Unreliable Emotions and Ethical Knowledge

James Hutton, University of Edinburgh

[7,272 words + 1,749 words bibliography]

Abstract

How is ethical knowledge possible? One of the most promising answers is the *moral sense view*: we can acquire ethical knowledge through emotional experience. But this view faces a serious problem. Emotions are unreliable guides to ethical truth, frequently failing to fit the ethical status of their objects. This threatens to render the habit of basing ethical beliefs on emotions too unreliable to yield knowledge. I offer a new solution to this problem, with practical implications for how we approach ethical decision-making. I argue that an agent can have a reliable belief-forming habit despite having unreliable emotions, so long as she is suitably attentive to defeaters. In *moderately virtuous* agents, unfitting ethical emotions are frequently accompanied by warning signs, such as negative metaemotions, epistemic feelings, moods, and social cues. An agent who cultivates the attentional skill of noticing these signs can prevent many of her unfitting emotions from translating into false beliefs. For a moderately virtuous agent with this skill, the habit of basing ethical beliefs on emotions will be reliable. The upshot is that emotion-based ethical knowledge is possible even for people whose emotions are unreliable, but only if we cultivate the skill of spotting defeaters.

Keywords

metaethics – moral epistemology – emotion – moral sense – unreliability – moral methodology

1. Introduction

How is ethical knowledge possible? One answer is the *moral sense view*: we can acquire ethical knowledge through emotional experience. Many of us find this view extremely attractive.[1] It identifies a non-inferential route to ethical knowledge, thus avoiding the threat of an infinite regress of moral justification. But, unlike other accounts of foundational ethical knowledge, it does so without positing any mysterious psychological capacities. On the contrary, the moral sense view invokes the familiar suite of emotions that punctuate our ethical lives—emotions like guilt, anger, gratitude, pride, and contempt—with which we're all acquainted. It thus fits nicely with leading empirical accounts of moral psychology, according to which emotions are the main driver of "intuitive" ethical belief-formation.[2] So, to many of us, the moral sense view looks like the best candidate for a non-sceptical moral epistemology that coheres with what we know about the human mind.

But there's a drawback. A belief-forming habit can only generate knowledge if it is reliable. However, our emotions are unreliable guides to ethical truth. As I'll illustrate below, people's emotions are frequently biased or unduly swayed by situational factors. Consequently, our emotions frequently fail to "fit" the ethical status of their objects. It follows that if an agent were to accept all the evaluative impressions her emotions convey, she would thereby form lots of false beliefs. For this reason, the habit of basing ethical beliefs on

[1] See Johnston (2001), Jacobson (2005), Cuneo (2006), Roeser (2011), Roberts (2013, 38–112), McGregor (2015), Tappolet (2016), Milona (2016; 2017), Furtak (2018), Milona and Naar (2020), and Hutton (2022). The weaker view that emotions can non-inferentially justify ethical beliefs is defended by Tolhurst (1990), Deonna (2006), Döring (2007), Kauppinen (2013), Pelser (2014), Sias (2014), Cowan (2016; 2018), Mitchell (2017), and Carter (2020). Historical precursors include Shaftesbury and Hutcheson (see Alanen 2013); Brentano (see Montague 2017); Scheler and Sartre (see Poellner 2016); Wiggins (1977); and McDowell (1985).

[2] For overviews, see Haidt (2012, 1–127), Greene (2013, 28–143), and Woodward (2016). This picture has been attacked by May (2018), but see Kurth (2019) and Kauppinen (2022, sec. 2.1) for convincing rebuttals.

emotions seems doomed to be seriously unreliable and not a means of acquiring ethical knowledge.[3]

In this article, I develop a new solution to this problem. I argue that an agent can have a reliable belief-forming habit despite having unreliable emotions, so long as she is suitably attentive to defeaters. In *moderately virtuous* agents, unfitting ethical emotions are frequently accompanied by warning signs, including negative metaemotions, epistemic feelings, moods, and social cues. An agent who cultivates the attentional skill of noticing these signs can prevent many of her unfitting emotions from translating into false beliefs. For a moderately virtuous agent with this skill, the habit of basing ethical beliefs on emotions when warning signs are absent will be reliable. The upshot is that emotion-based ethical knowledge is possible even for people whose emotions are unreliable, but only if we cultivate the skill of spotting defeaters.

