A Moralization Network Theory of Moral Judgment

Matthew E. Vanaman

A Moralization Network Theory of Moral Judgment

The average American from 50 years ago would probably not have guessed that bathroom use, of all things, would become one of America's front-and-center moral issues. This American might likewise have not imagined that recreational marijuana use, of all things, would become morally acceptable to American's conservative leaders. Yet just as recent polling reveals widespread division over bathroom laws (McCarthy, 2017), two Republican congressmen have proposed a bill to legalize recreational marijuana ("Bill to repeal federal prohibition filed by two republican lawmakers," 2021). Through these topical examples, we can see that the *moral domain,* or that social category colloquially named "moral issues," is a transient one. Moreover, these examples illustrate how seemingly any non-moral issue has the potential to enter the moral domain or, as in the latter case, to leave it, should just the right circumstances come about.

What is the Moral Domain, And How Do Things Get Into It?

Why did bathroom bills enter the moral domain? Why does marijuana seem to be leaving it? These are questions about *moralization,* or the process by which previously neutral stimuli[1] become morally relevant. This definition has given moral psychology a jumping-off point for understanding moralization, though it does not tell us very much about what to look for in our search for a moralization construct. Indeed, despite some important work (e.g., Rozin, 1999), recent work has called for a more integrated framework that explains what moralization is

[1] Throughout this article, I use "stimuli" or "stimulus" generically to refer to anything that has the potential to elicit a moral judgment. This may include specific behaviors, political policies, or judgments of someone's character, to name a few examples.

exactly, and when it will or will not occur (Rhee, Schein, & Bastian, 2019). As Vaidyanathan (2020) puts it, "strong theories need strong constructs" (p. 334). As such, I aim to persuade you of a particular vision for a strong construct of moralization.

This brings me to the central question of this article: what might a strong construct of moralization look like, in theory and in practice? Before I propose an answer, let me propose what a strong construct should be able to *do*. First, a strong construct of moralization should be integrative, meaning that it should be able to incorporate as much of our extant knowledge about moral judgment and decision making as possible. Second, a strong construct of moralization should be ecologically valid, meaning it should be minimally restrictive in which stimuli have the potential for moral relevance. After all, the moral status of a stimulus can change unexpectedly and over time (Hartman, 2019), as shown in bathroom choice and recreational marijuana. Third, a strong construct of moralization should make it easy for researchers to agree about what the "moral" in moralization refers to, be it moral as in *morally desirable*, in a prescriptive sense, or moral as in *morally relevant,* in a purely descriptive sense. As I will later show, previous perspectives have been limited in one or more of these three areas. With the aim of addressing those limitations, I propose here a new construct of moralization.

Moralization Network Theory: Morality is an Emergent Subjective Experience

To understand the moral domain and moralization, we must think of them as being two sides of the same coin. More specifically, the moral domain is a complex *network of cognitive affective, and psychological contexts within an individual* which, when operating in unison, give rise to the phenomenological state of experiencing a stimulus as morally relevant, whatever that stimulus may be. In other words, the experience that an issue is particularly *moral* is an emergent property, or a byproduct of the interactions among a suite of constituent parts - reasons, emotions,

and psychological contexts (henceforth *context*) - which are non-moral when taken by themselves, yet quintessentially moral when taken together.

Why, then, would bathroom bills enter the moral domain? Why would marijuana leave it? These can be explained by the *strengthening or weakening of the relationships among the network's constituent parts*, whereby this network becomes activated in just the right way such that a stimulus comes to be experienced as morally relevant, and likewise when deactivated in just the right way, comes to be experienced as morally irrelevant. Once a stimulus is moralized, the person will experience the stimulus in distinctly *moral* way, an experience unique from the way other stimuli are experienced. I call this perspective *moralization network theory*.

Before we explore moralization network theory in more detail, let us first review what moral psychology has revealed about moral judgment so far. We will see that moral psychology has often viewed questions about the moral domain and moralization through one of two broad perspectives: *cognitivism* and *emotivism*. After reviewing what those perspectives argue for, as well as their respective shortcomings, I will show how the cognitivist and emotivist perspectives could be integrated, along with more recent findings, by adopting moralization network theory. With an eye toward integrating the complexity of moral judgment, minimizing content constraints, and forcing agreement in what our flagship construct entails, I will argue that a moralization network theory is key to advancing, adjudicating between, or consolidating the variety of theoretical perspectives found throughout the field. But first, let us briefly review the foundations of a moralization network theory, beginning with cognitivism.

Cognitivism: Morality from Reason

The cognitivist perspective has focused on how the capacity for reason develops through childhood and goes on to give way to the moral judgments of adults. Cognitivism began with Piaget's interest in how humans develop intelligence (Huitt & Hummel, 2003), though its focus

looks a bit different today. Piaget argued that for moral judgments to emerge, children must first develop the capacity for reason. Once they've acquired that capacity, children will combine reason with social experience to form moral rules (Piaget, 1932; Von Glasersfeld, 1991). Eventually, children reach "autonomous morality," meaning they can recognize that moral rules serve the rational function of constructing and maintain social systems of cooperation. In other words, once a child is able, she will see that the principle of justice holds cooperative systems together and choose to enforce it at her own volition. To Piaget, the moral domain operates like a free market which inevitably finds an equilibrium around justice (Piaget, 1977), though that particular aspect of Piagetian morality is largely absent from today's cognitivism.

Building on Piaget, Lawrence Kohlberg proposed that the capacity for moral reasoning develops through stages (Kohlberg et al., 1973). Kohlberg's moral domain was inspired by the philosophical cognitivism of Immanuel Kant, who argued that individuals use reason to apprehend objective moral truth (Rosen, 2018). With this philosophical baseline in mind (Kohlberg, 1976, p. 194), Kohlberg probed for justice-based reasoning by having children make judgments of hypothetical dilemmas, such as whether another child should steal medicine from a doctor to save a sick parent. The youngest children occupied the *pre-conventional* stages, where judgments of dilemmas are based on personal preference (e.g., "will I be punished?"; "will I gain something I want?"). Older children occupied the *conventional* stages, where judgments are based on traditions or rules (e.g., "what is the local consensus?"; "what do the authorities say?"). Adults, usually college-aged, occupied the *post-conventional* stages, where judgments are based on social contracts ("what is best for the community?") and, by the sixth and final stage, on rational justice considerations ("what is the right thing to do, regardless of how I or anyone else feels about it?"; Kavathatzopoulos, 1991; Kohlberg, 1981; Kohlberg, 1971). Indeed, Kohlberg's

theory was intertwined with Kant's philosophy, which contrasted with Piaget's view that justice is a byproduct of cooperation.

Most recent work in cognitivism comes from social domain theory (Turiel, 1983). Social domain theory argues that the stage account is inaccurate; rather, the pre-conventional (personal preferences), conventional (consensus or authority), and post-conventional (Kantian justice principles) stages each indicate the presence of distinct domains of social knowledge. If personal preferences, conventions, and justice-based morality are unique domains in their own right, then children should distinguish among them, and this is indeed what social domain theorists observed (Turiel, 2008).

Strictly speaking, social domain theory was the first to propose a "moral domain" *per se*, or the idea that morality is its own type of judgment that exists in addition to, rather than superordinate to, other types of judgments. This opened up a new question, "what is special about *moral* judgments?" Social domain theory's answer to this question was that unlike personal preference and conventional judgments, moral judgments use the following criteria:

- Obligatory: if given the opportunity to do it or abide by it, one must always do so.
- Universality: the obligation applies to everyone equally, regardless of time and place.
- Impartiality: the obligation is immutable, fact-like, deriving from dispassionate rationality rather than opinion, sentiment, or emotion.
- Authority-independence: it is obligatory regardless of social consensus, prevailing law, or what authority figures have to say about it.

Social domain theory marked an important shift from the approaches of Piaget and Kohlberg in one critical respect: social domain theory allows for moral disagreement. When Piaget and Kohlberg observed two people giving diverging judgments toward something that was supposed to be a moral issue, they would have concluded that one person was morally immature;

that is, less developed or less rational than the other. Social domain theory, on the other hand, allowed that two people can be equally developed and rational, yet give different judgments toward moral stimuli nonetheless. Social domain theory argues that *informational assumptions,* or reasoned beliefs about whether some non-moral facts apply to a stimulus (Wainryb & Turiel, 1993), can cause disagreement, such as one person judging a stimulus as a moral issue, and another judging it as a conventional issue, for example. As an illustration, Wainryb (1991) found that parents will all agree that it is unjust to use corporal punishment on their children, as long as they are all told that corporal punishment does not help to teach children right from wrong. As such, social domain theory acknowledges that moral disagreement exists, and is not due to lack of maturity or reason; rather, people sometimes have different informational assumptions about the stimulus.

Weaknesses of Cognitivism

The cognitivist approach has been criticized for being normatively biased (Habermas, 1990), particularly with respect to limiting the moral domain to justice, and for limiting moral decision making to pure reason. Kohlberg at one point did argue his work was explicitly normative (Kohlberg, 1990), and indeed his measurements of stage progression comprise "a philosophical or ethical dimension as well as a psychological dimension" (Kohlberg, 1976, p. 194), illustrating Kohlberg's commitment to a prescriptive form of cognitivism. Social domain theory, while less clear on this front, has been criticized for being *descriptively* biased. Some argue that social domain theory too often limits the content of its stimuli to well-established Western legal and moral rules (Lourenço, 2014, p. 13). Moreover, studies supporting social domain theory tend to actively seek out evidence for justice-based reasoning, while other reasons are considered, ahead of time, to be non-moral forms of reasoning. This strategy could be said to subtly repeat Kohblerg's prescriptive approach, which dismisses non-justice reasoning as

inherently irrelevant to morality. Such methods may yield the impression that morality is universally restricted to justice reasoning, yet also miss other important aspects of morality.

One such illustration is seen in Shweder, Mahapatra, and Miller (1987), which tested social domain theory's prediction that as children age, they should recognize that some moral stimuli belong to conventional domain. Indeed, American children categorized less content as part of the moral domain, and more as part of the conventional domain, compared to younger children. However, the authors found the opposite pattern in India: older children categorized *more* content as part of the moral domain, which suggested that Indian society lacks a conventional domain. The authors concluded that the conventional domain is a product of American individualist morality, which prioritizes individual freedom and wellness over community concerns (Henrich, Heine, & Norenzayan, 2010). In the United States, community concerns become part of the conventional domain; in India, where individualist morality is absent, community concerns remain part of the moral domain.

Indeed, what social domain theory calls "convention," and the content that comes with it, may be more morally important outside of the United States than researchers have given credit for (Shweder & Haidt, 1993). Nisan (1987) found that traditional Arab children view social agreements, such as when to shake hands, as bad and authority independent, with weak relationships between moral judgments and justice (Nisan, 1987). Miller and Bersoff (1992) found that Indians frequently prioritize interpersonal relationships over justice, such as finding it preferable to e.g., steal a train ticket from a stranger (a justice violation) rather than arrive late to deliver a friend's wedding ring (a convention violation). Consistent with an interpersonalized morality, Indians view reciprocity in moral terms where Americans view it in conventional or personal preference terms and are much more forgiving of justice violations when caused by immaturity or emotionality (Bersoff & Miller, 1993). Taken together, Miller (1994) drew the

same conclusion as Shweder, namely that some content that Americans typically label as convention are often moralized by other cultures. These findings suggest that there could be more to morality than justice-based reasoning.

More importantly, cognitivism seems to lack an explicit account of moralization, or the idea that the moral status of a stimulus can change over time. More specifically, the idea of "culture wars," or heated political debates around changes in lifestyle, demographics, and societal values (Hartman, 2019), pose an explanatory challenge for the informational assumptions account of moral disagreement. For example, Turiel, Hildebrandt, Wainryb, and Saltzstein (1991) observed that opponents of pornography and homosexuality tended to waver in which informational assumptions they held about those issues, despite consistency in their opposition. Moreover, Haidt, Koller, and Dias (1993) found that disgusting and disrespectful behaviors, such as washing a toilet with the country flag, were judged as immoral even when subjects explicitly stated that the behaviors were not unjust. Similarly, moral stances toward cigarette smoking, which had swiftly taken on a moralized status during the 1990s (Rozin, 1999), is predicted better by emotional factors (e.g., disgust, irritation) than by explicit beliefs about health effects or harmfulness. These studies intimate that emotion, not just reason or informational assumptions, might have something to do with why stimuli become moralized, and why moral disagreement can emerge so suddenly like what we see with culture war issues like bathroom choice.

Indeed, the cognitivist perspective idea that people use cold and logical justice-based reasoning when making moral judgments seems at odds with more recent work. Moral judgments, it seems, are particularly inflexible (Russell & Giner-Sorolla, 2011), relatively resistant to information (Royzman, Kim, & Leeman, 2015), and share neural pathways with preferences, not facts (Theriault, Waytz, Heiphetz, & Young, 2017). There is even some evidence that moral judgments are treated as *exempt* from informational assumptions (Cusimano &

Lombrozo, 2021; but see also Monroe & Malle, 2019), rather than driven by them. Furthermore, the idea that people rationally apply informational assumptions to their moral judgments seems to lack ecological validity. In theory, expanded access to the internet provides greater opportunity than ever before to establish new informational assumptions, or to be disabused of old ones. Yet in practice, internet access seems to promote *mis*informational assumptions (Carpenter, Brady, Crockett, Weber, & Sinnott-Armstrong, 2021; Vosoughi, Roy, & Aral, 2018). While one might expect that the internet would change moral judgments by changing informational assumptions, what actually happens is that people's existing moral judgments are reinforced (Brady, McLoughlin, Doan, & Crockett, 2021), often by moral outrage (Brady, Wills, Jost, Tucker, & Bavel, 2017; but see also Burton, Cruz, & Hahn, 2019). This selective attendance to information, as well as the apparent amplification of judgments through emotion, raises the possibility that moral judgment involves processes beyond rational justice-based reasoning. Another perspective on morality - *emotivism* - seeks to address this possibility.

