

The mark of the moral: Beyond the sentimentalist turn

Frank Hindriks & Hanno Sauer

To cite this article: Frank Hindriks & Hanno Sauer (2020) The mark of the moral: Beyond the sentimentalist turn, *Philosophical Psychology*, 33:4, 569-591, DOI: [10.1080/09515089.2020.1731444](https://doi.org/10.1080/09515089.2020.1731444)

To link to this article: <https://doi.org/10.1080/09515089.2020.1731444>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 10 Mar 2020.



Submit your article to this journal [↗](#)



Article views: 322



View related articles [↗](#)



View Crossmark data [↗](#)

The mark of the moral: Beyond the sentimentalist turn

Frank Hindriks^b and Hanno Sauer^a

^aFaculty of Philosophy, University of Groningen, Groningen, Netherlands; ^bDepartment of Philosophy and Religious Studies, Universiteit Utrecht, Utrecht, Netherlands

ABSTRACT

In light of recent empirical data, many psychologists and philosophers have turned away from rationalism about moral judgment and embraced sentimentalism. In the process, they have rejected the classical “moral signature” as a way of distinguishing moral from conventional norms in favor of a sentimentalist approach to carving out the moral domain. In this paper, we argue that this sentimentalist turn has been made prematurely. Although we agree that the experiments reveal that the classical approach is flawed, we propose to replace it with an alternative, according to which a norm is moral precisely if it is justifiable to all. This does not hold for most norms based on disgust or loyalty to a particular community. We accommodate the fact that such norms are not merely conventional by introducing a third domain, the domain of ethics. Our proposal reveals that (psychological) rationalism is still a viable option, as a lot of the experimental evidence that features emotions concerns the domain of ethics rather than morality.

ARTICLE HISTORY

Received 10 May 2017

Accepted 13 May 2019

KEYWORDS

Metaethics; moral psychology; sentimentalism; moral foundations; moral signature

1. Introduction

What distinguishes moral from non-moral matters? According to a venerable tradition in moral psychology, this question is best addressed by investigating how people make moral judgments. Famously, Lawrence Kohlberg (1969) and Eliot Turiel (1983) studied the judgments people form in response to scenarios featuring norm violations. Turiel in particular singled out norms concerning harm and fairness as comprising the moral domain and contrasted judgments about these norms with judgments about prudential norms or social conventions. He proposed that the following four features are distinctive of moral judgments: strong condemnation, authority-independence, universal application, and justification in terms of well-being or harm, and fairness. Responses to violations of conventional norms have none of these four features. Thus, on their view, these features form the mark of the moral. Together, they are sometimes referred to as “the moral signature” (Kelly, Stich, Haley, Eng, &

Fessler, 2007). According to what we refer to as “the classical conception,” the moral signature *is* the mark of the moral.

Recently, the issue of how to carve out the moral domain has been gaining increasing attention again. Some authors favor a thoroughly naturalistic approach, according to which moral judgment is a natural kind which constitutes a homeostatic property cluster and which can be empirically investigated (Kumar, 2015, 2016). Others couldn't disagree more, arguing that the main questions regarding the nature of moral judgment cannot be decided empirically at all because all the metaethical heavy-lifting is done by normative assumptions which are conceptually prior to experimental investigation (Clipsham, 2014). Yet others hold that morality is a radically dis-unified domain, such that any search for a distinctive mark of the moral must be doomed to fail (Sinnott-Armstrong & Wheatley, 2012). Finally, functionalists maintain that the moral domain should be identified as the set of norms that allow groups to reap cooperative benefits (Wong [2006] holds that this is a weak conceptual constraint on the subject matter of moral psychology and philosophy; for impressive empirical evidence, see Curry, 2016). Our paper can be seen as a contribution to this recent debate.

The classical conception has been challenged in the past three decades by psychological sentimentalists in what could be characterized as “a sentimentalist turn” in moral psychology.¹ Just like Kohlberg and Turiel, psychological sentimentalists study features of moral judgments. However, proponents of the sentimentalist turn have arrived at radically different conclusions concerning the nature and scope of the moral domain, and thus the mark of the moral. Richard Shweder, Much, Mahapatra, and Park (1997), Paul Rozin et al., (1997), and Jonathan Haidt, Koller, and Dias (1993), Haidt, Rozin, McCauley, and Imada (1997), for instance, studied *affective responses* to scenarios that feature various kinds of norm violations. In addition to anger and indignation, which are commonly associated with violations of harm and fairness norms, these affects included contempt and disgust. They took contempt to be associated with norms concerning the community, and disgust with norms concerning purity. Accordingly, they proposed that the moral domain extends beyond harm and fairness and encompasses these norms as well. To the extent that they recognize a mark of the moral at all, they take it to consist in the fact that moral judgments about norm violations are caused by moral emotions – which is what makes the turn *sentimentalist*. We will refer to this as *the sentimentalist conception* of the moral domain.

In what follows, we propose to go back and move forward at the same time by providing an argument for a “neo-classical conception” of the mark of the moral. We ask whether the sentimentalist revolution has given us good reasons for abandoning the classical conception of the mark of the moral. Furthermore, we investigate whether they have provided a convincing alternative. Our answers to both of these questions are negative.

Just like the classical conception, the neo-classical view gives pride of place to the way in which responses to norm violations are justified. In contrast to the classical conception, we argue that moral justification need not refer to particular values such as harm and fairness. Instead, a moral norm needs to be *justifiable to all*. Thus, according to the neo-classical conception, justifiability to all is the basis for singling out moral norms and for condemning moral norm violations. Our proposal is consistent with rationalism, which focuses on reasoning rather than emotion.

In [Section 2](#), we introduce the classical conception of the mark of the moral. In [Section 3](#), we discuss the sentimental turn and how research concerning disgust is taken to pose a challenge to the classical conception. [Section 4](#) is devoted to the sentimentalists' alternative and addresses the role that emotions play in the sentimentalist conception of the moral domain. In [Section 5](#), we introduce our neo-classical view, arguing that universal justifiability (or acceptability) should be seen as the mark of the moral.

2. The classical conception of the mark of the moral

The classical conception of the moral signature is due in particular to Lawrence Kohlberg and Eliot Turiel's ground-breaking psychological work on the ontogenesis of moral judgment.² Unlike most modern sentimentalists, who focus on adult moral judgment, they were interested first and foremost in how the capacity of moral judgment develops in children and adolescents.³ They understood this capacity as a cognitive ability to distinguish moral norms – and their violation – from other types of social norms.

The distinction between moral and merely conventional norms plays an especially important role for this project. For Kohlberg, moral development consists in a progression in people's capacity of moral judgment and reasoning through three levels (with two substages each). At the pre-conventional level, morality is framed in terms of externally given norms that are backed up by sanctions. The conventional level understands morality as loyalty to a peer group and the norms of a particular community. Finally, full-blown post-conventional morality bases moral norms upon abstract, universal principles such as the social contract, human rights, a principle of utility, or the categorical imperative.⁴ The universality or impartiality of this standpoint is a feature we will frequently return to below. In light of his focus on argumentation and the fact that he took his inspiration from Immanuel Kant, Kohlberg is best described as a rationalist rather than a sentimentalist.