This result has important implications, both practical and metaethical. On the practical front, it has implications for how we conduct ethical inquiry. If I'm right, then intuitive judgments based on emotion deserve a central place in our endeavours to figure out how we should live. But in order to make good use of our emotions, we must cultivate the attentional skill of noticing signs that any given emotion is wide of the mark. On the metaethical front, my argument supports a non-sceptical, empiricist-friendly account of how ethical knowledge is possible. I thus hope to contribute to the broader project of showing how to resist ethical nihilism without resorting to coherentism or relativism, but also without invoking notions like rational intuition.

[3] This problem is pressed by Sinnott-Armstrong (1991, 91), Szigeti (2013), Pelser (2014, 114–16), and Brady (2014, 98–101). Of course, there are other objections one could raise against the moral sense view (see Hutton 2022, 593–96), but they'll have to wait for another day.

Overview: Section 2 fleshes out a framework for emotion-based ethical knowledge by describing an agent with idealized emotions. Section 3 details how we fall short of this ideal, which gives rise to the problem of unreliability. Section 4 uses an observation about defeaters to identify a way of acquiring ethical knowledge on the basis of unreliable emotions. Section 5 identifies a range of defeaters that are accessible to many agents. In doing so, it paints a psychologically realistic picture of a person who acquires ethical knowledge through emotional experience. Section 6 concludes by setting out the article's ramifications.

2. The Sage

Above, I sketched how the unreliability of our emotions poses a problem for the moral sense view. But we'll need a more precise understanding of what gives rise to this problem before we can see our way to a solution. With that goal in mind, I'll begin by characterising an idealized kind of agent who is capable of acquiring emotion-based ethical knowledge. By spelling out how we fall short of this ideal, we'll reach a precise specification of the unreliability problem. In turn, this will reveal the space for a solution.

The idealized agent in question is the figure of the sage (*shèng*) as envisaged in Confucian ethics. In a famous passage in *The Analects*, Kongzi describes what his mind was like once he achieved sagehood: "At seventy, I could follow my heart's desires without overstepping the bounds of propriety" (*Analects* 2.4; 2003, 9). At this final stage of ethical development, the sage's "heart" becomes perfectly attuned to the contours of the ethical landscape. Interpreting a bit, we can think of the sage as having emotions which are always perfectly aligned with the demands of ethics.[4]

[4] Compare Mengzi's ([300BC] 2008) and Wang Yangming's ([1572] 2009) elaborations of the Confucian ideal of sagehood.

One way of cashing this out is to say that the sage's ethical emotions are always *fitting*. Let me explain this terminology. Each token emotion is directed at some object, which we call its *target*. Whether a given emotion is fitting depends on what its target is like. Specifically, each type of emotion is paired with some evaluative property, and a token emotion is fitting iff it is directed at a target which instantiates the corresponding property. For example,

- An agent's guilt is fitting iff the deed about which she feels guilty is a wrongdoing for which she is culpable.
- An agent's admiration is fitting iff the thing she admires is excellent.
- An agent's anger is fitting iff the deed she is angry about is an offence.

As these examples illustrate, some types of emotion are paired with ethical properties (whereas others are paired with non-ethical evaluative properties).[5] Let's call such emotions *ethical emotions*. Applying this framework to the sage: to say that the sage's heart is perfectly attuned to the bounds of ethical propriety is to say that, whenever she experiences an ethical emotion, she does so towards a target which exemplifies the corresponding ethical property.[6]

Now, suppose that every time the sage experiences an ethical emotion, she forms the belief that its target instantiates the corresponding ethical property. E.g., when she experiences anger towards an action, she comes to believe that this action is an offence. Because the sage's emotions are all fitting, the ethical beliefs she forms on the basis of her emotions will all be true. In other words, the sage's habit of basing ethical beliefs on emotions will be a perfectly reliable belief-forming habit. Plausibly, this means that these beliefs amount

[5] Examples of emotions paired with non-ethical evaluative properties might include aesthetic emotions such as amusement and prudential emotions such as fear. I won't take a stand on how to delimit the ethical from the non-ethical; I'll just focus on paradigm cases of the former.