Emotivism: Morality from Emotion

While cognitivism was inspired by the philosophy of Kant, emotivism has much more in common with philosopher David Hume. Hume argued that moral judgments are caused by quick, emotion-laden impressions, which psychologists have called *intuitions* (Hume & Selby-Bigge, 1789). As with art, moral judgments are determined by what a person or group of people experiences as subjectively pleasing or offensive. Because moral judgments are caused by emotion, reason can only serve a secondary function.

Hume's ideas about emotion and reason are similar to the view of reason held by many social psychologists. Reason is not for finding objective truth; reason is *for doing* (Fiske, 1992; James, 1907). Said another way, people's judgments are based on intuitive flashes of emotion (Kuhn, 1991; Kunda, 1990; Margolis, 1987; Perkins, Farady, & Bushey, 1991; Zajonc, 1980).

Reason comes after, usually with the purpose of providing a *post-hoc rationalization*, or a retroactive explanation for a judgment that has already been made (Nisbett & Wilson, 1977).

Based on these ideas, Haidt (2001) proposed the *social intuitionist model*. This model argues that when a person encounters a stimulus with emotionally arousing content (e.g., something disgusting), this person will feel an intuition, or the experience that the stimulus as morally pleasing or offensive. This intuition, in turn, causes a conscious moral judgment moral goodness or badness. Consistent with the motivated reasoning idea, this person only uses reason after the judgment has been made, either to justify to oneself why the judgment was made, or to persuade others to share the intuition. Although the social intuitionist model does allow for reasoning to influence one's own judgment or intuition, this happens only rarely.

The social intuitionist model was highly influential, generating much theorizing and inspiring many other models of moral judgment seeking to incorporate "intuitive" or automatic processes into our understanding of moral judgments. Almost always, these models argue for an important role for emotion, though they vary both in how central and distinct emotion is from reason. Some models do away with reason altogether; the theory of dyadic morality, for example, argues that moral judgments arise from quick, emotion-laden intuitions of powerful agents harming vulnerable patients (or victims), but does not explicate a role for effortful reasoning (Gray, Schein, & Cameron, 2017; Gray & Wegner, 2012). Other models, such as Paxton and Greene (2010)'s dual process model, argues that emotion and reason are distinct but equally important systems of moral judgment, thus could be said to occupy a space somewhere between Kant and Hume.

Emotivism and Moralization

As we have seen, cognitivism does not propose an explanation for moralization *per se*, and its reliance on informational assumptions to explain moral disagreement leaves a lot

unexplained. Emotivism, however, talks a bit more directly about the importance of moral subjectivity, emphasizing that emotion can push and pull stimuli in and out of the moral domain.

The purest form of emotivism argues that emotion alone is sufficient for moralization. Moral foundations theory, for example, argues that *purity* values, or moral stances toward the body, sex, and food, are moralized through the emotion of disgust (Graham et al., 2013). With respect to emotion, Haidt, Koller, and Dias (1993) and Rozin and Singh (1999) provided preliminary evidence that emotion, particularly disgust, may drive individual differences in what people find morally wrong. Since then, two more specific hypotheses emerged with respect to the role of emotion in moralization: the amplification hypothesis, or the idea that emotion strengthens already-moralized judgments (e.g., increased condemnation), and the moralization hypothesis, or the idea that emotion causes neutral stimuli to take on a distinctly moral valence. Some studies supported the amplification hypothesis by exposing participants to incidental disgust (e.g., foul odors), finding that participants exposed to disgusting stimuli gave stronger moral judgments than those who were not (Jones & Fitness, 2008; Schnall, 2011; Seidel & Prinz, 2013). Another study found some evidence that moralization may be driven by a combination of moral shock (e.g., emotion-provoking images) and attitude-relevant disgust (Wisneski & Skitka, 2017). Moreover, Tracy, Steckler, and Heltzel (2019) found support for what one might call the *de*amplification hypothesis: participants who consumed ginger, which suppresses nausea (Ernst & Pittler, 2000; Lien et al., 2003), gave milder moral judgments on average than those in a control group. These studies provide some evidence that emotion might be pivotal for stimuli to take on moral relevance.

Weaknesses of Emotivism

Despite the constraints of cognitivism's justice-only account of the moral domain, it does provide a clear definition of what is morally relevant, why it is, and provides criteria for

recognizing whether a person recognizes it as morally relevant. To that end, one criticism of the emotivism approach is that its conception of morality lacks in specificity. Haidt (2001), for example, defines morality as "evaluations of good or bad of the actions or character of a person made with respect to a set of virtues held to be obligatory to a culture or subculture" (p. 817). Some have argued for broad definitions like these because they give researchers the flexibility to recognize the moral undertones of issues previously considered to be non-moral, or even anti-social (Haidt & Graham, 2007). Such flexibility, however, comes at the cost clarity; for example, obligations may be non-moral, creating ambiguity about whether certain content really captures quintessentially *moral* judgments (Turiel, 2006). As such, we are currently without a clean method for distinguishing the moral from non-moral, and any method we employ will need to be able to help settle questions about the moral status of diverse content.

Given the lax demarcation between moral and non-moral, it is perhaps unsurprising that emotivism's search for the "secret ingredient" of moralization has remained elusive. Evidence for emotivism's idea that emotion can moralize stimuli by itself is mixed and contradictory. For example, moral foundations theory's idea that disgust moralizes purity content cannot explain why purity content becomes moralized. Pilot data from my own work (Vanaman & Chapman, 2021), visualized in Figure 1, shows that some disgusting but non-moral stimuli (e.g., "a man privately keeps flakes of his skin in a container") are sometimes less immoral than immoral purity stimuli (e.g., "a woman is sexually aroused by an image of Jesus"), despite being more disgusting. Furthermore, a few studies have failed to induce moralization of neutral stimuli using emotion. Wisneski and Skitka (2017) found that their moral shock effects did not affect neutral stimuli, while Jylkkä, Härkönen, and Hyönä (2021) were unable to find evidence that disgust led to harsher judgments of neutral stimuli. Moreover, while moral emotions (disgust, guilt) are correlated with increases in moralization over time (Feinberg, Kovacheff, Teper, & Inbar, 2019),

other work suggests that they might not be (Skitka, Wisneski, & Brandt, 2018). Clearly emotion is important for moralization (Feinberg, Kovacheff, Teper, & Inbar, 2019; Horberg, Oveis, Keltner, & Cohen, 2009; Rozin, 1999), though it may not be sufficient for moralization (Skitka, Wisneski, & Brandt, 2018). It might not even amplify judgments of already-moralized stimuli (Landy & Goodwin, 2015). Overall, as Landy and Goodwin (2015)'s review shows, the relationship between emotion and moralization is inconclusive.

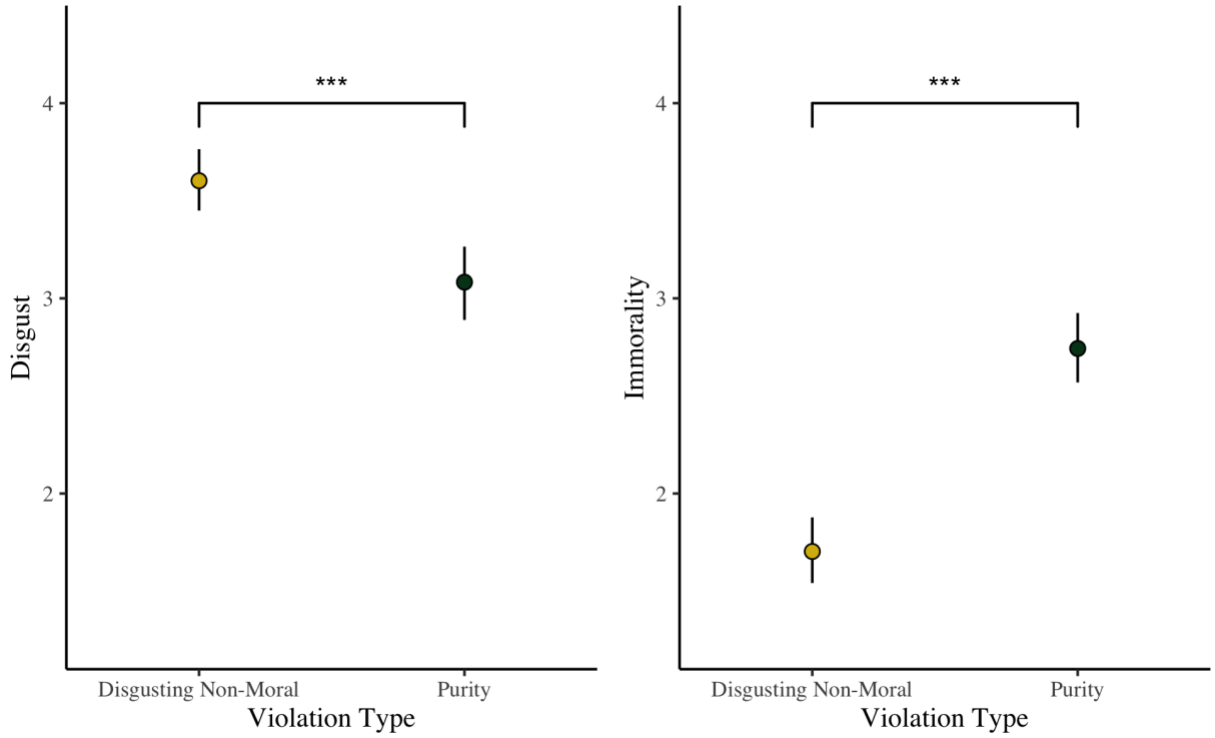

Figure 1. Pilot data from Vanaman and Chapman (2021), showing that disgusting but non-moral norms violations can be more disgusting, but less immoral, than purity violations. Significance tests were conducted with linear mixed effects models with participant random intercepts (n = 28).

On the Complexity of Morality

Cognitivist and emotivist perspectives have served as useful and generative heuristics for hypothesizing about what sorts of issues are in the moral domain, and how they get there. However, we have seen that they both have some weaknesses, at least with respect to each other. Cognitivism has used specific definitions of morality and methods for testing them, but has been criticized for not being expansive enough. Emotivism, on the other hand, sacrifices specificity for its expansiveness. In order to integrate them, along with more recent findings that don't fit neatly into either perspective, we will need characterize the construct of moralization in a way that allows for expansiveness while remaining specific enough to test. But what do we mean by "complexity?" I focus here on three types: morality's many inputs, the idea of feedback loops, and the realization by recent work that reason and emotion are not as separable as they have been traditionally described.

The (Non-)Necessity and (In)Sufficiency of Morality's Correlations

Cognitivism and emotivism have identified several variables as being correlated with moral judgment. Interestingly, many of these variables are only *sometimes* correlated with morality. That is, each of them show robust correlations with moral judgments in some studies, yet little to no correlation in others. At the very least, they each leave much to be explained.

We saw one example of this earlier in cognitivism's idea of informational assumptions: there is evidence that they can matter, but other evidence suggests they might not. Another example can be seen in the theory of dyadic morality, which argues that concerns about harm, combined with emotion, are the necessary ingredients to moralization (Gray, Schein, & Cameron, 2017; Schein & Gray, 2016). While harm does show a strong, if the not the strongest, correlation with moral judgment (Gray, Schein, & Ward, 2014; Schein & Gray, 2015b), at least some evidence suggests that moral concerns can persist in the absence of, or even subordination of,

concerns about harm (MacCoun, 2013; Rottman, Kelemen, & Young, 2014; Wylie, Alto, Flores-Robles, Satter, & Gantman, 2021).

Yet another example is the norms account, where descriptive norms (about how people generally behave) and injunctive norms (about how what people generally judge as right and wrong) often correlate with moral judgments (Dungan, Chakroff, & Young, 2017; Lindström, Jangard, Selbing, & Olsson, 2018; Nisan, 1987; Shweder, 1986). But such norms may only effect public, but not private, judgments (Paluck, 2009). Findings like these, along with the apparent mixed relationship between emotion and morality reviewed in the previous section, seem to give the impression that individual variables are neither necessary nor sufficient for explaining why moralization happens.

Context Is Important, But What Do We Do With It?

Context[2] is almost certainly important to morality (Schein, 2020). For example, Simpson, Laham, and Fiske (2016) found that relational context, such as whether a (im)moral actor is a sibling or authority figure, can moderate judgments of certain types of moral actions (e.g., unfair actions) but not others (e.g., harmful actions). As Willemsen, Newen, and Kaspar (2018) shows, changes in social roles may facilitate changes in a stimulus's moral status (for another example, Weidman, Sowden, Berg, & Kross, 2020). A person's moral identity, which can change over

[2] Note that *context* is constrained here to purely psychological terms: it refers to any perceived situational variable that varies within individuals, with or without respect to the true context (e.g., perceived social roles, a person's understanding of norms, construal level, to name a few).

time, can also bear on moral judgments (Aquino, 2002; Feinberg, Kovacheff, Teper, & Inbar, 2019).

These examples show that a person's context, real or perceived, can influence moral judgments; however, the perceived context of the stimulus itself can also influence them. For example, the construal level of a stimulus (abstract or concrete; nearby or far away) can alter attributions of moral responsibility (Hofer, Tamborini, & Ryffel, 2020), as can the identity of moral actors (Hester & Gray, 2020). Even incidental contexts, like availability of time and money, can influence moral judgments (Gino & Mogilner, 2014). While some theories do highlight a role for context in morality (Rai & Fiske, 2011), Schein (2020) has made a compelling case that a continued emphasis on context may tell us a lot about how stimuli come to be seen as moral.