Turiel is often described as a major critic of Kohlberg because his work shows that even very young children are able to distinguish between moral and conventional norms. This allegedly shows that there is no such thing as a pre-conventional stage, but this impression is, to a certain extent, mistaken. Turiel was interested in what distinguishes morality

and convention as social-cognitive domains. Kohlberg did not deny that morality is different from convention in any important way. The main objects of his studies were *the reasons* subjects would offer for why something is morally wrong or right. In this respect, subjects do appeal to pre-conventional, conventional, or post-conventional considerations to justify why something belongs to the moral domain, even if that domain is itself construed as non-conventional in nature. The moral domain is non-conventional, but people's reasoning about why a set of norms belongs to this domain progresses through Kohlbergian stages.

Turiel and his colleagues managed to determine on what basis subjects classify norm transgressions as moral or conventional on the basis of so-called *criterion judgments* and *justification categories* (Turiel, 1983, p. 52ff). The main idea behind this is that in the case of moral as opposed to non-moral norms – that is, acts involving lying, cheating, stealing, or hurting another individual rather than matters of etiquette or social mores – subjects will be more likely to view the wrongness of these acts as independent of the existence of an explicit rule or the will of an authority. They will also regard their wrongness as invariant across time, place, or culture, and they will judge transgressions of moral norms as more seriously wrong than conventional transgressions. Moreover, the justifications they proffer for why a given act was wrong are more likely to pertain to others' welfare (“Nobody wants to get their money taken because they like to have lunch”) or considerations of fairness (“I don't think it would be fair, [...] he didn't earn it”), rather than social coordination or punishment avoidance (Turiel, 1983, p. 67).

The moral signature thus delineates the moral domain on the basis of both formal and content-related criteria:

- (1) Strong condemnation
- (2) Authority-independence
- (3) Universal application
- (4) Justification in terms of well-being or harm, and fairness

On a strong interpretation, these four criteria are individually necessary and jointly sufficient conditions for moral judgment. The idea is that judgments about violations of moral norms exhibit all of these features, whereas judgments about violations of conventional norms exhibit none of them.⁵ On a weak interpretation, the features tend to occur together and form a nomological cluster (Kelly et al., 2007). This second interpretation allows for exceptions and fringe cases. There can be atypical moral judgments that lack one or another of the above features. With this proposal on the table, let us now turn to the sentimentalist turn.

3. The sentimentalist turn

Psychological sentimentalists want to rezone the moral territory. They reject the moral signature as the basis for distinguishing moral from non-moral judgments about norm violations. Thus, they challenge the moral–conventional distinction, or at least its usefulness for delineating the boundaries of the moral domain. Instead, they argue that what is distinctive of moral judgments about norm violations is that they are caused by *emotions* of a certain kind. We regard this as the defining thesis of psychological sentimentalism. In this section, we discuss how sentimentalists take recent research concerning disgust to challenge the moral–conventional distinction. Disgust research poses a challenge in particular to the fourth criterion, that of justification. In the next section, we discuss their alternative proposal, which consists in a taxonomy of moral *domains*, moral *emotions*, and the links between them as a way of distinguishing between moral and non-moral judgments.

Perhaps the main strand of research in the sentimentalists' case against the moral signature has been the empirical study of disgust reactions and taboo violations. Haidt's (2001) social intuitionist model, for instance, is based to a large extent on studies concerning subjects' judgments about harmless (or *allegedly* harmless; see Jacobson, 2012; Royzman, Kim, & Leeman, 2015) taboo violations, such as consensual incest or eating one's dead pet. Importantly, endorsement of these taboo norms seems to be independent of considerations of harm and fairness. If these judgments count as genuinely moral ones, then the notions of harm and fairness cannot be used to carve out the concept of moral judgment. This would mean that, as it stands, feature 4 is invalid as a criterion of morality.

Actually, the situation may be even more complicated. Recent findings by Landy (2016) suggest that Turiel's social domain theory, with its moral signature, and moral foundations theory are not in conflict with but, rather, orthogonal to each other. The moral signature is about the *wrong-making features of action*; moral foundations theory is about the *kinds of things* to which different people are inclined to attribute these features. This would mean that, at its core, moral foundations theory (and perhaps other forms of sentimentalism too) is a theory of individual differences regarding the issues which people are willing to moralize. We take this approach to support our proposal: at the end of the day, genuine morality is all about universally justifiable norms prohibiting serious harm. Feelings of disgust, respect for authority or loyalty to a community support ethical norms that attribute and frequently misattribute the moral signature to non-moral issues. The domain of ethics is thus populated by two things: first, teleological values that articulate one way among many of living a good life and are recognized as such and, second, improperly moralized but equally non-universally shareable values.

Our proposal is not meant to resolve once and for all which norms are moral and which are not. The reason for this is that the norms which are generally justifiable is a contested issue. For instance, some people may see intrinsic harm and, thus, general justifiability in actions that others may view as non-harmful offenses of mere taboos. Our main battle is a different one. If both parties to this debate invoke general justifiability to support their view, then we have made our main point. Apparently, people do indeed rely – explicitly or implicitly – on this criterion. The further step of taking a stance as to which of the two groups is mistaken requires a substantive normative benchmark. We do this to illustrate our framework. For instance, we think that norms of etiquette and disgust are good examples or norms that are sometimes moralized inappropriately. However, those who disagree with this are free to substitute in their own favorite examples. Doing so does not require a rejection of our proposed criterion as such. It merely amounts to the concession that the way in which people apply this criterion is contested and normatively charged. In this sense, our project is descriptive. It aims to supply a better conceptual framework for studying moral psychology.

Even so, there is some empirical support for regarding etiquette and disgust norms as non-moral norms, at least given the claim that the distinction turns on generic justifiability. People are less inclined to justify their responses to violations of disgust-based or purity norms as compared to harm or fairness norms. In other words, people feel less of a need to justify their responses of disgust as compared to their responses of anger. More specifically, they less frequently offer reasons concerning the causes and consequences of the relevant forms of transgressive behaviors (Russell & Giner-Sorolla, 2011). Instead, they sometimes cite the emotion of disgust as the basis for their response (Haidt et al., 1993). Thus, disgust sometimes functions as a conversation stopper. Disgust responses are relatively immune to arguments. In response, psychological sentimentalists have altogether abandoned justification as a criterion for the moral domain.⁶

Strikingly, people's judgments about disgust-based norms sometimes do exhibit one or more features of the moral signature, even though they do not concern harm or fairness. First, people sometimes regard violations of such norms as seriously wrong (feature 1: strong condemnation; Chapman & Anderson, 2014; Eskine, Kacinin, & Prinz, 2011; Haidt et al., 1993; Nichols, 2002; Schnall, Haidt, Clore, & Jordan, 2008).⁷ Second, they are sometimes deemed to be wrong independently of authority (feature 2: authority-independence; Nichols, 2002; Chapman & Anderson, 2014). Third, the norms are sometimes taken to be in force not only in the context at issue, but also everywhere else (feature 3: universal application; Haidt et al., 1993).