[6] We might interpret some of Aristotle's remarks about virtue in a similar way (e.g. *Nicomachean Ethics*, 1106b; Aristotle 2000, 30).

to ethical knowledge; this will be the case, for example, if we accept a virtue reliabilist account of knowledge.[7]

Note that this description of how ethical knowledge is possible for the sage is more or less neutral about the nature of emotion. It's compatible with any account on which some emotions are intentional, nondoxastic states with ethical fittingness conditions (e.g., perceptual accounts, feeling-towards accounts, attitudinal accounts). Similarly, the characterization is fairly neutral about the ontology of value. It meshes nicely with "fitting attitudes" accounts, but, since it doesn't appeal to any constitutive link between emotion and ethical value, it's available to realists and sophisticated antirealists alike.[8]

It's also worth noting that there's nothing supernatural about the sage. We can give a perfectly unmysterious explanation of her ability to respond to objects with emotions matching their ethical properties. This is because the ethical properties of any object are determined by its nonethical properties, and our emotions are differentially responsive to the nonethical properties of their targets. So long as the sage's patterns of emotional responsiveness match the patterns of nonethical-to-ethical determination, her emotions will fit

[7] See Sosa (1991; 2017). This suggestion is bolstered by various views according to which the content of ethical emotions enables them to justify ethical beliefs. For instance, many theorists argue that experiencing anger towards *x* involves it seeming to one that *x* is an offence (e.g., Roberts 1988, 190–95; Mitchell 2021, 30–69) and that this makes the emotion a suitable epistemic basis for the corresponding ethical belief (Tolhurst 1990, 85–86; Kauppinen 2013, 375–77). Since the issue of whether the content of emotion supports the moral sense view is orthogonal to the issue of reliability, I won't say anything more about it in this article. For further discussion, see Deonna and Teroni (2012, 118–25), Brogaard and Chudnoff (2016), Mitchell (2017), and Harrison (2021).

[8] My talk of reliable habits of ethical belief-formation is compatible with any forms of antirealism sophisticated enough to draw a distinction between true and false ethical attitudes, including modern versions of non-cognitivism. (This point is emphasized by Blackburn (1996, 86–88) and Sinnott-Armstrong (2011, 289).)

the ethical status of their targets.[9] Indeed, this is exactly how we would explain our own limited capacity to respond to objects with emotions that match their ethical properties. The sage doesn't have any special psychological capacities which normal human beings lack; it's just, as I'll argue in a moment, that the sage's emotional dispositions are unrealistically perfect.

3. The Unreliability of Emotions

The previous section's account helps us flesh out how, in principle, an agent could acquire ethical knowledge on the basis of emotion. But most, perhaps all, real human beings fall short of the ideal of sagehood. Unlike the sage, we experience unfitting ethical emotions, and this is far from a rare occurrence.

One reason we experience unfitting emotions is that, unlike the sage, our patterns of emotional response don't perfectly match the patterns of nonethical-to-ethical determination that make up the ethical landscape. Someone raised in a homophobic society will likely grow up to experience unfitting negative emotions towards gay people. For instance, they will tend to experience anger towards public displays of affection between gay people, even though these acts don't instantiate the property of being an offence. Since we all grew up in societies that are ethically flawed to one degree or another, it's reasonable to suppose that none of us have acquired emotional dispositions that match the ethical landscape perfectly, a supposition that's further supported by evidence of how prevalent implicit bias is in our cultures.[10]

---

[9] Compare McBrayer's (2010) account of moral perception.

[10] See, e.g., Brownstein (2018, 2) on the prevalence of implicit bias (though see Machery (2022) for some potential problems with this research paradigm). Baron *et al.* (2014, Experiment 1) illustrates the impact of implicit bias on ethical judgments. It's common among psychologists and philosophers to construe implicit bias as a partly "affective" phenomenon (Banaji and Heiphetz 2010; Gendler 2011; Brownstein and Madva 2012b; 2012a; Madva and Brownstein 2018; Brownstein 2018). Since the "affect" in question is directed at an external