The Relationship Between Reason and Emotion is Complicated

As some others have argued, the distinction between emotion and reason could be a valid one, but the distinction is at least more complicated than it gets credit for (Evans & Stanovich, 2013; Stanovich, West, & Toplak, 2011). The dual process model of moral judgment proposed by Paxton and Greene (2010), for example, argues that moral judgment indeed involves distinct systems based on reason and emotion, but these systems interact. When a person is making a moral judgment, she will be influenced by intuitive emotion responses, such as the feeling that it is wrong to harm someone, though reason will also influence her, such as through her effort to apply a moral rule consistently. Often, reason will be overwhelmed by intuition (e.g., physically killing one person to save multiple others activates aversion to harm; Paxton & Greene, 2010); in other situations, reason overrides intuitions (e.g., with time and effort, one realizes that allowing several people dying violates the principle that one should minimize harm; Paxton, Ungar, & Greene, 2012). Regardless of whether we should move on from intuition-reason dichotomies, or

try to reconcile them, one thing is clear: moral judgment is complex, and current trends suggest that the inputs to moral judgment are many, and we will probably discover more.

Feedback Loops Seem Likely - and Neglected

Another limitation of the cognitivist and emotivist perspectives is that they both neglect the possibility of psychological feedback loops, which are complex to think about but seem plausible nonetheless. For example, while emotivism adopts Zajonc (1980)'s view that emotion occurs quickly and in advance of cognition, contemporary clinical perspectives are more aligned with Lazarus (1982)'s view. To Lazarus, emotions arise neither solely from stimuli nor solely from one's environment, but from their interaction (Lazarus, 1991). This perspective, sometimes referred to as the *cognitive-appraisal* theory of emotion, is embedded into the theory behind cognitive behavioral therapy. As shown in Figure 2, cognitive behavioral therapy holds the view that emotion causes behavior (and vice-versa), behavior causes thoughts (and vice-versa), which then influence emotions (and vice-versa). Cognitive-behavioral therapy is one of the most powerful clinical interventions in patient outcomes (Barth et al., 2013; Beck, 2005; Cuijpers et al., 2020), so much so that it predominates first-line treatment in many countries (Hollon & Beck, 2013). Some argue that this is a good reason to take the cognitive-behavioral perspective, and its perspective on emotion along with it, as a strong theory (Fried, 2020). As such, it seems reasonable for moral psychology to take seriously the idea that reason and emotion are subject to feedback loops.

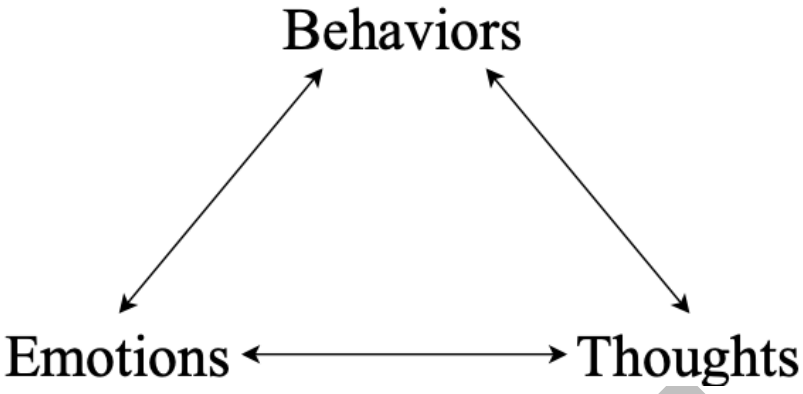

Figure 2. The "cognitive-behavioral triangle," where thoughts, feelings, and behavior influence each other in a self-reinforcing feedback loop. This is a reproduction of Figure 1 in Southam-Gerow, McLeod, Brown, Quinoy, and Avny (2011).

Another feedback loop that cognitivism and emotivism do not allow for, at least not to my knowledge, is the idea that stimuli and perception can influence each other. Cognitivism generally (if implicitly) argues for a "top-down" relationship between the mind and stimuli where reasoning alone determines how stimuli are judged. Cognitivism thus rules out a "bottom-up" relationship where stimuli, or judgments about stimuli, determine how people reason. In contrast, emotivist theories generally argue for a bottom-up relationship. That is, stimuli determine emotional responses, which in turn determine how people judge and reason about the stimuli. Emotivism has, for this reason, been criticized for neglecting the various ways in which the mind imposes patterns *onto* stimuli (Mikhail, 2007).

Although most people at this point would probably agree that stimuli can arouse emotion, recent work reveals how top-down processes can affect how people perceive stimuli in the first place. The *moral pop-out effect*, for example, shows that morally valanced stimuli exceed the threshold of perception earlier than morally-neutral stimuli (Gantman & Van Bavel, 2014). In such cases, the stimulus is not the only input into judgment; perception seems attuned, in advance, toward guiding attention toward some stimuli over others (Gantman & Van Bavel,

2015, 2016). Although it has not been tested directly, this evidence seems to suggest a stimulus-perception feedback loop. One could imagine that developing moral emotions toward a stimulus (Rozin, 1999) could in turn influence how much attention our perceptual system allots to the stimulus (Brady, Gantman, & Van Bavel, 2020). Indeed, at least some theorists argue that feedback loops are integral to moralization (Schein & Gray, 2016), though to my knowledge such accounts have yet to be fully developed and tested.

On the Need to Focus on Phenomenology Over Content

Beyond the challenges of complexity, there is another limitation of theories of moral judgment. Theories have traditionally emphasized *content* when trying to understand the moral domain, meaning they have attended mostly to what qualities moral stimuli have in common that non-moral stimuli do not have. This leads to two avoidable problems: intractability, and confusion over what we mean by "moral" in our theories.

The Intractability of Content Counting

Cataloging the content of the moral domain threatens an endless search for the exhaustive list of moral content categories. Numerous theories have offered candidate categories: dyadic morality proposes one category (Schein & Gray, 2015b); Shweder, Much, Mahapatra, and Park (1997) and Asao and Buss (2016) propose three; relational models theory proposes four (Rai & Fiske, 2011); the model of moral motives proposes six (Janoff-Bulman & Carnes, 2013); morality-as-cooperation proposes seven (Curry, Jones Chesters, & Van Lissa, 2019); the theory of universal basic values proposes ten (Schwartz, 1994), and moral foundations theory proposes somewhere between five (Graham et al., 2011) and eleven (Graham, Meindl, Koleva, Iyer, & Johnson, 2015; Iyer, Koleva, Graham, Ditto, & Haidt, 2012; Meindl, Iyer, & Graham, 2019), to name just a few examples. Each of these content categories are essentially distinct from each other, or so theories argue; this adds up to 45 potential slices of the moral domain pie.

Why has the list of prototypical moral concerns grown so large, and seem to be growing larger? One explanation is that there are indeed numerous unique content categories of morality, and that we are simply discovering them. Another explanation, and perhaps a more interesting and realistic one, is that there is no inherently non-moral content (Shweder, Mahapatra, & Miller, 1987). If this is true, then debates around the exact number of moral content categories represent hermeneutical debates rather than theoretical debates about which categories exist. As long as we focus on content, we will likely continue to uncover new content categories, because *any* stimulus content has the potential to become moralized. It is more interesting to focus on how and why people experience content as moral. Doing so may also be useful, insofar as it might force clarity about what our theories mean by "moral" while facilitating the suspension of moral disbelief necessary to truly understand what morality is and where it comes from.

The Challenge of Suspending Moral Disbelief

What does the "moral" in "moral domain" refer to? Cognitivism and emotivism have often trod a fine line between the definitions "morally desirable" and "morally relevant." Too often, it is not clear whether a theory intends to defend its moral domain in a prescriptive or descriptive sense, which limit what we can learn from our theoretical debates. This sort of thing largely happens as a result of content counting, but let us first start with some examples from the literature.

Two examples come from moral foundations theory. One theorist once wrote, "if Democrats want to understand what makes people vote Republican, they must first understand the full spectrum of American moral concerns. They should then consider whether they can use more of that spectrum themselves" (Haidt, 2008). Similarly, another article proposed that if moral foundations theory is correct that conservatives have moral motivations behind their negative attitudes toward gays, researchers should consider describing "Christian and other conservatives

in less derogatory terms" (Rosik, Dinges, & Saavedra, 2013). Strictly speaking, these arguments are prescriptive, not descriptive: a) moral foundations theory says conservatism is morally motivated; b) "moral" means desirable; therefore c) if we accept moral foundations theory, we cannot view conservatism as morally undesirable. This argument tells us nothing about how morality works; rather, it distracts us by drawing attention to *our* values, which we should hope to keep in the background if we want an accurate account of the world's diverse moralities.

Another example is seen in Kugler, Jost, and Noorbaloochi (2014), which examined the aforementioned quote that Democrats should use more of the conservative spectrum of morality. The authors' findings suggested that conservative values are better described as authoritarianism and social dominance orientation. They conclude that their findings "might lead some to question the wisdom and appropriateness of efforts to 'broaden' scientific conceptions of morality in such a way that preferences based on authoritarianism and social dominance are treated as *moral*—rather than amoral or even immoral" (emphasis theirs).

On the one hand, it is valid to ask whether moral foundations theory is factually correct in categorizing some conservative content as part of the moral domain. Yet the reader is asked to consider whether accepting moral foundations theory would be *wise* or *appropriate*. On the one hand, this study offers a theoretically interesting alternative description to moral foundations theory, namely that conservative moral content is outside of the moral domain. Yet to do so, it relies on a prescriptive argument: a) conservative moral values are actually authoritarianism and social dominance orientation; b) authoritarianism and social dominance orientation are morally undesirable, therefore c) it would be unwise or inappropriate to accept that conservative views are moral.

Sometimes, moral psychology will try to explain behavior that many would consider shocking. In these cases, ambiguity in the meaning of "moral" becomes especially dangerous to a

dispassionate understanding. One example comes from a New York Times interview, in which a social domain theorist says of an Islamic State member's rape of a 12-year-old Yazidi girl[3]: "[Moral foundations theory] is stuck trying to defend ISIS…one of these moralities (purity and harm) isn't legitimate, so how in the world is purity a 'morality?'" (Cazzell, 2019). Writing in a peer-reviewed article, the theorists argues, "if the ethic of purity and the ethic of harm are interchangeably supreme, we cannot condemn religiously based 'honor' killings of female family members, or the pronouncements by ISIS regarding the capture, enslavement and rape of non-Muslim women" (Nucci, 2016, p. 293). If one wishes to condemn ISIS, it would indeed be unwise to accept moral foundations theory. But only if one takes the "moral" in "moral foundations theory" to mean "desirable" - otherwise, it is a purely empirical question as to whether this Islamic State member subjectively felt morally motivated in doing what he did.

David Hume called these sorts of arguments *is-ought fallacies*: questions of moral desirability or undesirability are unrelated to questions of factual correctness (Hume, 2018). These arguments leave the reader wondering whether the theory is intended to offer a description of the moral domain, or to prescribe one, or perhaps to do both. Content counting shoulders at least some of the responsibility, because it is designed to lure the researcher into this ambiguity. This is because content counting follows a specific order for constructing a moral domain. Step one is to use one's own morals as the first candidate of moral content. In the case of Western, educated researchers, this will usually be justice (Haidt, Koller, & Dias, 1993). Step two, which

[3] This theorist was referring to moral foundations theory's use of the rape case as an extreme example of how moral values about purity can override harm.

often happens behind the scenes, is to test the plausibility of the candidate content by trying to imagine another person could experience that content as morally relevant. Step two might be referred to as "suspending moral disbelief."

This strategy is destined for bias, because any moral content in tension with justice will fail the test of plausibility. If justice is part of the moral domain, how could someone be unjust and feel justified about it? Entertaining that idea would require a herculean ability to suspend moral disbelief, as even mere disagreement over whether a stimulus is immoral leads to intolerance of those who disagree (Skitka, 2010). Importantly though, the researcher should have never been asked to suspend moral disbelief in the first place. That they do so is an artifact of content counting, has its order for constructing a moral domain backwards.

If we want to accurately describe the moral domain, it would be preferable that the researcher not be forced to imagine whether a rapist could feel moral. To avoid that, we must reverse the order of content counting: begin by understanding the cause moral experience and let that understanding do the work of identifying moral content.

Toward A Phenomenological Perspective on Moral Judgment

There is an alternative perspective to content counting that might solve some the problems which I have argued are caused by a content approach. I propose that we should orient our focus away from content and toward the *phenomenological experience of our research participants* - that is, morality as it is lived and felt by individuals.

The idea behind phenomenological approach is that if the moral domain is a special kind of domain, separate and above e.g., strong attitudes (Skitka, Bauman, & Sargis, 2005), then there must be a unique psychological essence to the conscious experience of morality. While the content counting approach prompts us to first ask, "what content looks to us like it might be morally relevant?" the phenomenological approach prompts to ask, "what is it like to have the

conviction that bathroom choice should be restricted?" This, of course, is no small task. Just as we might not be able to appreciate what it's like to be a bat without actually *being* a bat (Nagel, 1974), perhaps the only way to know for sure what it is like to moralize bathroom choice is to moralize it ourselves. I propose, though, that there is another way to isolate the phenomenology of morality.