However, disgust-based norms do not consistently display all features of the moral signature. Because of this, psychological sentimentalists take this

research to pose a challenge to the moral signature as a means for delineating the moral domain. But why regard disgust-based judgments as moral at all? If this idea were rejected, the moral signature could perhaps be retained. Sentimentalists could respond to this challenge by claiming that research concerning moral disgust reveals that the moral–conventional distinction is not exhaustive (i.e., not bivalent). Disgust-based responses reveal that there are norms that are neither moral nor conventional in Turiel’s sense. These norms are not justified in terms of harm or fairness. Even so, they are not mere conventions such as driving on the left or right side of the road. As just discussed, norm-violations that trigger disgust could be seriously wrong, and the relevant norms might be authority-independent or universally applicable. The upshot is that they might have one, two, or even three features of the moral signature, but not all.⁸

As we discuss in [Section 5](#), the alternative is to relax the assumption that the distinction between moral and conventional is exhaustive. The thing to do, we propose, is to distinguish three domains and recognize an intermediate domain between the moral and the conventional: the ethical. Disgust-based normative judgments would then be part of this third domain and would be qualified as ethical rather than moral. Before defending this alternative, we discuss how sentimentalists who have rejected the moral signature have proposed an alternative mark of the moral: the taxonomic value–emotion approach.

4. The sentimentalist alternative: Domains, emotions, and moralization

Psychological sentimentalists argue that the moral domain extends beyond harm and fairness. According to the alternative they propose, the moral and the non-moral can be distinguished in terms of domains coupled with emotions (Haidt, 2012, p. 112ff.). The domain of *harm* is linked with the emotion of *care*, the domain of *fairness* is connected to the emotion of *anger* or *indignation*, *community* is coupled with *loyalty* and *contempt*, *authority* is backed up by *respect*, and, finally, *purity* is underwritten by feelings of *disgust*.⁹ It seems plausible to describe well-being, fairness, community, authority, and purity as values. Hence, the domain–emotion pairs which psychological sentimentalists use to distinguish the moral from the non-moral can be seen value–emotion pairs.

How useful is such a taxonomy of value–emotion pairs? There is one obvious advantage the sentimentalist approach has over the classical conception. The latter features only social-cognitive domains, whereas the former presents a richer picture of the psychological foundations of moral cognition, as it combines these social-cognitive domains with affective mechanisms that explain their motivational clout.

This taxonomic approach receives support from research concerning moralization, moral judgment, and moral reasoning. Consider moralization first. Roughly speaking, a type of action is moralized when negative responses come in the form of condemnation rather than an expression of dislike. Smoking is a famous example. It used to be seen as a matter of preference and is now, particularly in the United States, frequently viewed as a vice rather than a nuisance (Rozin & Singh, 1999). Demoralization occurs when morally neutral or acceptable acts (such as masturbation) cease to be the target of moral condemnation.¹⁰

Someone who moralizes an action condemns it as wrong and believes that it should not have been performed. Furthermore, she takes that kind of action to be governed by a norm that prohibits it. Norm-violation warrants condemnation of the action and, possibly, of the person, who can be seen to pose a threat to the values and way of life at issue (Keizer, Lindenberg, & Steg, 2008). Sentimentalists claim that such condemnation involves a suitable emotional response.

Psychological sentimentalists claim that the norms that govern moralized actions are moral norms. Although this claim has its attractions, the empirical research which has just been discussed cannot be used to support it. Violations of conventional norms are also condemned and might also elicit an emotional response. Hence, the fact that people have come to respond to a kind of action by condemning it rather than by expressing their dislike establishes only that it is now governed by a norm. It does not settle whether the norm is conventional or moral. At this point, psychological sentimentalists invoke some value–emotion pair that they regard as moral in support of their claim that the norm is a moral norm. However, they thereby assume what needs to be established: that certain value–emotion pairs are distinctive of morality. Of course, the examples that sentimentalists give for such pairings are not made up. Rather, they rely on laypeople’s judgments about which considerations should count as moral and which should not. If no one thought that community or relations of authority had any moral significance whatsoever, then sentimentalists would not presuppose that they do. Instead, their presupposition lies elsewhere: namely, in the assumption that we can glean, without much further ado, the mark of the moral from laypeople’s judgments. We merely wish to emphasize that an apparent strength of the sentimentalist approach comes at a price in the form of a list of moral values and emotions. As any such list would be controversial, this makes it somewhat arbitrary. The upshot is that it would be question-begging to claim that research concerning moralization supports the taxonomic approach. All that can be concluded is that the two are in line with each other.

Second, the sentimentalist turn seems to be supported by evidence concerning the emotional basis of moral judgment (Prinz, 2006; Sauer, 2012). A battery of studies seems to suggest that there is a tight link between (incidental) disgust

and moral judgment: inducing disgust in subjects makes their moral judgments more severe.¹¹

However, this way of supporting the sentimentalist turn is not as convincing as it initially appears. The significance of these findings is now being questioned, as the effect seems to be found only for some of the used vignettes, only in very specific subgroups of the population, and virtually never manages to change the polarity of people's moral judgments from right to wrong (May, 2014). A recent meta-analysis found that the influence of artificially induced disgust on moral judgment is small or nonexistent (Landy & Goodwin, 2015). Moreover, it is not at all obvious that the alleged influence of disgust on moral judgment generalizes to other moral domains and emotions.

A third line of support for the sentimentalist turn comes from research on the alleged frailty of moral reasoning. Haidt's famous "dumbfounding" studies seem to show that people make moral judgments intuitively and automatically. For the most part, moral reasoning consists of confabulatory post hoc rationalizations. This claim has attracted significant philosophical criticism (Hindriks, 2014, 2015; Jacobson, 2012; Railton, 2014; Sauer, 2017). We, however, wish to focus on the fact that the majority of subjects cannot offer reasons for their moral intuitions yet still refuses to abandon them, for it remains unclear why this fact should reveal much about the significance of reasoning for moral judgment. After all, a minority in the dumbfounding experiments *is* sensitive to a lack of arguments and does change their opinion (Hindriks, 2014). Moreover, there is simply no reason to assume *a priori* that emotional and intuitive responses generate only post hoc reasons (Kurth, 2015). This assumption becomes particularly implausible when processes of (affectively charged) intuitive judgment formation can be shown to incorporate and update on morally salient information (Nichols, Kumar, Lopez, Ayars, & Chan, 2016; Railton, 2017).¹² This seems to be more than enough for the rationalist.