Another reason we experience unfitting emotions is that our emotional dispositions are inherently noisy. Your tendency to experience anger towards a certain remark will depend partly on the remark's ethically relevant features—the content, context, and consequences of what was said. But your emotional response will frequently also be influenced by factors that are extraneous to the remark's ethical status. Such factors include your recent run of positive or negative emotional experiences: if you're having a bad day, you're more likely to be in an irritable mood and get angry about an innocuous remark. Other extraneous factors that affect our ethical emotions seem to include being low on blood-sugar,[11] being in the grips of a hangover,[12] having to listen to irritating sounds,[13] and perhaps even smelling unpleasant odours.[14] All these factors work by changing our moods,[15] as a result of which our emotional dispositions get "temporarily put 'out of tune'".[16] While it's possible, on occasion, for these extraneous factors to put you in just the right mood to respond with a

object, it's plausible that being implicitly biased against a group involves being disposed to experience unfitting negative emotions towards them.

[11] See Danziger *et al.* (2011) for preliminary evidence that being "hangry" can influence one's ethical judgments. All the empirical evidence cited in this paragraph should be treated as provisional, owing to the replication crisis in social psychology.

[12] See Fjær (2015) and Milton *et al.* (2019).

[13] See Seidel and Prinz (2013); compare Mathews and Canon (1975).

[14] See Schnall *et al.* (2008, Experiment 1) and Inbar *et al.* (2012). The validity of the "incidental disgust" research programme has been questioned (e.g., May 2014), but the effect of disgusting smells in particular seems to hold up. As one influential meta-analysis puts it, "gustatory and olfactory disgust inductions exert a reliable, small- to medium-sized effect on moral judgments" (Landy and Goodwin 2015, 529; compare May 2014, 132).

[15] On the influence of recent emotions on mood, see Kontaris *et al.* (2020); on hunger and mood, see Gold *et al.* (1995); on hangovers and mood, see McKinney and Coyle (2006) and Penning *et al.* (2012); on sound and mood, see Smith *et al.* (1997); on odour and mood, see Herz (2009) and Kontaris *et al.* (2020, 7).

[16] This nice phrase comes from Goldie (2004, 257).

fitting emotion to the situation at hand, the general tendency of noise in any system of human judgment is to reduce its reliability.[17] Consequently, the fact that our emotional dispositions are sensitive to these extraneous factors means an increased prevalence of unfitting emotions.

These two effects—on the one hand, shortcomings due to the imperfections in our patterns of emotional response; on the other, noise due to the influence of ethically extraneous factors—work cumulatively to reduce the reliability of our emotions. Therefore, it's hard to deny that most, if not all, human beings experience unfitting ethical emotions fairly frequently. If, like the sage described in the previous section, you or I were to form the corresponding ethical belief every time we experienced an ethical emotion, we would thereby form lots of false beliefs. This means that our habit of forming emotion-based ethical beliefs wouldn't be reliable and *ipso facto* wouldn't be a means of acquiring ethical knowledge. This, in detail, is the problem the moral sense view must overcome.

At this point, one possible move would be to claim that, for all that's been said, there might still be real-life sages among us.[18] There are interpersonal differences in how well people's emotions match the ethical landscape; not all societies are equally plagued by bias, and in any case, individuals' emotional dispositions can be reshaped through habituation. So, there might be some exceptional people whose hearts match the bounds of propriety perfectly. But this way of defending the moral sense view seems extremely perilous. Interpersonal differences notwithstanding, it would be a tall order to make a convincing case that any real-life agent's ethical emotions are perfect (*pace* traditional views about Kongzi and

[17] See Kahneman *et al.* (2021).

[18] Compare Cuneo (2006, 82–85), Pelser (2014, 116), and Sias (2014, 544–47). Another route would be to deny that reliability is a necessary condition for knowledge, e.g., by endorsing epistemological disjunctivism. However, even on this latter route, the unreliability of emotion will resurface as an undercutting defeater for any ethical beliefs based on emotion unless more can be said.

other sages). Theorists in both psychology and philosophy of emotion present the influence of mood as a ubiquitous, structural feature of humans' emotional dispositions.[19] Consequently, the burden of proof lies with anyone who wants to claim that some people's emotions are immune to the influence of extraneous factors. It thus seems unwise to tie the fate of the moral sense view to the existence of real-life sages. Moreover, it's obvious that most of us aren't sages—the moral sense view would lose much of its interest and appeal if it only credited a tiny minority of us with first-hand ethical knowledge (i.e., ethical knowledge that isn't derived from moral testimony). Therefore, would-be defenders of the moral sense view should try to find a way to square emotion-based ethical knowledge with the unreliability of emotion.[20] The upshot will be a nonideal theory of the cognitive habits that make emotion-based knowledge possible. As we'll see, this gives us practical guidance for how real agents can improve their epistemic conduct, guidance that's hidden so long as we're fixated on idealized agents.