A Moralization-Network Theory of Morality

Why did bathroom bills enter the moral domain? Why does marijuana seem to be leaving it? As a brief review, moral psychology has adopted two broad approaches to answering this question: the cognitivist approach, which emphasizes that moral judgments result from effortful reasoning about justice, and the emotivist approach, which emphasizes that moral judgments result from emotional intuitions caused by stimuli. I reviewed these two approaches, along with their limitations both with respect to each other, and with respect to more recent work highlighting the complexity of moral judgment, work that does not fit neatly into either perspective. Next, I argued that theoretical attempts to define the moral domain have relied too heavily on content counting, which causes two limitations to our theorizing: one, potentially endless searches for the finite set of moral content categories, and two, to ambiguity about what we mean by "moral" and, by extension, ambiguity about whether theories of moral judgment are meant to be prescriptive, descriptive, or both. I proposed that we can reconcile the limitations of content counting by replacing it with an explicit phenomenological approach. This approach would emphasize that the moral domain is best characterized by the experience that a stimulus is moral, rather than by the content of stimuli. I now propose a new perspective on morality that may alleviate some or all of this diverse set of issues, which I refer to as a *moralization-network theory* of morality.

What is Moralization-Network Theory's Chief Claim?

Moralization-network theory argues that because the process of moralization happens within a person, not within a stimulus, morality should be taken to be fundamentally about the *experiences of people* rather than about stimuli content. The moral domain is no longer a collection of content, existing independent of people; rather, it refers specifically to whichever content a particular person has moralized.

Where does the phenomenological experience of morality come from? Moralization-network theory hopes to fill out the details, but for now we will focus on its guiding idea: moral experience is an emergent property of interacting lower-level constituent parts. Although there are varying definitions of "emergence" (Baysan, 2020), I use a general definition: an *emergent property* is a secondary phenomenon whose existence is caused by the collective activity of a set of primary phenomena. Huxley (1874), in his classic analogy, described emergence as akin to the relationship between the locomotion of a steam train and its whistle. In order for a train to move, several mutually reinforcing processes must be working in unison: coal must be burning, pistons must be firing, and wheels must be rotating, to name a few. Eventually, the collective activity among the burning coal, firing pistons, and rotating wheels stabilizes, at which point the train is moving. It is under these specific circumstances that the steam train whistles.

Despite being individually correlated with the movement of a train, the causal network of coal, pistons, wheels, and their interrelations must work in unison for the train to whistle. Likewise, despite that reason and emotion are individually correlated with moral judgments of e.g., bathroom choice, the network of reasoned thoughts, emotions, and their interrelations must work in unison for bathroom choice to become moralized. Once the network stabilizes, the train whistles; once the moralization network of reasons and emotions stabilize, the person experiences unrestricted bathroom choice as *just wrong*.

Is There Precedent That Moralization-Network Theory Could Be Informative?

The idea of emergence from networks is not new to psychology. Over 100 years ago, Gestalt psychologists proposed that the *principle of totality* is crucial for understanding the mind, arguing that consciousness as a subjective experience is a system (or network, if you will) of interacting parts, best understood when considered globally rather than on the basis of each part separately (Wagemans et al., 2012). Consciousness and perception, to Gestalt psychologists, are "greater than the sum of [their] parts" (Hsiao, 1934).

The idea that psychological phenomena emerge from *networks* is not new either (Barabási, 2012). One prominent example is in the study of psychological disorders and pathologies. Traditionally, psychological disorders as characterized in the *Diagnostic and Statistical Manual* (American Psychiatric Association, 2013), or DSM-V, have been thought of as brain disorders (Insel & Cuthbert, 2015). In this *latent variable* conceptualization (see Figure 3, left), the brain disorder is taken as a hidden, or latent, common cause of several symptoms, which indicate the presence of the brain disorder when taken together. After over a half-century of investigation, researchers have found limited evidence that pathologies could be brain disorders (Fullana et al., 2020; Stein et al., 2010), leading many contemporary theorists to propose alternatives (Borsboom, Cramer, & Kalis, 2019a; Borsboom, Cramer, & Kalis, 2019b; Olthof, Hasselman, Oude Maatman, Bosman, & Lichtwarck-Aschoff, 2020).

In place of the latent variable perspective came the network perspective (Figure 3, right; Borsboom, 2017; Borsboom & Cramer, 2013; Nelson, McGorry, Wichers, Wigman, & Hartmann, 2017). This perspective proposes that pathologies are networks of mutually reinforcing interrelations among symptoms themselves. This network perspective on pathology has been used to describe depression (D'Ostilio & Garraux, 2016; Hoorelbeke, Marchetti, De Schryver, & Koster, 2016), obsessive-compulsive disorder (Jones, Mair, Riemann, Mugno, &

McNally, 2018), panic disorder (Robinaugh et al., 2019), post-traumatic stress disorder (Russell, Neill, Carrión, & Weems, 2017), borderline personality disorder (Berdahl, 2010; Southward & Cheavens, 2018), suicidal ideation (De Beurs, 2017), psychosis (Isvoranu et al., 2017), and pathology comorbidity (Cramer, Waldorp, Maas, & Borsboom, 2010), among others (for excellent reviews, see Fried, 2017, and Robinaugh, Hoekstra, Toner, & Borsboom 2020).

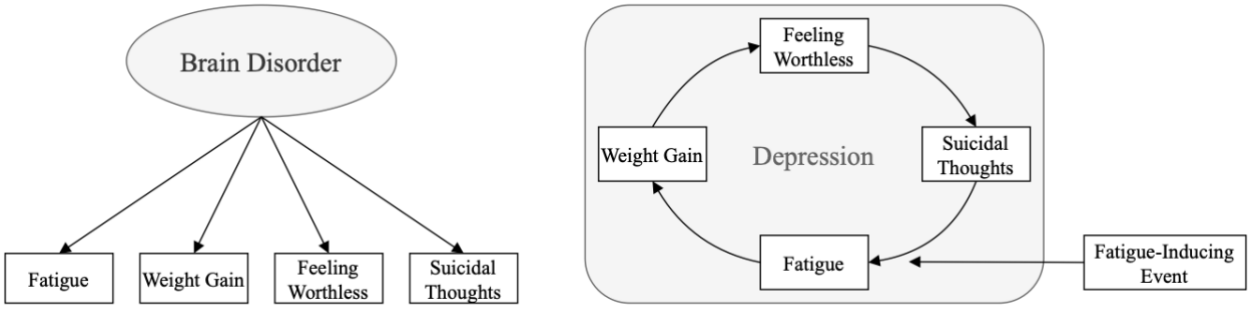

Figure 3. An example of a latent-variable, or "common cause" perspective on depression (left), contrasted with an example of a network perspective (right). In the common cause perspective, a yet-to-be-identified brain disorder causes depression, the presence of which is inferred by a suite of indirect indicators (e.g., fatigue). In the network perspective, depression is the whole of a suite of interrelations among a suite of thoughts and behaviors, creating a self-reinforcing vicious cycle.

As a matter of theoretical utility, the network perspective aligns its interpretation of symptoms away from the elusive latent variable and toward the more well-supported cognitive behavioral perspective (shown earlier in Figure 2). The network perspective has proven generative and revealing. We now know, for example, that remittance from depression is driven largely by resilience but not cognitive control (Hoorelbeke, Marchetti, De Schryver, & Koster, 2016); that loneliness or rejection drives recidivism among sex offenders (van den Berg et al., 2020); that the persistence of major depression is a function of the strength of the relationship

between guilt and fatigue (van Borkulo et al., 2015); and that post-ideation suicide attempts emerge mostly from internal entrapment and perceived burdensomeness (De Beurs, 2017).

The network perspective has also seen a productive application in areas more closely related to moral psychology. The *causal attitude network model* (Dalege et al., 2016), which draws from cognitive consistency theory (Gawronski, 2012; Gawronski & Strack, 2012), proposes that causal relationships among judgments are driven by a need for consistency. For example, a person may judge a snake as dangerous, which leads the person to judge the snake as scary (dangerous things are scary), in turn leading to the judgment that snakes should be avoided (one should avoid dangerous and scary things). These judgments feed back into the original judgment: if a snake is scary and to be avoided, it must be dangerous too. This process continues until the causal relationships have become stabilized through self-reinforcement, at which point a stable network, or "strong attitude," has formed. This vicious cycle account explains how attitudes become polarized (Dalege, Borsboom, van Harreveld, & van der Maas, 2018), why attitudes remain stable over time (Dalege, Borsboom, van Harreveld, Waldorp, & van der Maas, 2017), and how attitude (network) strength connects political interest with voting choices (Dalege, Borsboom, van Harreveld, & van der Maas, 2019).

Network models of emotion, to use another example, were introduced over 30 years (Singer & Salovey, 1988). Some current theories characterize emotion as highly distributed, or recruiting from many psychological (Lindquist & Barrett, 2012) and neurological (Van Bavel, FeldmanHall, & Mende-Siedlecki, 2015) processes, which is consistent with the network approach. On the other hand, theories that propose discrete mechanisms for discrete emotions (anger, disgust; e.g., Cosmides & Tooby, 1994), or "modular" theories, also point toward a network perspective (Bertolero, Yeo, & D'Esposito, 2015; Kragel & LaBar, 2016; Saarimäki et al., 2018). Pessoa (2017), for example, proposed that the behavioral and psychological attributes

of emotion arise from a network of neural pathways, the strength of which indicates the strength of one's experience of an emotion. The strength of an emotion can also be expressed as the strength of a network of behaviors and judgments (Bringmann et al., 2016). Taken together, these ideas are consistent with the finding that neurological pathways are reinforced when behaviors and judgments are repeated (Fields, 2008; Patel, Spreng, & Turner, 2012; Schlaug, Jäncke, Huang, Staiger, & Steinmetz, 1995). Indeed, there is an elegant theoretical symmetry to the idea that emotions can be constituted as networks of brain regions at one level of analysis (Pessoa, 2017), and as networks of behaviors and judgments at another (Bringmann et al., 2016). Given the preponderance of evidence from neighboring subfields of psychology, moralization-network theory may provide a path forward for understand moralization.

How Does Moralization-Network Theory Solve The Problems Examined In This Review?

Complexity. Cognitivism is the earliest and perhaps most-straightforward model of moral judgment. In cognitivism, reason determines which stimuli enter the moral domain, while informational assumptions alter how people reason about stimuli. Emotivism, which came later, allowed for a bit more complexity insofar as it posited separate roles for reasoning and emotion, implied a moderating role for culture, and introduced the possibility of non-justice moral content categories. More recent perspectives have drawn our attention to how morality's complexities may exceed cognitivist and emotivist models (e.g., Van Bavel, FeldmanHall, & Mende-Siedlecki, 2015; Cunningham, Zelazo, Packer, & Van Bavel, 2007; Schein, 2020; Van Bavel, Xiao, & Cunningham, 2012).

As Cramer, Waldorp, van der Maas, and Borsboom (2010) note, "complex realities require complex theories." With this complexity in mind, moralization-network theory proposes an organized theory of how relationships are established among reasoning, emotion, a potentially

moral stimulus, along with the potential to include roles for many more variables, such as context. In so doing, it forces specificity about how these variables relate.

Figure 4 shows an example network that illustrates how exposure to a stimulus can initiate changes in one or more of reason, emotion, or context; likewise, changes in reasoning, emotion, or context can influence the way a stimulus is judged. In this particular example, each relationship is mutually reinforcing, such that activation of one part (e.g., disgust) elicits a cascade of causal effects throughout the rest of the network. Said another way, a change in (perceived) context causes changes in the way the stimulus is judged, which causes changes in how the stimulus is reasoned about, and so on. This is moralization-network theory's process of *moralization*.

As the network strengthens, the person's subjective experience that the stimulus is morally relevant will likewise strengthen. In other words, once moralization of a stimulus has produced a stable network with strongly reinforced interrelationships, the subjective moral experience will begin, and the stimulus will become part of the *moral domain* for that person.

