. *Emory Law Journal* 64 (6): 1905-1975.
- Turiel E (1983) *The Development of Social Knowledge: Morality and Convention*. Cambridge University Press.
- Turiel E, Hildebrandt C, and Wainryb C (1991) Judging social issues: Difficulties, inconsistencies, and consistencies. *Monographs of the Society for Research in Child Development* 56 (2): 1-103.
- Tybur J (2009) *Dissecting Disgust: An Investigation of the Validity of the Three Domain Disgust Scale*. Doctoral dissertation, University of New Mexico.
- Tybur JM, Çınar Ç, Karinen AK, and Perone P (2018) Why do people vary in disgust? *Philosophical Transactions, Series B, Biological Sciences* 373 (1751): 1-10.

- Tybur J, Lieberman D, Griskvicius V (2009) Microbes, mating, and morality. *J Pers Soc Psychol.* 97 (1): 103-122.
- Tybur JM, Lieberman D, Kurzban R, and DeScioli P (2013) Disgust: Evolved function and structure. *Psychological Review* 120 (1): 65-84.
- Wall S (2013) Enforcing morality. *Criminal Law and Philosophy* 7 (3): 455-471
- Weatherford RC (1995) Liberalism. In T Honderich (Ed.) *The Oxford Companion to Philosophy*: 485.
- Wheatley T and Haidt J (2005) Hypnotic disgust makes moral judgments more severe. *Psychol Sci* 16 (10): 780-784.
- White DR et al. (1988) Rethinking polygyny: co-wives, codes, and cultural systems. *Current Anthropology* 29 (4): 529-572.
- Wisnewski JJ (2004) A defense of cannibalism. *Public Affairs Quarterly* 18 (3): 265-272.
- Young L and Saxe R (2011) When ignorance is no excuse: Different roles for intent across moral domains. *Cognition* 120 (2): 202-214.
- Youngner SJ (2014) Why would it be wrong to create a human-animal chimera? In A Akabayashi (Ed.) *The Future of Bioethics: International Dialogues*: 358-365.