How is it possible that proponents of the sentimentalist turn arrive at conclusions which are so far removed from those made by psychologists with more rationalist inclinations, who embrace the moral signature? To a certain extent, this appears to be an artifact of their respective research designs. A focus on complex hypothetical moral dilemmas that elude easy answers generates the type of moral reasoning that Kohlberg and his followers studied, whereas violations of affect-laden taboo norms that have already happened tend to trigger quick and emotionally charged intuitive responses. The idea that moral reasoning plays little or no role in how subjects arrive at their moral judgments may be due partly to a one-sided focus in recent moral psychology on third-person dilemmas at the expense of first-person dilemmas. First-person moral dilemmas tend to trigger moral reasoning. Investigating how people decide (first-personally) which action to perform as well as how people react (third-

personally) to how other people have acted provides for a richer and more balanced picture of moral judgment and reasoning (Monin, Pizarro, & Beer, 2007). Thus, there is ample reason to be critical of the sentimentalist turn and explore an alternative conception of the mark of the moral. Note that we do not intend to dismiss the research done by psychologists after the sentimentalist turn, nor do we reject the evidence they have produced as irrelevant to the study of morality – far from it. Instead, we accommodate it within our neo-classical view.

5. A neo-classical perspective: Justifiability to all

5.1. *Morality and ethics*

Earlier, we concluded that the moral–conventional distinction should not be seen as marking domains that are mutually exclusive and jointly exhaustive (bivalent). Some judgments about norm violations – in particular those that elicit disgust – fit in neither category. Sentimentalists have responded by broadening the domain of the moral and abandoning the moral signature. The alternative that we propose is based on a distinction between morality and ethics. Rather than broadening the conception of morality, we propose to introduce a third domain, that of ethics. According to the view we defend, norms are not only moral or conventional; they can also be ethical. Norms the violation of which elicits disgust are ethical norms rather than moral or conventional norms. This move is not merely semantic. We argue that there is a psychologically real distinction here, and we propose to use the terms ‘morality’ and ‘ethics’ to keep this distinction in place.

What is ethics, then, and how does it differ from morality? It is sometimes said that moral norms govern interpersonal behavior, and ethical norms are personal in that they concern what constitutes a good life. The problem with this way of drawing the contrast is that ethical norms are not personal but shared. In light of this, a better way of drawing the contrast is this: the domain of ethics is optional in the sense that people have some room for choice as to which norms and values to commit to, depending on how they conceive of the good life; moral norms are independent of such choices (Habermas, 1990; Williams, 1985).¹³ We argue that the distinction between morality and ethics has empirical bite, in that it marks a psychologically robust difference.¹⁴

In particular, the evidence supports this conception of the two domains in two closely related and mutually supporting ways. First, the five kinds of norms that the sentimentalists distinguish can be classified as moral or ethical in an intuitive manner. Norms concerning harm and fairness belong to the moral domain, just as the classical conception stipulates. More parochial norms concerning purity, community, and authority belong to

the domain of ethics. Second, the empirical findings concerning the moral signature fall into place. In particular, judgments that are neither moral in this sense nor conventional in Turiel's sense are part of the ethical domain. Recall that the disgust-based moral judgments discussed in the previous section are neither moral nor conventional in nature. According to our proposal, they are *ethical* judgments. Similarly, norms whose violation elicits an emotional response but that do not exhibit all features of the moral signature are ethical norms.

Thus, the theoretical distinction between morality and ethics turns out to be empirically robust. This supports our hypothesis that the evidence concerning norms of purity, community, and authority does not uniquely favor the sentimentalist alternative discussed. However, it leaves one thing to be desired. We have characterized the moral-ethical distinction (in part) in terms of different kinds of norms. This suggests that our approach is taxonomic, just like that of our rivals. In order to establish that it is more than that, we develop our rationalist alternative in terms of a particular kind of justification: a norm is moral exactly if it is justifiable to all, and an ethical norm is a non-conventional norm that does not have to meet such a high bar in order to be acceptable.

In [Section 5.2](#), we present evidence concerning the role of justification in relation to moral judgment. The rationalist proposal just mentioned is introduced more fully in [Section 5.3](#), where we present a neo-classical conception of morality, and in [Section 5.4](#), we discuss how it reflects on the ethical domain.

5.2. Evidence

Consider the following study done by Haidt et al. (1993): cross-cultural evidence shows that the most important factor predicting differences in subjects' moral judgments is not cultural background, but socioeconomic status (SES). When confronted with stories featuring harmless taboo violations (such as eating your dead pet dog or masturbating with a chicken carcass), judgments made by low SES subjects from both the US and Brazil exhibit features of the moral signature (in particular, universal application). Unlike low SES subjects from either background, high SES subjects tend to view the transgressive nature of taboo actions as a matter of social convention.

In an intriguing variation on the classic dumbfounding experiment, Paxton, Ungar, and Greene (2012) were able to show that subjects are significantly more likely to revise their moral intuitions when confronted with good counterarguments and enough time to reflect on them. Subjects themselves recognized the relevance of the fact that the disgusting aspect of an act (such as consensual incest between siblings) can be overridden by the fact that no one was harmed and that nobody's rights were breached. This

shows that something very much like the moral signature plays an important role in people's evaluations of actions.

Moreover, the phenomenon of moralization may also provide support for the view we are defending here. In a recent study, for instance, Gray, Schein, and Ward (2014) provide evidence for what they refer to as “dyadic completion” in moral judgment. When no obvious victim of an action that is intuitively deemed morally wrong can be found, subjects simply invent one. This suggests that there is considerable psychological pressure to frame moral issues in terms of harm, which shows that notions of harm play an especially central role in people's normative cognition. When people perceive an action as wrong, they automatically search for a victim. When no obvious victim can be found, and the perception of wrongness is strong enough, a suitable victim may be confabulated to make sense of the action's perceived wrongness. This, too, supports the idea that harm-based judgments occupy a special place in people's normative cognition. These judgments are based on norms that are justifiable to all.

These three considerations support the hypothesis that the moral signature is a robust feature of our psychology. Distinguishing between the moral domain and the domain of ethics allows us to preserve this idea, as well as accommodate the evidence concerning disgust-based judgments. What remains to be done is to provide a more principled and non-taxonomic way of drawing the distinction.

5.3. Morality and justification: The neo-classical view

Our proposal is to zoom in on justification as the mark of the moral. According to the classical view, moral judgments are justified in terms of well-being or harm and fairness. Thus, it designates particular values as moral values, which makes it somewhat arbitrary in the same way that the taxonomic approach is arbitrary. Instead, we propose that moral judgments are subject to public justification in ways that conventional and (mere) ethical judgments are not.

At the most general level, a norm is publicly justified exactly if it is acceptable to all – in the sense that, upon reflection, each comes to regard the norm as legitimate. Before commenting on more specific versions of this idea, we should note that what is justifiable to all moral agents may well be a moral matter itself, and if it is, our proposal is not reductive. The point to appreciate, however, is that it can be non-reductive without being viciously circular. It still makes a substantive contribution. In particular, it enriches our understanding of what is distinctive of the moral – namely, it has a special tie to interpersonal justification.