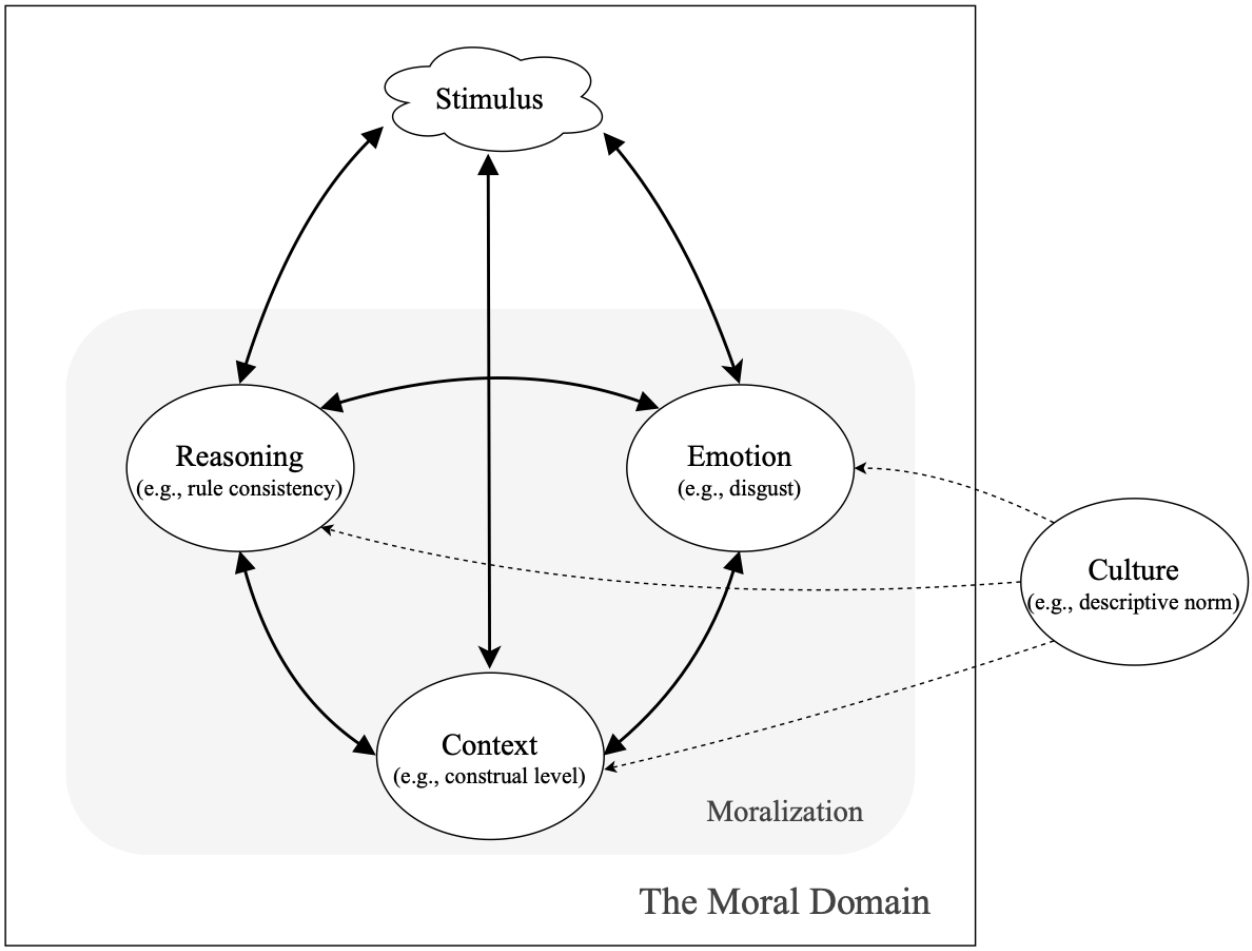

Figure 4. Moralization and the moral domain according to moralization-network theory. The gray box is moralization, where reasoning, emotion, and context mutually influence each other in a vicious cycle. Here, the qualities of a stimulus can affect reasoning, emotion, and context, which each affect each other in a feedback loop. Changes in reasoning, emotion, and context can in turn change how the stimulus is perceived or appraised. The moral domain is the experience of the stimulus as moral, which results from a stable network of reasons, emotions, and context, such that the stimulus is perceived as morally relevant. Culture can intervene on reasoning, emotion, and context, which can serve to catalyze or disrupt the network, resulting in moralization or demoralization, respectively.

One could imagine several candidate reasons, emotions, and contexts relevant to such a network. We might, for example, systemically test competing theoretical predictions about which specific reasons, emotions, or contexts should be included or excluded in the network. For example, dyadic morality's conception of morality argues that moral experience emerges from the perception of harm combined with a norms violation and any emotion that is high in arousal and negative in valence (Schein & Gray, 2016). Does a moralization network consist of these parts alone, or more? Or does one or more of them not belong?

Lastly, while emotivism implies a role for culture, it typically does not explicate how culture affects morality beyond the general idea that it results from socialization, broadly construed. In moralization-network theory, culture is said to intervene on the individual constituent parts of the network in predictable ways, at least in principle. For example, a cultural change might expose someone to a new norm, such as that a disgusting food is now delicious, at least by consensus. This new norm decreases the amount of disgust this person feels toward that food, which causes the person to view the food less abstractly (Chowdhry, Winterich, Mittal, & Morales, 2015). This shift away from abstract thinking in turn causes a milder moral judgment toward eating the food (Eyal, Liberman, & Trope, 2008). This process seems plausible, given that Americans initially found sushi disgusting before it became mainstream in the United States (Ruby & Rozin, 2019). In summary, moralization-network theory attempts to capture this kind of complexity using a conceptualization, and method, that has worked well for other subfields in psychology, and does so in a way that has the potential to integrate the many influences of moral judgment not explicitly captured by cognitivism and emotivism.

Phenomenology as the Gatekeeper of the Moral Domain. An important thrust of Moralization-network theory is that morality is a subjective experience, not a collection of content. For clarity's sake, *subjective* is meant here to convey that morality is an attribute of the

subject, or person, as opposed to being an object, or content category. This is not the same thing as relativism. It could very well be an objective fact about nature that all humans would share the subjective experience of morality toward a specific stimulus, under the same circumstances. Indeed, if we accept that humans are primates, as most psychologists probably do, then in principle any human exposed to the same stimulus under the same circumstances should experience the stimulus the same way. That is far from relativistic; in fact, unlike with content approaches, focusing on phenomenology orients our study to the aspects of morality that are universal (moral experience) and away from the aspects of morality that are descriptively relativistic (moral content). How, though, does moralization-network theory incorporate phenomenology, and how can a focus on phenomenology demystify the moral domain?

Moralization-network theory makes the assumption that for any two people for whom the same stimulus elicits the same strength of the same network, these two people will also share the same internal experience. This assumption carries much utility for distinguishing between the moral and non-moral. Whether morality is something special (Skitka, Bauman, & Sargis, 2005) or not (Knobe, 2018) likely turns on whether morality is experienced differently from strong adherence to a norm (Stich, 2018), strong attitudes (Petty & Krosnick, 2014) or personal preferences (Theriault, Waytz, Heiphetz, & Young, 2017).

Ultimately, we cannot know for sure what it is like to morally value something that we currently do not value. But with moralization-network theory, we can infer whether two stimuli elicit the same experience based on whether they elicit the same network structure and strength. For example, comparing networks could provide educated guesses about whether stimuli are experienced as strong attitudes or as a uniquely *moral* experience. An obvious starting point might be to compare one or more proposed network structures of moralization with the causal attitude model reviewed earlier. Do judgments of, say, sports and bathroom choice policies both

correspond relatively better to an attitude network? If so, the proposal that morality is special might be in trouble. Moralization-network theory also promotes fine-grained comparisons between individuals: whoever has the strongest network toward a stimulus must moralize that stimulus the most.

Importantly, this strategy encourages a more straightforward way of thinking about morality than content approaches. With content approaches, we are required to suspend our moral disbelief that someone could feel moral about content that is decidedly non-moral to the us, or perhaps vice-versa. With the phenomenological approach of moralization-network theory, our focus is directly on the experience of the person. We don't have to imagine very much - either the person responds to the stimulus in a moral way, or they do not. Deciding whether there could be something essentially moral about e.g., social dominance orientation no longer requires an egalitarian to stretch her imagination in pretending that "some groups deserve to be at the bottom" makes moral sense to someone else. Instead, the question "does social dominance orientation entail a moral experience?" is answered through empirical observation of the network and its behavior.

What Would Moralization-Network Theory Look Like in Practice?

The first task for exploring morality as a network is to identify a candidate set of constituent parts. In previous applications of the network perspective, this job was easy: all you had to do was take an existing set of scale items and transition from a latent variable interpretation (see Figure 3 right), where each item shares a common cause, to a network interpretation, where items share causal relationships with each other. Using a network model of depression as an example (see Figure 3 left), researchers inherited a full suite of constituent parts from scale development work done over the course of decades prior (e.g., Cattell & Scheier,

1961; McCrae & Costa, 1987). All they had to do was change the way they thought about the relationships among the items, and adapt the way they model the data.

Morality is a different story. No theory[4], to my knowledge, has explicitly proposed something like a latent variable model of moralization, though many have done so implicitly. For example, it seems reasonable that social domain theory's criteria of obligation, universality, impersonality, and authority-independence could be taken as indirect indicators of a moralized state. Likewise, it seems reasonable to interpret Gray and Keeney (2015)'s argument that correlations between harm and other moral concerns are so high that "most psychometricians would question their distinctness" (Schein & Gray, 2015a) as arguing for a common cause of harm. Because most models of moralization have merely hinted at a possible set of indicators, constructing such a network will require a rigorous examination of the extant literature along with an embracing of its many ambiguities and uncertainties.

---

[4] Instruments like the moral foundations questionnaire (Graham et al., 2011) do use latent variable modeling, at least to measure their moral domain. However, it is not clear how we should interpret these factors. They might indicate lower levels of a heritable foundation: a trait of *the person* (Haidt, 2016). Alternatively, the factors could simply index how much that person's culture emphasizes a particular foundation (Haidt & Joseph, 2004). Which is the common cause - person-level traits, or culture-level emphasis on some given values? It's not clear. In fairness to moral foundations theory, this sort of thing mat be an inherit problem of latent variable theory, at least given the limitations of their applications within psychological science (Borsboom, Mellenbergh, & van Heerden, 2003).

For the sake of simplicity, I have thus far limited the candidates of the network to emotion, reasoning, and context. Emotion, though, could be further divided up into anger, disgust, sadness, and other basic emotions (Ekman, 1992; but see also Russell & Barrett, 1999); reasoning into effortful thinking about harm (Cushman, Young, & Hauser, 2006), justice (Turiel, 1983), rule-consistency (Paxton & Greene, 2010), or mental states (Cushman, 2015), to name a few; and context into social roles (Willemsen, Newen, & Kaspar, 2018), status (Gilbert & McGuire, 1998), psychological distance (Eyal, Liberman, & Trope, 2008), or construal level (Hofer, Tamborini, & Ryffel, 2020), among others. Thus, the task of figuring out what should go into a network will be a lengthy one. As such, I do not propose a specific theoretical formulation for the real network of moralization looks like; instead, moralization-network theory's current status is more in line with what McGuire (1983) called a "guiding-idea theory."

To this end, the example candidate set shown in illustrates a hypothetical end goal of constructing an explicit network of moralization. In this example, the stimulus is bathroom choice, or the proposition that anyone should be allowed to use whichever bathroom, usually constrained to men or women's bathrooms, that they feel most comfortable using, a highly divisive issue in the United States (McCarthy, 2017). Participants are asked, "regardless of any policies, how morally wrong would be for someone who was born into a [male/ female] body to use the [women's/men's] bathroom, if they feel they are a [woman/man]?" to which they respond on a 6-point Likert scale where a rating of 1 = *never morally wrong* and 6 = *always morally wrong* (Vanaman & Chapman, 2020). For the sake of a simple illustration, the moralization network has been limited to the constituent parts of emotion of disgust, effortful reasoning about harmful consequences, and concrete versus abstract construal level, though in reality it would likely include many other parts.

In Figure 5, the solid lines represent causal pathways; lines with red *X*'s indicate cases where literature suggests that the causal path should be ruled out or minimized, at least after accounting for the other parts of the network (such as the path from reasoned harm concerns to bathroom choice; Vanaman & Chapman, 2020). Dotted lines connect supporting or contradicting literature to the proposed causal paths. We see that in some cases, theorists have made clear cases for causal paths in both directions; for example, Vanaman and Chapman (2020) suggests that disgust may predict moral judgments of bathroom choice, while Hutcherson and Gross (2011) and Chapman (2018) make compelling cases that disgust happens downstream of moral judgments. In this case, empirical evidence and theory suggest a feedback loop. In other cases, the literature seems to suggest contradictory pathways.

One benefit of the network perspective of moralization-network theory is that it provides clear and testable hypotheses about the causal relationships among constituent parts, and also integrates these relationships alongside relationships among several other parts. Similar to how adding a covariate to a linear regression may alter the existing relationships between the dependent variable and the independent variables (Simpson, 1951), the relationship between e.g., disgust and harm may become clear once situated within a theoretically motivated complex network. Beyond clarifying currently murky understandings of how constituent parts relate to each other, moralization-network theory may also inspire new questions. Once we begin to ask ourselves how each constituent part ought to relate to the other in theory, we will quickly find cases in which our current body of empirical evidence and theorizing has nothing to say on the matter. For example, while Napier and Luguri (2013) has shown that inducing an abstract construal level will increase concerns about harm, I was unable to find an example of someone testing whether inducing harm-based reasoning increases one's propensity to engage in abstract reasoning. In the typical psychological study, we are almost always investigating unidirectional

hypotheses, with no interest in what the theory might say about an effect that goes in the opposite direction tested in the theory. From the perspective of moralization-network theory, a feedback loop seems plausible, such that an induction of abstract thinking makes it easier to imagine potential harm, which in turn alerts one to think more abstractly, which in turn prompts one to imagine potentially harmed parties, and so on.