4. Improving Reliability by Responding to Defeaters

In this section, I'll describe a kind of agent whose habit of forming emotion-based ethical beliefs is reliable despite the fact that her emotions are unreliable.

It's commonplace, across all cognitive domains, that agents are capable of treating certain cues as defeaters. One treats a cue as a defeater when one responds to it by refraining from forming a belief one would otherwise have formed. (One also treats a cue as a defeater when one responds to it by relinquishing a belief one formed previously, but our focus is on

[19] E.g. Deonna and Teroni (2012, 104–5), Kontaris *et al.* (2020).

[20] Another advantage of the strategy pursued below is that the sage's abilities show up as an edge case of a more general route to emotion-based ethical knowledge (see fn. 21). This increase in generality indicates that my account provides a deeper understanding of the boundaries within which the moral sense view is viable.

belief-formation rather than belief-maintenance.) Although the idea of responding to defeaters is commonplace, it's less commonly observed that treating cues as defeaters can increase the reliability of one's belief-forming habits. Let's illustrate this fact with a toy example.

NIGHT VISION: Amina and Saira both have good eyesight under normal conditions, but terrible night vision: 99% of their visual experiences are veridical during the daytime, but only 50% of them are veridical during the night-time. (For simplicity, assume that Amina and Saira both have an equal number of visual experiences in the daytime and in the night-time, however visual experiences are to be individuated.) Every time Amina has a visual experience, she believes what she sees. Consequently, Amina's habit of forming vision-based beliefs is 74.5% reliable. Every time Saira has a visual experience, she believes what she sees, unless it's night-time (i.e., she treats the presence of night-time as a defeater). Consequently, Saira's habit of forming vision-based beliefs is 99% reliable even though her visual experiences are veridical only 74.5% of the time.

Amina's responsiveness to defeaters filters out all the visual experiences that occur during the night-time, and the pool of visual experiences that remains is more reliable than her total pool of visual experiences. The lesson is as follows: if there are cues which are more likely to accompany a visual experience if it's non-veridical than if it's veridical, and if you treat these cues as defeaters, then you will form a lower proportion of false vision-based beliefs than you would have otherwise. This improves the reliability of your habit of forming vision-based beliefs without requiring any improvement in your vision (see fig. 1).

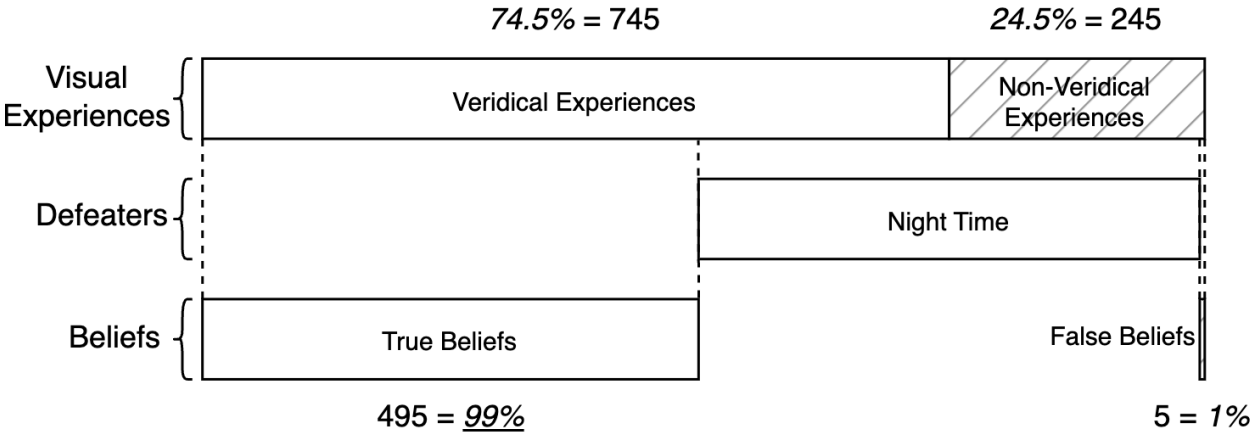

Figure 1. Amina's habit of forming vision-based beliefs while treating the presence of night time as a defeater.