This notion of justifiability to all can be made more precise in a number of ways. One is the Kantian conception according to which moral norms apply

irrespective of the goals a particular agent happens to have. According to Kant's categorical imperative, a norm is impartially justified precisely if anyone can will that it becomes a universal law. Kantians take such norms to be acceptable for all. A second way to develop this proposal is the rule-utilitarian conception, according to which moral norms apply precisely if they tend to issue in the greatest happiness for the greatest number of people. For instance, the norm to do as promised is binding because, simply put, breaking promises often pains people. According to utilitarians, norms that maximize happiness are justifiable for all.¹⁵ Note that the goal of the paper is to propose a new criterion for distinguishing moral norms from other kinds of norms. As act-utilitarianism does not acknowledge any norms (other than maximizing overall utility), it falls outside the scope of our paper.

A third proposal is contractualist. Whereas Kantians and utilitarians offer a criterion for what makes a norm justifiable, contractualists regard the very fact that a norm can be justified to others as the distinguishing feature of moral norms (Darwall, 2006; Gauss, 1990; Scanlon, 1998).¹⁶ These are norms that concern what is justifiable to all. At the heart of contractualism lies the notion of mutual respect. To respect other people is to regard them as independent sources of reasons. This "horizontal" respect differs from the hierarchical respect for authority that features in Haidt's sentimentalism, as it concerns respect for someone as having moral standing independent of any role she occupies. In virtue of your moral standing as someone I have to respect, I need to take heed of your interests, needs, and values, and you of mine. At least some of these form inputs for a (hypothetical) contract that embodies moral norms. Thus, contractualism is social in a way in which Kantianism is not. According to contractualism, moral norms are public not only in the sense that they are accessible to all, but also in the sense that they need to be sensitive to any standpoint that bears on the issue at hand.

Irrespective of one's preferred normative theory, the proposal that public justification is the mark of the moral is useful for our purposes because it explains why norms might have some or all of the features of the moral signature. Norms that are justifiable to all – that are impartial, that maximize happiness, or that concern what we owe to each other – will typically concern harm or fairness. Transgressions of such norms are as such more serious than violations of conventional norms. Furthermore, norms that are subject to public justification will apply independently of particular authorities, and in many cases, such norms will apply universally. Note that this also accounts for the traditional way of drawing the moral–ethical distinction, as it explains why moral norms are not optional.

Interestingly, Scanlon allows for norms that are moral but that do not have universal application. More precisely, he argues that some norms concerning what we owe to each other are sensitive to the practices established in particular contexts (Scanlon 1998 pp. 338–342). For instance, the

norms of privacy that can be accepted – or, more precisely, not reasonably rejected – depend in part on the conception of dignity that prevails in a society and on what is regarded as embarrassing or shameful. As a consequence, the neo-classical conception is less rigid than the classical conception. People’s judgments about a norm need not have all four features of the moral signature in order for that norm to count as a moral norm. In particular, universal application is optional. Thus, instead of the moral signature, the mark of a moral norm is that it is subject to public justification.¹⁷

The classical conception invokes harm and fairness as the considerations in terms of which people justify moral judgments. The problem with this, however, is that by linking the moral domain to substantive considerations such as harm and fairness, this criterion can hardly help us move beyond the taxonomic approach: it always remains open for modern sentimentalists to argue for an expansion of the list of substantive considerations to characterize the moral domain. There is nothing about grounding the mark of the moral in substantive considerations such as harm or fairness that precludes this extension. This is what makes the idea of public justification so attractive, because such a non-substantive specification of the mark of the moral imposes constraints on which other substantive considerations can be added to the list of morally relevant notions – namely, only those which can figure in public justifications.

We have argued that psychological sentimentalists have overstated their case, and that there is more to be said in favor of rationalism than they allow. The notion of public justification can be regarded as a rationalist notion. In spite of this, the neo-classical view is congenial to certain versions of *philosophical* rather than psychological sentimentalism, which postulate a justificatory rather than merely a causal link between emotions and moral judgments. Antti Kauppinen, for instance, writes that “an attitude is moral only if it characteristically results from a process of simulating the non-moral reactive attitudes that any informed and unbiased participant would have in the circumstances of those affected by the action” (Kauppinen, 2010, p. 231). Here, too, the link between moral judgments and justification is stressed.¹⁸

Recently, Quigley (2015) has criticized the sentimentalist expansion of the moral domain, which extends the moral beyond considerations of harm and fairness to include norms pertaining to purity, community, and authority. His immanent critique supports our neo-classical perspective. Firstly, he notes that, importantly, the latter three considerations remain tied to notions of harm and welfare in that, for instance, loyalty to an in-group or respect for an authority figure largely consists in not wanting to harm and, instead, wanting to benefit one’s community and its leaders. Secondly, the most plausible evolutionary account of why community-, authority-, and purity norms emerged at all is that under the special circumstances of our

ancestral environment, such norms served as indirect proxies for the minimization of harm. Under these contingent conditions, these norms were indeed genuinely moral ones: since adhering to them happened to be the best way of securing the welfare of the tribe and its members, they could be publicly justified. Importantly, this same idea also debunks the relevance of such norms under modern conditions of large-scale societies with complex institutions, where the original rationale for these norms no longer obtains. Considerations of purity, authority, or community only enjoy moral relevance if and to the extent that they contribute to welfare and the avoidance of harm, and thus remain tied to public justifiability. If this “dependence thesis” is true, however, then the following question remains: if norms regarding purity, community, and authority are not moral, then how should they be classified? Evidently, they are not merely conventional. Here, the notion of a separate domain of ethical norms proves to be helpful.

5.4. *The ethical domain*

How, then, should we characterize the domain of ethics as opposed to morality? The domain of ethics concerns norms, values, and ideals that are “worthy of adoption and adherence” (Scanlon, 1998, p. 345). They can but need not be shared. In contrast to the moral domain, they are not directly concerned with what is justifiable to all. Many ethical norms and values are optional in the sense that people are at liberty to adopt them or not (Berlin, 1969). Whereas a value such as respect is mandatory, we can be pluralists in this sense about other values such as beauty, friendship, or work. Not everybody has to be actively concerned with art, for instance. The extent to which people value friendship or work is to a large extent up to them. For some, work is just that: work. Thus, we propose, ethical norms are norms that are based on ethical values that do not require public justification.

We can use this proposal to make the notion of moralization less arbitrary because it is independent from a taxonomic specification of moral emotions: to moralize a norm is to regard it as justifiable to all. Ethical norms include the disgust-based norms discussed above. Note that on our account, morality and ethics are not completely opposed: ethical norms will exhibit one or more of the other features of the moral signature, such as authority-independence, strong condemnation, or universal application.¹⁹ Ethics and morality are complementary, in the sense that ethical norms and values are, in principle, consistent with moral norms and values. Particular communities might, for instance, attribute particular significance to patriotic values, and there need not be anything wrong with doing so. It is not a moral issue, however: patriotic norms are not justifiable to all. Even so, they may be accepted by some. It is perfectly legitimate for communities that value patriotism to accept patriotic norms that condemn using a national

flag to clean a toilet. Even so, their normative force is different: rather than depending on general justifiability, they depend on particular valuations.