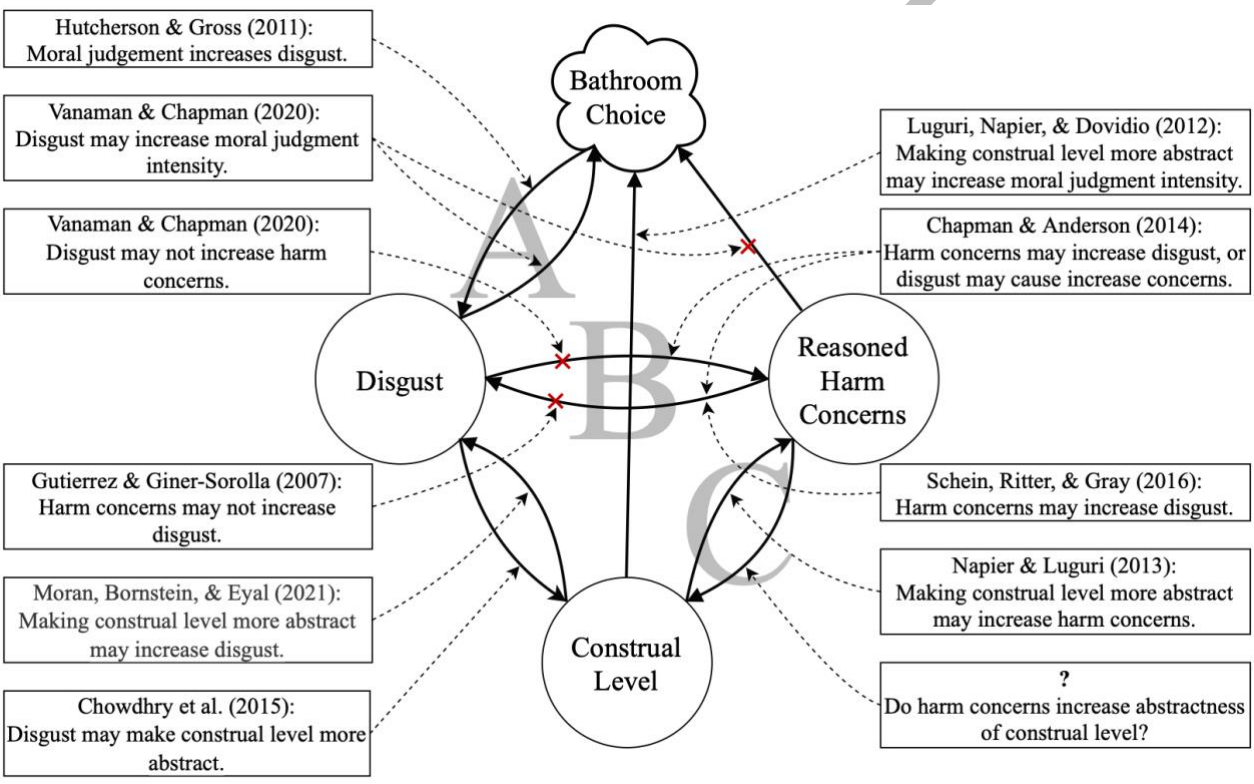

Figure 5. Example network informed by the extant literature on moral judgment. Section A represents relationships supported by evidence and theory, which a network perspective may validate and integrate. Section B represents relationships with mixed findings, which a network perspective may help clarify. Section C represents relationships that have not yet been explored, but may be suggested by, and tested with, a network perspective.

Looking Toward the Future

As we see, moralization-network theory requires a careful theoretical consideration of each detail of the network. We are forced to cease ignoring areas of the literature outside of our

immediate research, as moralization-network theory demands that we acknowledge and integrate the numerous and multifaceted characteristics of morality if we are to critically identify what is known or unknown, suggested or not suggested, about how these characteristics relate to each other. By thinking of moralization-network theory as a window into subjective, phenomenological experience, we are forced to at least entertain Shweder, Mahapatra, and Miller (1987)'s idea that there is no inherently non-moral content, flinging wide open the door to discover new parties to the process of moralization. Moralization-network theory also comes with a set of tools hand-built for investigating psychological phenomena as networks (Dalege, Borsboom, van Harreveld, & van der Maas, 2017). As others have argued for and demonstrated, network perspectives like moralization-network theory lend themselves naturally to computational modeling and the more formal theorizing seen in the mature sciences (Haslbeck, Ryan, Robinaugh, Waldorp, & Borsboom, 2019), which advance theory by forcing researchers to formalize the components of cognition and how they operate at an algorithmic level (Crockett, 2016). The network perspective of moralization-network theory, though still an infant, provides a jumping-off point for facilitating the more exhaustive theories that reveal new knowledge - as well as new mysteries.

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition). American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Aquino, K. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 14231440.

Asao, K., & Buss, D. M. (2016). The tripartite theory of machiavellian morality: Judgment, influence, and conscience as distinct moral adaptations (p. 325). Springer.

Barabási, A.-L. (2012). The network takeover. *Nature Physics*, *8*(1), 14–16. https://doi.org/10.1038/nphys2188

Barth, J., Munder, T., Gerger, H., Nüesch, E., Trelle, S., Znoj, H., … Cuijpers, P. (2013). Comparative Efficacy of Seven Psychotherapeutic Interventions for Patients with Depression: A Network Meta-Analysis. *PLOS Medicine*, *10*(5), e1001454. https://doi.org/10.1371/journal.pmed.1001454

Baysan, U. (2020). Causal Emergence and Epiphenomenal Emergence. *Erkenntnis*, *85*(4), 891–904. https://doi.org/10.1007/s10670-018-0055-z

Beck, A. T. (2005). The current state of cognitive therapy: A 40-year retrospective. *Archives of General Psychiatry*, *62*(9), 953–959. https://doi.org/10.1001/archpsyc.62.9.953

Berdahl, C. H. (2010). A neural network model of Borderline Personality Disorder. *Neural Networks*, *23*(2), 177–188. https://doi.org/10.1016/j.neunet.2009.10.007

Bersoff, D. M., & Miller, J. G. (1993). Culture, context, and the development of moral accountability judgments. *Developmental Psychology*, *29*(4), 664.

Bertolero, M. A., Yeo, B. T. T., & D'Esposito, M. (2015). The modular and integrative functional architecture of the human brain. *Proceedings of the National Academy of Sciences*, *112*(49), E6798–E6807. https://doi.org/10.1073/pnas.1510619112

Bill to repeal federal prohibition filed by two republican lawmakers. (2021). Retrieved from https://norml.org/blog/2021/05/13/bill-to-repeal-federal-prohibition-filed-by-two-republican-lawmakers/

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annual Review of Clinical Psychology*, *9*(1), 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Borsboom, D., Cramer, A. O. J., & Kalis, A. (2019). Brain disorders? Not really: Why network structures block reductionism in psychopathology research. *Behavioral and Brain Sciences*, *42*. https://doi.org/10.1017/S0140525X17002266

Borsboom, D., Cramer, A. O. J., Link to external site, this link will open in a. new window, Kalis, A., & Link to external site, this link will open in a. new window. (2019). Reductionism in retreat. *Behavioral and Brain Sciences*, *42*. https://doi.org/http://dx.doi.org/10.1017/S0140525X18002091

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. https://doi.org/10.1037/0033-295X.110.2.203

Brady, W. J., Gantman, A. P., & Van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, *149*(4), 746–756. https://doi.org/10.1037/xge0000673

Brady, W. J., McLoughlin, K., Doan, T. N., & Crockett, M. (2021). How social learning amplifies moral outrage expression in online social networks. https://doi.org/10.31234/osf.io/gf7t5

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Bavel, J. J. V. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313–7318. https://doi.org/10.1073/pnas.1618923114

Bringmann, L. F., Pe, M. L., Vissers, N., Ceulemans, E., Borsboom, D., Vanpaemel, W., … Kuppens, P. (2016). Assessing temporal emotion dynamics using networks. *Assessment*, *23*(4), 425–435. https://doi.org/10.1177/1073191116645909

Burton, J. W., Cruz, N., & Hahn, U. (2019). How Real is Moral Contagion in Online Social Networks?, 7.

Carpenter, J., Brady, W., Crockett, M., Weber, R., & Sinnott-Armstrong, W. (2021). Political Polarization and Moral Outrage on Social Media, 15.

Cattell, R. B., & Scheier, I. H. (1961). *The meaning and measurement of neuroticism and anxiety*. Oxford, England: Ronald.

Cazzell, A. (2017). The moral science podcast: "Stupid rules": Social domains and moral development with Larry Nucci on apple podcasts. Retrieved from https://podcasts.apple.com/us/podcast/stupid-rules-social-domains-moral-development-larry/id1468695145?i=1000442667838

Chapman, Hanah A. (2018). A component process model of disgust, anger, and moral judgment. In K. Gray & J. Graham (Eds.) (pp. 70–80).

Chapman, Hanah A., & Anderson, A. K. (2014). Trait physical disgust is related to moral judgments outside of the purity domain. *Emotion*, *14*(2), 341348. https://doi.org/10.1037/a0035120

Chowdhry, N., Winterich, K. P., Mittal, V., & Morales, A. C. (2015). Not all negative emotions lead to concrete construal. *International Journal of Research in Marketing*, *32*(4), 428–430. https://doi.org/10.1016/j.ijresmar.2015.03.006

Cosmides, L., & Tooby, J. (1994). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfeld & S. A. Gelman (Eds.) (1st ed., pp. 85–116). Cambridge University Press. https://doi.org/10.1017/CBO9780511752902.005

Cramer, A. O. J., Waldorp, L. J., Maas, H. L. J. van der, & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137–150. https://doi.org/10.1017/S0140525X09991567

Cramer, A. O. J., Waldorp, L. J., van der Maas, H. L. J., & Borsboom, D. (2010). Complex realities require complex theories: Refining and extending the network approach to mental disorders. *Behavioral and Brain Sciences*, *33*(2-3), 178–193. https://doi.org/http://dx.doi.org/10.1017/S0140525X10000920

Crockett, M. J. (2016). Computational Modeling of Moral Decisions. In J. P. Forgas, L. Jussim, & P. A. M. Van Lange (Eds.) (1st ed., pp. 71–90). Routledge. https://doi.org/10.4324/9781315644189-5

Cuijpers, P., Noma, H., Karyotaki, E., H. ?Vinkers, C., Cipriani, A., & Furukawa, T. A. (2020). A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry*, *19*(1), 92–107. https://doi.org/https://doi.org/10.1002/wps.20701

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The Iterative Reprocessing Model: A Multilevel Framework for Attitudes and Evaluation. *Social Cognition*, *25*(5), 736–760. https://doi.org/10.1521/soco.2007.25.5.736

Curry, O. S., Jones Chesters, M., & Van Lissa, C. J. (2019). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *Journal of Research in Personality*, *78*, 106–124. https://doi.org/10.1016/j.jrp.2018.10.008

Cushman, F. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology*, *6*, 97–103. https://doi.org/10.1016/j.copsyc.2015.06.003

Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, *17*(12), 1082–1089. https://doi.org/10.1111/j.1467-9280.2006.01834.x

Cusimano, C., & Lombrozo, T. (2021). Morality justifies motivated reasoning in the folk ethics of belief. *Cognition*, 104513. https://doi.org/10.1016/j.cognition.2020.104513

D'Ostilio, K., & Garraux, G. (2016). The network model of depression as a basis for new therapeutic strategies for treating major depressive disorder in parkinson's disease. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00161

Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. J. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological Review*, *123*(1), 2–22. https://doi.org/10.1037/a0039802

Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. J. (2017). Network analysis on attitudes: A brief tutorial. *Social Psychological and Personality Science*, *8*(5), 528–537. https://doi.org/10.1177/1948550617709827

Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. J. (2018). The Attitudinal Entropy (AE) Framework as a General Theory of Individual Attitudes. *Psychological Inquiry*, *29*(4), 175–193. https://doi.org/10.1080/1047840X.2018.1537246

Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. J. (2019). A network perspective on attitude strength: Testing the connectivity hypothesis. *Social Psychological and Personality Science*, *10*(6), 746–756. https://doi.org/10.1177/1948550618781062

Dalege, J., Borsboom, D., van Harreveld, F., Waldorp, L. J., & van der Maas, H. L. J. (2017). Network Structure Explains the Impact of Attitudes on Voting Decisions. *Scientific Reports*, *7*(1), 4909. https://doi.org/10.1038/s41598-017-05048-y

De Beurs, D. (2017). Network Analysis: A Novel Approach to Understand Suicidal Behaviour. *International Journal of Environmental Research and Public Health*, *14*(3), 219. https://doi.org/10.3390/ijerph14030219

Dungan, J. A., Chakroff, A., & Young, L. (2017). The relevance of moral norms in distinct relational contexts: Purity versus harm norms regulate self-directed actions. *PLOS ONE*, *12*(3), e0173405. https://doi.org/10.1371/journal.pone.0173405

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169–200. https://doi.org/10.1080/02699939208411068

Ernst, E., & Pittler, M. H. (2000). Efficacy of ginger for nausea and vomiting: a systematic review of randomized clinical trials. *British Journal of Anaesthesia*, *84*(3), 367–371. https://doi.org/10.1093/oxfordjournals.bja.a013442

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Eyal, T., Liberman, N., & Trope, Y. (2008). Judging near and distant virtue and vice. *Journal of Experimental Social Psychology*, *44*(4), 1204–1209. https://doi.org/10.1016/j.jesp.2008.03.012

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of moralization: How eating meat becomes a moral issue. *Journal of Personality and Social Psychology*, *117*(1), 50–72. https://doi.org/10.1037/pspa0000149

Fields, R. D. (2008). White matter in learning, cognition and psychiatric disorders. *Trends in Neurosciences*, *31*(7), 361–370. https://doi.org/10.1016/j.tins.2008.04.001

Fiske, S. T. (1992). Thinking is for doing: Portraits of social cognition from daguerreotype to laserphoto. *Journal of Personality and Social Psychology*, *63*(6), 877.

Fried, E. I. (2020). Theories and Models: What They Are, What They Are for, and What They Are About. *Psychological Inquiry*, *31*(4), 336–344. https://doi.org/10.1080/1047840X.2020.1854011

Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, *12*(6), 999–1020. https://doi.org/10.1177/1745691617705892

Fullana, M. A., Abramovitch, A., Via, E., López-Sola, C., Goldberg, X., Reina, N., … Radua, J. (2020). Diagnostic biomarkers for obsessive-compulsive disorder: A reasonable quest or ignis fatuus? *Neuroscience & Biobehavioral Reviews*, *118*, 504–513. https://doi.org/10.1016/j.neubiorev.2020.08.008

Gantman, A. P., & Van Bavel, J. J. (2014). The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition*, *132*(1), 22–29. https://doi.org/10.1016/j.cognition.2014.02.007

Gantman, A. P., & Van Bavel, J. J. (2015). Moral Perception. *Trends in Cognitive Sciences*, *19*(11), 631–633. https://doi.org/10.1016/j.tics.2015.08.004

Gantman, A. P., & Van Bavel, J. J. (2016). See for yourself: Perception is attuned to morality. *Trends in Cognitive Sciences*, *20*(2), 76–77. https://doi.org/10.1016/j.tics.2015.12.001

Gawronski, B. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition*, *30*(6), 652–668. https://doi.org/10.1521/soco.2012.30.6.652

Gawronski, B., & Strack, F. (2012). Cognitive consistency as a basic principle of social information processing (pp. 1–16). New York, NY, US: Guilford Press.