This result has an obvious corollary for emotion-based beliefs: if there are cues which are more likely to accompany an emotion if it's unfitting than if it's fitting, and if you treat these cues as defeaters, then you will form a lower proportion of false emotion-based beliefs than you would have otherwise. This improves the reliability of your habit of forming emotion-based beliefs without requiring any improvement in your emotional dispositions.[21] It follows that there's a kind of agent who is capable of forming reliable emotion-based beliefs despite having unreliable emotions. This will be the case under precisely specifiable conditions, namely so long as

(i) a sufficient proportion of the agent's unfitting emotions are accompanied by accessible cues, such that

(ii) she treats these cues as defeaters, with the result that

[21] There are other contexts in which parallel lessons can be drawn. For instance, this observation about defeaters and reliability contains materials for defending nonreductionism about testimonial justification against the charge of endorsing gullibility (see Fricker 1994). But such applications must wait for another day.

(iii) the proportion of fitting emotions among her remaining pool of undefeated emotions equals or exceeds the proportion of true beliefs a belief-forming habit must produce in order to count as reliable.

(By an "accessible" cue, I mean one it is psychologically possible for the agent to treat as a defeater. Plausibly, this means any cue it is psychologically possible for the agent to attend to.)

Let's use the phrase *moderately virtuous agent* to denote any agent whose unfitting emotions are adequately covered by accessible cues. For a moderately virtuous agent who possesses the attentional skill of noticing the relevant cues and treating them as defeaters, her habit of basing ethical beliefs on emotions will be reliable. Plausibly, this means that some of her emotion-based beliefs will amount to ethical knowledge. (For example, if we accept a virtue reliabilist account of knowledge, then those of her ethical beliefs which are correct in virtue of stemming from this reliable belief-forming habit will constitute knowledge.) This, in schematic form, is how we can uphold the moral sense view while acknowledging the unreliability of emotion.[22]

Nothing I've said so far indicates that any of *us* are moderately virtuous agents. It's less perilous to stake the moral sense view on the existence of moderately virtuous agents than on the existence of sages, but more needs to be said to flesh out the moral psychology underlying this account. This is the next section's task: to fill in the details, making it plausible that emotion-based ethical knowledge is within reach for human beings.

[22] The sage's capacity for ethical knowledge turns out to be an edge case of this more general schema. Since all the sage's emotions are fitting emotions, it's trivially true that a sufficient proportion of her undefeated emotions are fitting to meet the threshold of reliability.

5. Defeaters for Ethical Emotions

If enough of our unfitting ethical emotions are accompanied by accessible cues which indicate unfittingness, then we can achieve ethical knowledge by treating these cues as defeaters. But what kinds of cue fit the bill? This is a partly empirical matter and there will be variation from person to person. Indeed, there's no *a priori* guarantee that a person's emotions will be adequately covered by accessible cues. Nevertheless, there are a range of cues we can point to which plausibly, for many agents, are more likely to accompany unfitting than fitting emotions. By laying out these cues, we can paint a psychologically realistic picture of a moderately virtuous human being—one who is capable of acquiring ethical knowledge on the basis of her unreliable emotions.

5.1 Negative Metaemotions

The first kind of cue is provided by *metaemotions*. A metaemotion is an emotion, the target of which is another emotion.[23] For example, I might feel angry about a remark someone has made, but also feel embarrassed about getting angry. It's common sense to use negatively valenced metaemotions as defeaters for first-order ethical emotions; this is something most of us probably do already; if I feel embarrassed to find myself getting angry, I'm less likely to endorse the evaluative impression the anger conveys. Moreover, for most agents, negative metaemotions are more likely to accompany unfitting first-order emotions than fitting ones. In the same way that one's first-order emotions intelligently combine multiple streams of information to produce holistic evaluative impressions,[24] one's metaemotions are likely to be sensitive to a range of factors that bear on the appropriateness of the first-order emotion. Of course, there's no reason our metaemotions should be any less fallible than our first-order

[23] See Howard (2017) and Mitchell (2019a).