There can, however, be clashes between ethics and morality. In [Section 2](#), we mentioned the relation between disgust and ‘us’ versus ‘them’ divides. Many of these – such as ethnic divides or divides based on gender or sexual preference – cannot be publicly justified, or so we assume. The values that support disqualifying such outsiders are either abused, in the sense that they are misinterpreted, or they should be disqualified altogether. As the norms based on those values cannot be justified to all, it appears that they belong to the ethical domain: they are accepted by some. However, as these norms cannot be justified to anyone – in fact, alternative norms are justifiable to all – they do not even have a proper place there.²⁰

One might object to how we have classified particular norms. For instance, both disgust norms and norms of etiquette are sometimes seen as universally applicable (Nichols, 2004). This puts pressure on the proponent of the classical conception of the moral signature to regard them as moral norms. Note, however, that according to the neo-classical conception, universal applicability is not a defining characteristic of moral norms, even if many Kantians will say that it is. This leaves open the possibility that some ethical norms are universally applicable. Furthermore, it may be that the relevant norms are not justifiable to all. If they are not, then they are ethical norms rather than moral norms.

The main problem for any account which, like ours, aims to draw a distinction between the domains of the moral and the ethical on the basis of some feature x is how, and on what grounds, one can describe cases in which some people think that x applies to both domains when, in fact, it merely applies to one. The main example for this probably comes from cases where people come to believe that one or more of their ethical values do exhibit the mark of the moral, such as general justifiability.

Here, one option would be to note that our twofold claim is that (a) there is a distinction between the moral and the ethical, and (b) general justifiability is the best criterion by which to distinguish the former from the latter. Cases in which people mistakenly believe that some of their ethical norms are not generally justifiable when, in fact, they actually are generally justifiable don’t undermine either claim. What these cases show is that the way to apply the distinction between the two domains in practice is contested. It does not show that there is no such distinction or that the criterion we propose for drawing it is flawed. In fact, when people claim the feature of general justifiability for their “merely” ethical norms, they seem to implicitly accept this criterion as the right one. The remaining disagreement, then, is merely about *where* to draw line, not about the *whether* or the *how*.

The other salient option is to discount the responses Nichols encountered in his experiment, to propose that those who deem such norms to be universally

applicable are mistaken, and to point out that there are often good reasons for doing so. First, the rules of etiquette have often been seen as paradigmatic cases of conventional norms. Second, disgust is highly malleable: what used to be perfectly acceptable can become disgusting to people over time (Rozin & Singh, 1999). Third, disgust-norms display a relatively large degree of variation across individual and cultural contexts (Haidt et al., 1997; Rozin & Singh, 1999). Just as the first option, this second one entails that these norms are not moral norms. Rather, they are ethical or conventional norms.

We should emphasize at this point that we can accept that some disgust norms are moral norms. In Section 2, we suggested that taboo norms are, in all likelihood, not moral norms. Disgust can be and frequently has been used to create or maintain pernicious ‘us’ versus ‘them’ divides (Kelly, 2011; Nussbaum, 2009). Members of the out-group are then portrayed as repulsive vermin or some other kind of sticky, oozy, subhuman creature.²¹ This leaves open that other norms that trigger disgust when violated are moral norms. According to the neo-classical conception, this will be the case exactly if they are justifiable to all. Victor Kumar (2017) observes that disgust originally served to detect and sanction those who violated in-group norms but that this emotion has also become attuned to genuine moral wrongs, such as cheating, dishonesty, and exploitation. It may be that, in some cases, disgust has become a fitting response to violations of norms that are justifiable to all.

6. Conclusion

We have argued that many proponents of the sentimentalist turn have been too quick to discard the moral–conventional distinction. The way in which they themselves characterize the moral domain is problematic because it is based merely on a taxonomy of value–emotion pairs, which makes for a fairly shallow characterization. According to our proposal, the classical conception of the mark of the moral has to be reconceived. Rather than the four features of the moral signature – strong condemnation, authority-independence, universal application, and justification in terms of well-being or harm and fairness – we single out one feature: a norm is a moral norm exactly if it is publicly justifiable.

A lot of experimental research does suggest that the distinction between conventional and moral norms is not so clear-cut as it has been taken to be. Judgments about many norm violations exhibit at least one of the features of the moral signature. Rather than broadening the domain of morality, we have accommodated this by distinguishing a third domain in addition to the conventional and the moral domain: the ethical domain. This domain includes norms and values that are not publicly justifiable but may well exhibit one or more features of the moral signature.

This way of carving up norm domains is conducive to experimental research in a number of ways. For instance, it provides for a more adequate conceptual framework, and it serves to avoid drawing invalid conclusions, such as the claim that disgust-based norms are part of morality. Finally, an important motivation for categorizing domains in this way is that it facilitates condemning certain norms that appear to be ethical. This holds, for instance, for disqualifying certain outsiders in particular ways. Thus, by resolving certain confusions, our proposal also facilitates the formation of better normative judgments.

Notes

1. Earlier, the classical conception had already been challenged by authors such as Gilligan (1982), whose approach is in some ways congenial to what we describe as “the sentimentalist turn.” Others have referred to this as “the affective revolution” (Ditto, Pizarro, & Tannenbaum, 2009; Greene & Haidt, 2002; Haidt, 2007).
2. Piaget (1932) was an important predecessor here. See also Judith Smetana’s (1989) work.
3. Exceptions are nicely summarized in Bloom (2013).
4. In later work, Kohlberg included an intermediate stage 4 ½ to accommodate a transitional phase of relativism or nihilism often found between the conventional and post-conventional stage.
5. It was long thought that psychopathic individuals lacked mastery of the moral-conventional distinction (Blair, 1995). Recent research, however, shows that this is probably not the case (Aharoni, Sinnott-Armstrong, & Kiehl, 2012). The moral signature and the criteria on which it is based seem to be psychologically deeply entrenched.
6. To be sure, these differences concerning justification hold in particular for disgust responses toward food and sex. When disgust concerns betrayal, exploitation, or bigoted behavior, people often do offer reasons, and these reasons typically concern harm and rights (Russell & Giner-Sorolla, 2011). Furthermore, people sometimes infer (often symbolic) harm or rights violations in response to harmless taboo violations (Gutierrez & Giner-Sorolla, 2007). Finally, some people – particularly liberals – sometimes reappraise their disgust response, in which case they end up being more accepting of impure behavior and forming more deliberative judgments (Feinberg, Antonenko, Willer, Horberg, & John, 2014).
7. Disgust increases the extent to which people regard an action as wrong only when the disgust experience is subtle, or not so concrete and salient that people would easily notice an effect on their judgments (Schnall et al., 2008).
8. The relation between disgust and social polarization could be taken to count against regarding disgust as a moral emotion. Disgust can be – and frequently has been – used to create or maintain pernicious ‘us’ versus ‘them’ divides (Kelly, 2011; Nussbaum, 2009). Members of the out-group are then portrayed as repulsive vermin or some other kind of sticky, oozy, subhuman creature. This makes it easier, psychologically speaking, for perpetrators to morally disengage and victimize others (Bandura, 2016).
9. More recent versions of moral foundations theory include liberty–oppression as a sixth foundation (Haidt, 2012, p. 170ff).
10. Despite its somewhat negative connotation, when done properly, moralization can also constitute progress (Buchanan & Powell, 2016). Think, for instance, of certain