Gilbert, P., & McGuire, M. (1998). Shame, social roles and status: The psychobiological continuum from monkey to human. *Shame: Interpersonal Behavior, Psychopathology and Culture*, 99125.

Gino, F., & Mogilner, C. (2014). Time, Money, and Morality. *Psychological Science*, *25*(2), 414–421. https://doi.org/10.1177/0956797613506438

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2013). Chapter two - moral foundations theory: The pragmatic validity of moral pluralism. In P. Devine & A. Plant (Eds.) (Vol. 47, p. 55130). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00002-4

Graham, J., Meindl, P., Koleva, S., Iyer, R., & Johnson, K. M. (2015). When Values and Behavior Conflict: Moral Pluralism and Intrapersonal Moral Hypocrisy. *Social and Personality Psychology Compass*, *9*(3), 158–170. https://doi.org/https://doi.org/10.1111/spc3.12158

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366385. https://doi.org/10.1037/a0021847

Gray, K., & Keeney, J. E. (2015). Impure or just weird? Scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*, *6*(8), 859868. https://doi.org/10.1177/1948550615592241

Gray, K., Schein, C., & Cameron, C. D. (2017). How to think about emotion and morality: Circles, not arrows. *Current Opinion in Psychology*, *17*, 4146. https://doi.org/10.1016/j.copsyc.2017.06.011

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), 16001615. https://doi.org/10.1037/a0036149

Gray, K., & Wegner, D. M. (2012). Morality takes two: Dyadic morality and mind perception. (p. 109127). American Psychological Association. https://doi.org/10.1037/13091-006

Gutierrez, R., & Giner-Sorolla, R. (2007). Anger, disgust, and presumption of harm as reactions to taboo-breaking behaviors. *Emotion*, *7*(4), 853868. https://doi.org/10.1037/1528-3542.7.4.853

Habermas, J. (1990). Justice and solidarity: On the discussion concerning stage 6. *The Moral Domain: Essays in the Ongoing Discussion Between Philosophy and the Social Sciences*, 224254.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. In J. E. Adler & L. J. Rips (Eds.) (p. 10241052). Cambridge University Press.

Haidt, J. (2008). What makes people vote republican? Retrieved from https://www.edge.org/conversation/what-makes-vote-republican

Haidt, J. (2016). Are moral foundations heritable? probably. Retrieved from https://righteousmind.com/are-moral-foundations-heritable-probably/

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research*, *20*(1), 98116. https://doi.org/10.1007/s11211-007-0034-z

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 5566. https://doi.org/10.1162/0011526042365555

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, *65*(4), 613628. https://doi.org/10.1037/0022-3514.65.4.613

Hartman, A. (2019). *A War for the Soul of America, Second Edition: A History of the Culture Wars*. University of Chicago Press.

Haslbeck, J., Ryan, O., Robinaugh, D., Waldorp, L., & Borsboom, D. (2019). Modeling psychopathology: From data models to formal theories. https://doi.org/10.31234/osf.io/jgm7f

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 6183.

Hester, N., & Gray, K. (2020). The Moral Psychology of Raceless, Genderless Strangers. *Perspectives on Psychological Science*, *15*(2), 216–230. https://doi.org/10.1177/1745691619885840

Hofer, M., Tamborini, R., & Ryffel, F. A. (2020). Between a rock and a hard place: The role of moral intuitions and social distance in determining moral judgments of an agent in a moral dilemma. *Journal of Media Psychology: Theories, Methods, and Applications*, No Pagination Specified–No Pagination Specified. https://doi.org/10.1027/1864-1105/a000284

Hollon, S. D., & Beck, A. T. (2013). Cognitive and cognitive-behavioral therapies. In M. J. Lambert (Ed.) (Vol. 6, pp. 393–442). John Wiley & Sons.

Hoorelbeke, K., Marchetti, I., De Schryver, M., & Koster, E. H. W. (2016). The interplay between cognitive risk and resilience factors in remitted depression: A network analysis. *Journal of Affective Disorders*, *195*, 96–104. https://doi.org/10.1016/j.jad.2016.02.001

Horberg, E. J., Oveis, C., Keltner, D., & Cohen, A. B. (2009). Disgust and the moralization of purity. *Journal of Personality and Social Psychology*, *97*(6), 963976. https://doi.org/10.1037/a0017423

Hsiao, H. H. (1934). *Principles of gestalt psychology*. Oxford: National Editing Bureau. Retrieved from http://ezproxy.gc.cuny.edu/login?url=

Huitt, W., & Hummel, J. (2003). Piaget's theory of cognitive development. *Educational Psychology Interactive*, *3*(2), 15.

Hume, D. (2018). *An Enquiry Concerning the Principles of Morals*. Yale University Press. Retrieved from https://www.degruyter.com/document/doi/10.12987/9780300240504-006/html

Hume, D., & Selby-Bigge, L. A. (1789). *A treatise of human nature, 3 volumes, reprinted from the original ed.* The Clarendon Press. Retrieved from http://ezproxy.gc.cuny.edu/login?url=

Hutcherson, C. A., & Gross, J. J. (2011). The moral emotions: A socialfunctionalist account of anger, disgust, and contempt. *Journal of Personality and Social Psychology*, *100*(4), 719737. https://doi.org/10.1037/a0022408

Huxley, T. H. (1874). On the Hypothesis that Animals are Automata, and its History *. *Nature*, *10*(253), 362–366. https://doi.org/10.1038/010362a0

Insel, T. R., & Cuthbert, B. N. (2015). Brain disorders? Precisely. *Science*, *348*(6234), 499–500. https://doi.org/10.1126/science.aab2358

Isvoranu, A.-M., van Borkulo, C. D., Boyette, L.-L., Wigman, J. T. W., Vinkers, C. H., Borsboom, D., & Group Investigators. (2017). A Network Approach to Psychosis: Pathways Between Childhood Trauma and Psychotic Symptoms. *Schizophrenia Bulletin*, *43*(1), 187–196. https://doi.org/10.1093/schbul/sbw055

Iyer, R., Koleva, S., Graham, J., Ditto, P., & Haidt, J. (2012). Understanding Libertarian Morality: The Psychological Dispositions of Self-Identified Libertarians. *PLOS ONE*, *7*(8), e42366. https://doi.org/10.1371/journal.pone.0042366

James, W. (1907). *Pragmatism: A new name for some old ways of thinking.* Longmans, Green; Co. Retrieved from http://ezproxy.gc.cuny.edu/login?url=

Janoff-Bulman, R., & Carnes, N. C. (2013). Surveying the Moral Landscape: Moral Motives and Group-Based Moralities. *Personality and Social Psychology Review*, *17*(3), 219–236. https://doi.org/10.1177/1088868313480274

Jones, A., & Fitness, J. (2008). Moral hypervigilance: the influence of disgust sensitivity in the moral domain. *Emotion (Washington, D.C.)*, *8*(5), 613–627. https://doi.org/10.1037/a0013435

Jones, P. J., Mair, P., Riemann, B. C., Mugno, B. L., & McNally, R. J. (2018). A network perspective on comorbid depression in adolescents with obsessive-compulsive disorder. *Journal of Anxiety Disorders*, *53*, 1–8. https://doi.org/10.1016/j.janxdis.2017.09.008

Jylkkä, J., Härkönen, J., & Hyönä, J. (2021). Incidental disgust does not cause moral condemnation of neutral actions. *Cognition and Emotion*, *35*(1), 96–109. https://doi.org/10.1080/02699931.2020.1810639

Kavathatzopoulos, I. (1991). Kohlberg and piaget: Differences and similarities. *Journal of Moral Education*, *20*(1), 4754.

Knobe, J. (2018). There is no important distinction between moral and nonmoral cognition. In K. Gray & J. Graham (Eds.) (pp. 556–564).

Kohlberg, L. (1971). *From is to ought: How to commit the naturalistic fallacy and get away with it in the study of moral development.* New York, New York: Academic Press.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. *Moral Development and Behavior*, 171205.

Kohlberg, L. (1981). The philosophy of moral development moral stages and the idea of justice.

Kohlberg, L., & others. (1973). *Moral development*. McGraw-Hill Films.

Kragel, P. A., & LaBar, K. S. (2016). Decoding the Nature of Emotion in the Brain. *Trends in Cognitive Sciences*, *20*(6), 444–455. https://doi.org/10.1016/j.tics.2016.03.011

Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another Look at Moral Foundations Theory: Do Authoritarianism and Social Dominance Orientation Explain Liberal-Conservative Differences in "Moral" Intuitions? *Social Justice Research*, *27*(4), 413–431. https://doi.org/10.1007/s11211-014-0223-5

Kuhn, D. (1991). *The Skills of Argument*. Cambridge University Press.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480498. https://doi.org/10.1037/0033-2909.108.3.480

Landy, J. F., & Goodwin, G. P. (2015). Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence, 19.

Lawrence Kohlberg, D. R. (1990). The return of stage 6: Its principle and moral point of view. *The Moral Domain: Essays in the Ongoing Discussion Between Philosophy and the Social Sciences*, 151.

Lazarus, Richard S. (1982). Thoughts on the relations between emotion and cognition. *American Psychologist*, *37*(9), 1019–1024. https://doi.org/10.1037/0003-066X.37.9.1019

Lazarus, Richard S. (1991). Progress on a Cognitive-Motivational-Relational Theory of Emotion. *American Psychologist*, 16.

Lien, H.-C., Sun, W. M., Chen, Y.-H., Kim, H., Hasler, W., & Owyang, C. (2003). Effects of ginger on motion sickness and gastric slow-wave dysrhythmias induced by circular vection. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *284*(3), G481–G489. https://doi.org/10.1152/ajpgi.00164.2002

Lindquist, K. A., & Barrett, L. F. (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends in Cognitive Sciences*, *16*(11), 533–540. https://doi.org/10.1016/j.tics.2012.09.005

Lindström, B., Jangard, S., Selbing, I., & Olsson, A. (2018). The role of a "common is moral" heuristic in the stability and change of moral norms. *Journal of Experimental Psychology: General*, *147*(2), 228–242. https://doi.org/10.1037/xge0000365

Lourenço, O. (2014). Domain theory: A critical review. *New Ideas in Psychology*, *32*, 1–17. https://doi.org/https://doi.org/10.1016/j.newideapsych.2013.08.001

Luguri, J. B., Napier, J. L., & Dovidio, J. F. (2012). Reconstruing Intolerance: Abstract Thinking Reduces Conservatives' Prejudice Against Nonnormative Groups. *Psychological Science*, *23*(7), 756–763. https://doi.org/10.1177/0956797611433877

MacCoun, R. J. (2013). Moral Outrage and Opposition to Harm Reduction. *Criminal Law and Philosophy*, *7*(1), 83–98. https://doi.org/10.1007/s11572-012-9154-0

Margolis, H. (1987). *Patterns, thinking, and cognition: A theory of judgment*. University of Chicago Press.

McCarthy, J. (2017). Americans split over new LGBT protections, restroom policies. *Gallup.com*. Retrieved from https://news.gallup.com/poll/210887/americans-split-new-lgbt-protections-restroom-policies.aspx

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81–90. https://doi.org/10.1037/0022-3514.52.1.81

McGuire, W. J. (1983). A Contextualist Theory of Knowledge: Its Implications for Innovation and Reform in Psychological Research**The writing of this chapter was substantially aided by a grant to the author, Grant Number 5 ROI MH 32588 from the National Institutes of Mental Health, Interpersonal Processes and Problems Section (BR-S). In L. Berkowitz (Ed.) (Vol. 16, pp. 1–47). Academic Press. https://doi.org/10.1016/S0065-2601(08)60393-7

Meindl, P., Iyer, R., & Graham, J. (2019). Distributive justice beliefs are guided by whether people think the ultimate goal of society is well-being or power. *Basic and Applied Social Psychology*, *41*(6), 359–385. https://doi.org/10.1080/01973533.2019.1663524

Mikhail, J. (2007). *Moral Cognition and Computational Theory*. Rochester, NY. Retrieved from https://papers.ssrn.com/abstract=1029511

Miller, J. G. (1994). Cultural diversity in the morality of caring: Individually oriented versus duty-based interpersonal moral codes. *Cross-Cultural Research*, *28*(1), 339.

Miller, J. G., & Bersoff, D. M. (1992). Culture and moral judgment: How are conflicts between justice and interpersonal responsibilities resolved? *Journal of Personality and Social Psychology*, *62*(4), 541.

Monroe, A. E., & Malle, B. F. (2019). People systematically update moral judgments of blame. *Journal of Personality and Social Psychology*, *116*(2), 215. https://doi.org/10.1037/pspa0000137

Moran, T., Bornstein, O., & Eyal, T. (2021). The level of construal involved in the elicitation of core versus moral disgust. *Emotion*, *21*(2), 391–404. https://doi.org/10.1037/emo0000709

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450. https://doi.org/10.2307/2183914

Napier, J. L., & Luguri, J. B. (2013). Moral mind-sets: Abstract thinking increases a preference for "individualizing" over "binding" moral foundations. *Social Psychological and Personality Science*, *4*(6), 754759. https://doi.org/10.1177/1948550612473783

Nelson, B., McGorry, P. D., Wichers, M., Wigman, J. T. W., & Hartmann, J. A. (2017). Moving from static to dynamic models of the onset of mental disorder: A review. *JAMA Psychiatry*, *74*(5), 528–534. https://doi.org/10.1001/jamapsychiatry.2017.0001

Nisan, M. (1987). Moral norms and social conventions: A cross-cultural comparison. *Developmental Psychology*, *23*(5), 719.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231.

Nucci, L. (2016). Recovering the role of reasoning in moral education to address inequity and social justice. *Journal of Moral Education*, *45*(3), 291–307. https://doi.org/10.1080/03057240.2016.1167027

Olthof, M., Hasselman, F., Oude Maatman, F. J. W., Bosman, A. M. T., & Lichtwarck-Aschoff, A. (2020). *Complexity Theory of Psychopathology*. https://doi.org/10.31234/osf.io/f68ej

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, *96*(3), 574–587. https://doi.org/10.1037/a0011989

Paxton, J. M., & Greene, J. D. (2010). Moral Reasoning: Hints and Allegations. *Topics in Cognitive Science*, *2*(3), 511–527. https://doi.org/10.1111/j.1756-8765.2010.01096.x

Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and Reasoning in Moral Judgment. *Cognitive Science*, *36*(1), 163–177. https://doi.org/10.1111/j.1551-6709.2011.01210.x

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. *Informal Reasoning and Education.*, 83–105.

Pessoa, L. (2017). A Network Model of the Emotional Brain. *Trends in Cognitive Sciences*, *21*(5), 357–371. https://doi.org/10.1016/j.tics.2017.03.002

Petty, R. E., & Krosnick, J. A. (2014). *Attitude Strength: Antecedents and Consequences*. Psychology Press.

Piaget, J. (1932). *The moral judgment of the child.* Harcourt, Brace. Retrieved from http://ezproxy.gc.cuny.edu/login?url=

Piaget, Jean. (1977). *The development of thought: Equilibration of cognitive structures.(trans a. rosin).* Viking.

Patel, R., Spreng, R.N., & Turner, G.R. (2012). Functional brain changes following cognitive and motor skills training: A quantitative meta-analysis. *Neurorehabilitation and Neural Repair*, *27*(3), 187–199. https://doi.org/10.1177/1545968312461718

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *118*(1), 57.

Rhee, J. J., Schein, C., & Bastian, B. (2019). The what, how, and why of moralization: A review of current definitions, methods, and evidence in moralization research. *Social and Personality Psychology Compass*, *13*(12), e12511. https://doi.org/10.1111/spc3.12511

Robinaugh, D., Haslbeck, J., Waldorp, L., Kossakowski, J. J., Fried, E., Millner, A., … Borsboom, D. (2019). Advancing the network theory of mental disorders: A computational model of panic disorder. https://doi.org/10.31234/osf.io/km37w

Robinaugh, D. J., Hoekstra, R. H. A., Toner, E. R., & Borsboom, D. (2020). The network approach to psychopathology: A review of the literature 20082018 and an agenda for

future research. *Psychological Medicine*, *50*(3), 353–366. https://doi.org/10.1017/S0033291719003404

Rosen, A. (2018). *Kant's theory of justice*. Cornell University Press.

Rosik, C. H., Dinges, L. J., & Saavedra, N. (2013). Moral Intuitions and Attitudes toward Gay Men: Can Moral Psychology Add to Our Understanding of Homonegativity? *Journal of Psychology and Theology*, *41*(4), 315–326. https://doi.org/10.1177/009164711304100404

Rottman, J., Kelemen, D., & Young, L. (2014). Purity matters more than harm in moral judgments of suicide: Response to gray (2014). *Cognition*, *133*(1), 332334. https://doi.org/10.1016/j.cognition.2014.06.008

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), 18.

Rozin, P. (1999). The process of moralization. *Psychological Science*, *10*(3), 218–221. https://doi.org/10.1111/1467-9280.00139

Rozin, P., & Singh, L. (1999). The Moralization of Cigarette Smoking in the United States. *Journal of Consumer Psychology*, *8*(3), 321–337. https://doi.org/10.1207/s15327663jcp0803_07

Ruby, M. B., & Rozin, P. (2019). Disgust, sushi consumption, and other predictors of acceptance of insects as food by Americans and Indians. *Food Quality and Preference*, *74*, 155–162. https://doi.org/10.1016/j.foodqual.2019.01.013

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*(5), 805–819. https://doi.org/10.1037/0022-3514.76.5.805

Russell, J. D., Neill, E. L., Carrión, V. G., & Weems, C. F. (2017). The Network Structure of Posttraumatic Stress Symptoms in Children and Adolescents Exposed to Disasters.

Journal of the American Academy of Child & Adolescent Psychiatry, *56*(8), 669–677.e5. https://doi.org/10.1016/j.jaac.2017.05.021

Russell, P. S., & Giner-Sorolla, R. (2011). Moral anger is more flexible than moral disgust. *Social Psychological and Personality Science*, *2*(4), 360364.

Saarimäki, H., Ejtehadian, L. F., Glerean, E., Jääskeläinen, I. P., Vuilleumier, P., Sams, M., & Nummenmaa, L. (2018). Distributed affective space represents multiple emotion categories across the human brain. *Social Cognitive and Affective Neuroscience*, *13*(5), 471–482. https://doi.org/10.1093/scan/nsy018

Schein, C. (2020). The importance of context in moral judgments. *Perspectives on Psychological Science*, *15*(2), 207215. https://doi.org/10.1177/1745691620904083

Schein, C., & Gray, K. (2015a). Making sense of moral disagreement: Liberals, conservatives, and the harm-based template they share. Retrieved from https://www.spsp.org/news-center/blog/making-sense-of-moral

Schein, C., & Gray, K. (2015b). The unifying moral dyad. *Personality and Social Psychology Bulletin*, *41*(8), 11471163. https://doi.org/10.1177/0146167215591501

Schein, C., & Gray, K. (2016). Moralization and harmification: The dyadic loop explains how the innocuous becomes harmful and wrong. *Psychological Inquiry*, *27*(1), 6265. https://doi.org/10.1080/1047840x.2016.1111121

Schein, C., Ritter, R. S., & Gray, K. (2016). Harm mediates the disgust-immorality link. *Emotion*, *16*(6), 862876. https://doi.org/10.1037/emo0000167

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia*, *33*(8), 1047–1055. https://doi.org/10.1016/0028-3932(95)00045-5

Schnall, S. (2011). Clean, proper and tidy are more than the absence of dirty, disgusting and wrong. *Emotion Review*, *3*(3), 264266. https://doi.org/10.1177/1754073911402397

Schwartz, S. H. (1994). Are There Universal Aspects in the Structure and Contents of Human Values? *Journal of Social Issues*, *50*(4), 19–45. https://doi.org/10.1111/j.1540-4560.1994.tb01196.x

Seidel, A., & Prinz, J. (2013). Sound morality: Irritating and icky noises amplify judgments in divergent moral domains. *Cognition*, *127*(1), 15. https://doi.org/10.1016/j.cognition.2012.11.004

Shweder, R. A. (1986). Divergent rationalities. *Metatheory in Social Science: Pluralisms and Subjectivities*, 163196.

Shweder, R. A., & Haidt, J. (1993). The future of moral psychology: Truth, intuition, and the pluralist way. *Psychological Science*, *4*(6), 360365.

Shweder, R. A., Mahapatra, M., & Miller, J. G. (1987). Culture and moral development. *The Emergence of Morality in Young Children*, 183.

Shweder, R., Much, N., Mahapatra, M., & Park, L. (1997). Divinity and the "big three" explanations of suffering. *Morality and Health*, *119*, 119169.

Simpson, A., Laham, S. M., & Fiske, A. P. (2016). Wrongness in different relationships: Relational context effects on moral judgment. *The Journal of Social Psychology*, *156*(6), 594–609. https://doi.org/10.1080/00224545.2016.1140118

Simpson, E. H. (1951). The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *13*(2), 238–241. https://doi.org/https://doi.org/10.1111/j.2517-6161.1951.tb00088.x

Singer, J. A., & Salovey, P. (1988). Mood and memory: Evaluating the network theory of affect. *Clinical Psychology Review*, *8*(2), 211–251. https://doi.org/10.1016/0272-7358(88)90060-8

Skitka, L. J. (2010). The Psychology of Moral Conviction. *Social and Personality Psychology Compass*, *4*(4), 267–281. https://doi.org/10.1111/j.1751-9004.2010.00254.x

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, *88*(6), 895917. https://doi.org/10.1037/0022-3514.88.6.895

Skitka, L. J., Wisneski, D. C., & Brandt, M. J. (2018). Attitude moralization: Probably not intuitive or rooted in perceptions of harm. *Current Directions in Psychological Science*, *27*(1), 913. https://doi.org/10.1177/0963721417727861

Southam-Gerow, M. A., McLeod, B. D., Brown, R. C., Quinoy, A. M., & Avny, S. B. (2011). Cognitive-behavioral therapy for adolescents. In B. B. Brown & M. J. Prinstein (Eds.) (pp. 100–108). San Diego: Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-373951-3.00106-X

Southward, M. W., & Cheavens, J. S. (2018). Identifying core deficits in a dimensional model of borderline personality disorder features: A network analysis. *Clinical Psychological Science*, *6*(5), 685–703. https://doi.org/10.1177/2167702618769560

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011). The complexity of developmental predictions from dual process models. *Developmental Review*, *31*(2), 103–118. https://doi.org/10.1016/j.dr.2011.07.003

Stein, D. J., Phillips, K. A., Bolton, D., Fulford, K. W. M., Sadler, J. Z., & Kendler, K. S. (2010). What is a mental/psychiatric disorder? From DSM-IV to DSM-v. *Psychological Medicine*, *40*(11), 1759–1765. https://doi.org/10.1017/S0033291709992261

Stich, S. (2018). The moral domain. In K. Gray & J. Graham (Eds.) (pp. 547–555).

Theriault, J., Waytz, A., Heiphetz, L., & Young, L. (2017). Examining overlap in behavioral and neural representations of morals, facts, and preferences. *Journal of Experimental Psychology: General*, *146*(11), 1586. https://doi.org/10.1037/xge0000350

Tracy, J. L., Steckler, C. M., & Heltzel, G. (2019). The Physiological Basis of Psychological Disgust and Moral Judgments, 18.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.

Turiel, E. (2006). Thought, emotions, and social interactional processes in moral development. *Handbook of Moral Development*, *2*.

Turiel, E. (2008). The development of children's orientations toward moral, social, and personal orders: More than a sequence in development. *Human Development*, *51*(1), 2139.

Turiel, E., Hildebrandt, C., Wainryb, C., & Saltzstein, H. D. (1991). Judging social issues: Difficulties, inconsistencies, and consistencies. *Monographs of the Society for Research in Child Development*, *56*(2), i116. Retrieved from http://www.jstor.org/stable/1166056

Vaidyanathan, U. (2020). Theory, in practice. *Psychological Inquiry*, *31*(4), 334335.

Van Bavel, J.J., FeldmanHall, O., & Mende-Siedlecki, P. (2015). The neuroscience of moral cognition: From dual processes to dynamic systems. *Current Opinion in Psychology*, *6*, 167172. https://doi.org/10.1016/j.copsyc.2015.08.009

Van Bavel, J.J, Xiao, Y., & Cunningham, W. A. (2012). Evaluation is a Dynamic Process: Moving Beyond Dual System Models: Evaluation is a Dynamic Process. *Social and Personality Psychology Compass*, *6*(6), 438–454. https://doi.org/10.1111/j.1751-9004.2012.00438.x

van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W. J. H., Waldorp, L. J., & Schoevers, R. A. (2015). Association of Symptom Network Structure With the Course of Depression. *JAMA Psychiatry*, *72*(12), 1219. https://doi.org/10.1001/jamapsychiatry.2015.2079

van den Berg, J. W., Smid, W., Kossakowski, J. J., van Beek, D., Borsboom, D., Janssen, E., & Gijs, L. (2020). The Application of Network Analysis to Dynamic Risk Factors in Adult Male Sex Offenders. *Clinical Psychological Science*, *8*(3), 539–554. https://doi.org/10.1177/2167702620901720

Vanaman, M. E., & Chapman, H. A. (2020). Disgust and disgust-driven moral concerns predict support for restrictions on transgender bathroom access. *Politics and the Life Sciences*, *39*(2), 200214. https://doi.org/10.1017/pls.2020.20

Vanaman, M. E., & Chapman, H. A. (2021). *Where (and when) is the harm? Judgments of purity under time pressure. Manuscript in preparation.*

Von Glasersfeld, E. (1991). An exposition of constructivism: Why some like it radical (p. 229238). Springer.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012). A century of gestalt psychology in visual perception II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*(6), 1218–1252. https://doi.org/10.1037/a0029334

Wainryb, C. (1991). Understanding differences in moral judgments: The role of informational assumptions. *Child Development*, *62*(4), 840851.

Wainryb, C., & Turiel, E. (1993). Conceptual and informational features in moral decision making. *Educational Psychologist*, *28*(3), 205218.

Weidman, A. C., Sowden, W. J., Berg, M. K., & Kross, E. (2020). Punish or Protect? How Close Relationships Shape Responses to Moral Violations. *Personality and Social Psychology Bulletin*, *46*(5), 693–708. https://doi.org/10.1177/0146167219873485

Willemsen, P., Newen, A., & Kaspar, K. (2018). A new look at the attribution of moral responsibility: The underestimated relevance of social roles. *Philosophical Psychology*, *31*(4), 595–608. https://doi.org/10.1080/09515089.2018.1429592

Wisneski, D. C., & Skitka, L. J. (2017). Moralization Through Moral Shock: Exploring Emotional Antecedents to Moral Conviction. *Personality and Social Psychology Bulletin*, *43*(2), 139–150. https://doi.org/10.1177/0146167216676479

Wylie, J., Alto, A., Flores-Robles, G., Satter, L., & Gantman, A. (2021). Moralization predicts opposition to incremental social policy. Online Session.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151.