- sexist behaviors or cruelty toward animals that used to be considered acceptable and is more and more becoming part of moral discourse and practice.
11. However, see Landy and Goodwin (2015). In addition to their meta-analysis, it is important to mention that there are now numerous studies that have failed to support the amplifying effect of disgust on moral judgment: see Case, Oaten, and Stevenson (2012); David and Olatunji (2011); Johnson, Cheung, and Donnellan (2014); and Johnson et al. (2016).
 12. Thanks to an anonymous referee for pointing this out; however, see Gjesdal (2017) for some of the limits of this strategy.
 13. The ethical domain played a central role in the debate about liberalism and communitarianism (MacIntyre, 1984).
 14. A disadvantage of relying on this distinction is that these terms are often used interchangeably. However, alternatives such as ‘proto-moral,’ ‘pseudo-moral,’ or ‘semi-moral’ could be taken to mistakenly suggest that ethics has a derivative status.
 15. Although we are concerned with the question of which norms are moral norms, our answer could be extended so as to accommodate the utilitarian view, according to which an action is right exactly if it is justifiable to all. The act-utilitarian will claim that this is the case when the action maximizes overall happiness (on the assumption that equal consideration is given to the concerns of all individuals; Brink, 1993).
 16. Scanlon (1998, pp. 199–201) is concerned with principles about acceptable reasons that support regarding an action as right or wrong. Norms, or succinct statements about what is right or wrong, are simplifications of such principles that do not explicitly refer to reasons. Rather than being acceptable to all, Scanlon argues that moral norms that can be justified to others are such that no one can reasonably reject them.
 17. Scanlon’s contractualism can explain why people feel the pressure to justify their moral judgments in terms of harm, a pressure that – as discussed above – leads some to see harm where there is none. He observes that “we rarely, if ever, ‘see’ that an action is wrong without having some idea *why* it is wrong,” and he argues that “there is pressure to come up with an explanation or else withdraw the judgment if we cannot explain what our objection is” (Scanlon, 1998, p. 198). The contractualist test of public justification explains this in a natural way.
 18. Our approach is conservative in the theoretical or nonpolitical sense that we conserve what is valuable in the classical conception of the mark of the moral.
 19. We regard it as rather attractive to hold a sentimentalist view of ethical values (D’Arms & Jacobson, 2000). Gauss (1990) combines sentimentalism about such values with rationalism about moral principles.
 20. As Scanlon (1998, pp. 352–353) notes, people might mistakenly believe that ethical norms have the same authority as moral norms. For instance, some people might believe that it is wrong to clean toilets with national flags everywhere, which need not be the case in a less patriotic country.
 21. This makes it easier, psychologically speaking, for perpetrators to morally disengage and victimize others (Bandura, 2016).

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Frank Hindriks is Professor of Ethics, Social and Political Philosophy at the Faculty of Philosophy of the University of Groningen, where he is also director of the Centre for Philosophy, Politics and Economics. He has published about topics such as collective responsibility, social institutions, and moral responsibility in journals ranging from *Philosophical Studies* to the *Monist*, from *Philosophical Quarterly* to the *Journal of Institutional Economics*. He is a co-founder of the *Journal of Social Ontology* and the *International Social Ontology Society*.

Hanno Sauer is Assistant Professor of philosophy at Utrecht University. His main research interests are in metaethics and moral psychology. He is the PI of the project *The Enemy of the Good. Towards a Theory of Moral Progress* (ERC Stg 851043, PROGRESS). His most recent book is *Debunking Arguments in Ethics*, Cambridge University Press 2018.

References

- Aharoni, E., Sinnott-Armstrong, W., & Kiehl, K. A. (2012). Can psychopathic offenders discern moral wrongs? A new look at the moral/conventional distinction. *Journal of Abnormal Psychology, 121*(2), 484–497.
- Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves*. New York: Worth Publishers.
- Berlin, I. (1969). *Four essays on liberty*. Oxford: Oxford University Press.
- Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition, 57*, 1–29
- Bloom, P. (2013). *Just babies: The origins of good and evil*. New York: Crown.
- Brink, D. (1993). The separateness of persons, distributive norms, and moral theory. In C. Morris & R. Frey (Eds.), *Value, welfare, and morality*. CUP.
- Buchanan, A., & Powell, R. (2016). Toward a naturalistic Theory of moral progress. *Ethics, 126*(4), 983–1014.
- Case, T. I., Oaten, M. J., & Stevenson, R. J. (2012). Disgust and moral judgment. In *Emotions, imagination, and moral reasoning* (pp. 195–218).
- Chapman, H. A., & Anderson, A. K. (2014). Trait physical disgust is related to moral judgments outside of the purity domain. *Emotion, 14*(2), 341. doi:10.1037/a0035120
- Clipsham, P. (2014). Does empirical moral psychology rest on a mistake? *Philosophical Studies, 170*(2), 215–233.
- Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. In T. K. Shackelford & R. D. Hansen (Eds.), *The evolution of morality* (pp. 27–51). Springer.
- D'Arms, J., & Jacobson, D. (2000). Sentiment and value. *Ethics, 110*(4), 722–748.
- Darwall, S. L. (2006). *The second-person standpoint: Morality, respect, and accountability*. Harvard University Press.
- David, B., & Olatunji, B. O. (2011). The effect of disgust conditioning and disgust sensitivity on appraisals of moral transgressions. *Personality and Individual Differences, 50*(7), 1142–1146.
- Ditto, P. H., & Pizarro, D. A., et al. (2009). "Motivated Moral Reasoning". In: *The psychology of learning and Motivation*. D. M. Bartles, C. W. Bauman, L. J. Skitka and D. L. Medin (eds). Burlington, Academic Press. 50: 307–338.
- Eskine, K. J., Kacirik, N. A., & Prinz, J. J. (2011). A bad taste in the mouth: gustatory disgust influences moral judgment. *Psychological Science, 22*(3), 295–299. doi:10.1177/0956797611398497

- Feinberg, M., Antonenko, O., Willer, R., Horberg, E. J., & John, O. P. (2014). Gut check: Reappraisal of disgust helps explain liberal–conservative differences on issues of purity. *Emotion, 14*(3), 513.
- Gauss, G. (1990). *Value and Justification*. Cambridge: Cambridge University Press.
- Gilligan, C. (1982). *A different voice. psychological theory and women's development*. Cambridge, Mass.: Harvard University Press.
- Gjesdal, A. (2017). Moral learning, rationality, and the unreliability of affect. *Australasian Journal of Philosophy, 1*–14. doi:10.1080/00048402.2017.1366535
- Gray, K., Schein, C., & Ward, A. F. (2014). The myth of harmless wrongs in moral cognition: Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General, 143*(4), 1600.
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences, 6*(12), 517–523.
- Gutierrez, R., & Giner-Sorolla, R. (2007). Anger, disgust, and presumption of harm as reactions to taboo-breaking behaviors. *Emotion, 7*(4), 853.
- Habermas, J. (1990). *Moral consciousness and communicative action*. Cambridge, MA: MIT Press.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*(4), 814.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science, 316*(5827), 998–1002.
- Haidt, J. (2012). *The righteous mind: Why good people are divided by religion and politics*. New York, NY: Pantheon.
- Haidt, J., Koller, S., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613–628.
- Haidt, J., Rozin, P., McCauley, C., & Imada, S. (1997). Body, psyche, and culture: The relationship between disgust and morality. *Psychology and Developing Societies, 9*, 107–131.
- Hindriks, F. (2014). Intuitions, rationalizations, and justification: A defense of sentimental rationalism. *The Journal of Value Inquiry, 48*(2), 195–216.
- Hindriks, F. (2015). How does reasoning (fail to) contribute to moral judgment? Dumbfounding and disengagement. *Ethical Theory and Moral Practice, 18*(2), Springer Netherlands, 237–250.
- Jacobson, D. (2012). Moral dumbfounding and moral stupefaction. In M. Timmons (Ed.), *Oxford studies in normative ethics* (Vol. 2, pp. 289–316). New York, NY: Oxford University Press.
- Johnson, D. J., Cheung, F., & Donnellan, M. B. (2014). Does cleanliness influence moral judgments? *Social Psychology, 45*(3), 209–215.
- Johnson, D. J., Wortman, J., Cheung, F., Hein, M., Lucas, R. E., Donnellan, M. B., ... Narr, R. K. (2016). The effects of disgust on moral judgments: Testing moderators. *Social Psychological and Personality Science, 7*(7), 640–647.
- Kauppinen, A. (2010). What makes a sentiment moral? *Oxford Studies in Metaethics, 5*, 225–256. In.
- Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science, 322*(5908), 1681–1685.
- Kelly, D. (2011). *Yuck!: The nature and moral significance of disgust*. MIT Press.
- Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language, 22*(2), 117–131.
- Kohlberg, L. (1969). Stage and sequence. The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Chicago: Rand McNally.
- Kumar, V. (2015). Moral judgment as a natural kind. *Philosophical Studies, 172*(11), 2887–2910. doi:10.1007/s11098-015-0448-7

- Kumar, V. (2016). The empirical identity of moral judgment. *The Philosophical Quarterly*, 66(265), 783–804.
- Kumar, V. (2017). Foul behavior. *Philosopher's Imprint*.
- Kurth, C. (2015). *Moral anxiety and moral agency*. *Oxford studies in normative ethics*. M. Timmons (ed) (Vol. 5). OUP.
- Landy, J. F. (2016). Representations of moral violations: Category members and associated features. *Judgment & Decision Making*, 11, 5.
- Landy, J. F., & Goodwin, G. P. (2015). Does incidental disgust amplify moral judgment? A meta-analytic review of experimental evidence. *Perspectives on Psychological Science*, 10(4), 518–536.
- MacIntyre, A. (1984). *After virtue*. Notre-Dame: University of Notre Dame Press.
- May, J. (2014). Does disgust influence moral judgment?. *Australasian Journal Of Philosophy*, 92(1), 125–141. doi:10.1080/00048402.2013.797476
- Monin, B., Pizarro, D. A., & Beer, J. S. (2007). Deciding versus reacting: Conceptions of moral judgment and the reason-affect debate. *Review of General Psychology*, 11(2), 99.
- Nichols, S. (2002). Norms with feeling: towards a psychological account of moral judgment. *Cognition*, 84(2), 221–236. doi:10.1016/S0010-0277(02)00048-3
- Nichols, S. (2004). *Sentimental rules. On the natural foundations of moral judgment*. New York, NY: Oxford University Press.
- Nichols, S., Kumar, S., Lopez, T., Ayars, A., & Chan, H. Y. (2016). Rational learners and moral rules. *Mind & Language*, 31, 530–554.
- Nussbaum, M. C. (2009). *Hiding from humanity: Disgust, shame, and the law*. Princeton University Press.
- Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, 36(1), 163–177.
- Piaget, J. (1932). *The moral development of the child*. London: Kegan Paul.
- Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, 9(1), 29–43.
- Quigley, J. G. (2015). Moral psychology and the unity of morality. *Utilitas*, 27(2), 119–146.
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, 118, 57–75.
- Railton, P. (2014). The affective dog and its rational tale: Intuition and attunement. *Ethics*, 124(4), 813–859.
- Railton, P. (2017). Moral learning: Conceptual foundations and normative relevance. *Cognition*, 167, 172–190.
- Royzman, E. B., Kim, K., & Leeman, R. F. (2015). The curious tale of Julie and Mark: Unraveling the moral dumbfounding effect. *Judgment and Decision Making*, 10(4), 296.
- Rozin, P., Markwith, M., & Stoess, C. (1997). Moralization and becoming a vegetarian: the transformation of preferences into values and the recruitment of disgust. *Psychological Science*, 8(2), 67–73. doi:10.1111/j.1467-9280.1997.tb00685.x
- Rozin, P., & Singh, L. (1999). The moralization of cigarette smoking in the United States. *Journal of Consumer Psychology*, 8(3), 321–337.
- Russell, P. S., & Giner-Sorolla, R. (2011). Social justifications for moral emotions: When reasons for disgust are less elaborated than for anger. *Emotion*, 11(3), 637.
- Sauer, H. (2012). Psychopaths and filthy desks. *Ethical Theory and Moral Practice*, 15(1), 95–115.
- Sauer, H. (2017). *Moral judgments as educated intuitions*. Cambridge, MA: MIT Press.
- Scanlon, T. (1998). *What we owe to each other*. Harvard University Press.
- Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin*, 34, 1096–1109.

- Shweder, R., Much, N., Mahapatra, M., & Park, L. (1997). The 'big three' of morality (autonomy, community, and divinity) and the "big three" explanations of suffering. *Morality and Health*, 119, 119–169.
- Sinnott-Armstrong, W., & Wheatley, T. (2012). The disunity of morality and why it matters to philosophy. *The Monist*, 95(3), 355–377.
- Smetana, J. G. (1984). Toddlers' social interactions regarding moral and conventional transgressions. *Child Development*, 55, 1767–1776.
- Smetana, J. G. (1989). Toddlers' social interactions in the context of moral and conventional transgressions in the home. *Developmental Psychology*, 25(4), 499–508.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.
- Williams, B. (1985). *Ethics and the limits of philosophy*. Routledge.
- Wong, D. B. (2006). *Natural moralities: a defense of pluralistic relativism*. Oxford University Press.