[24] See Allman and Woodward (2008) and Railton (2014). Compare Lacewing (2005).

emotions, and some agents might have totally unreliable metaemotions. (Remember that I'm not trying to identify cues that come with an *a priori* guarantee of being useful, or even available, to all agents.) Nevertheless, for agents with moderately reliable first-order emotions and moderately reliable metaemotions, treating negative metaemotions as a defeater will lead to a remarkable boost in reliability, as illustrated by the following case.

AMINA'S METAEMOTIONS: Amina has moderately reliable ethical emotions. To be precise, 80% of her ethical emotions are fitting while the rest are unfitting. She also has moderately reliable metaemotions. 80% of her unfitting first-order emotions are accompanied by negative metaemotions, while 20% of her negative metaemotions are misleading, in the sense that they accompany fitting first-order emotions. By default, when Amina experiences an ethical emotion, she forms the corresponding ethical belief. However, she treats negative metaemotions as defeaters. The upshot is that, for every thousand ethical emotions she experiences, she forms 760 true ethical beliefs and 40 false ones. Consequently, her habit of forming emotion-based ethical beliefs is 95% reliable even though her emotions are fitting only 80% of the time.

By treating negative metaemotions as defeaters, Amina filters out all the emotions that are accompanied by this cue. The remaining pool of undefeated emotions is more reliable than her total pool of emotions. Notice that Amina's metaemotions are no more reliable than her first-order emotions: they too are only 80% reliable. Nevertheless, even by attending to this one kind of cue, a fallible agent like Amina can significantly enhance the reliability of her habit of forming ethical beliefs based on her emotions (see fig. 2).

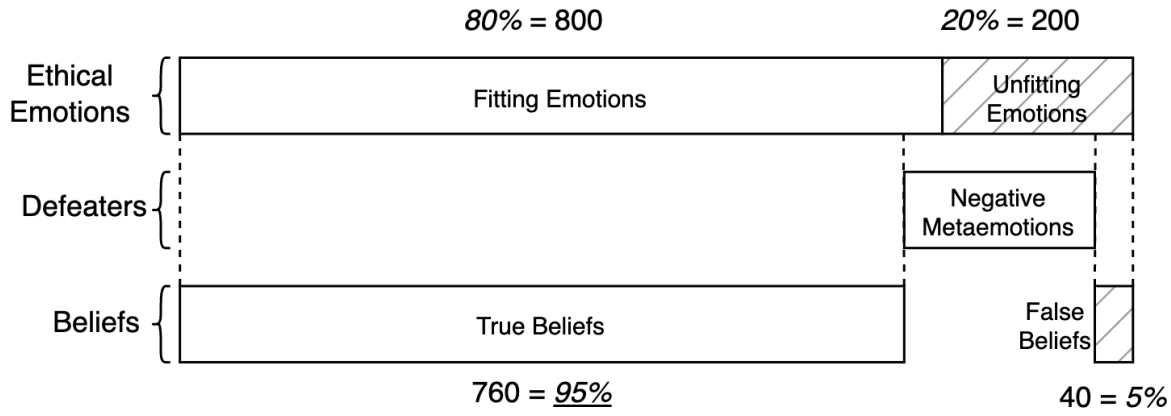

Figure 2. Amina's habit of forming emotion-based ethical beliefs while treating negative metaemotions as defeaters.

5.2 Epistemic Feelings

The second kind of cue is provided by *epistemic feelings*, such as feelings of uncertainty, hesitancy, or doubt.[25] For example, one might feel angry about someone's remark but simultaneously feel a sense of uncertainty about whether the remark really constitutes an offence. As with negative metaemotions, it's common sense to treat negative epistemic feelings as defeaters, and most of us probably do this already. It's also plausible that, for most agents, a negative epistemic feeling will be a fallible but useful sign that one should withhold judgment. This is because the capacity to experience epistemic feelings is yet another way in which the mind combines multiple streams of information to reach a useful (though fallible) impression of how things are going. Your feeling of uncertainty will result from various kinds of information that bear on the likelihood that you would go wrong if you went ahead and made the judgment.[26] Once again, there are no guarantees here—agents will differ in the extent to which their unfitting ethical emotions are accompanied by feelings of uncertainty, and some agents may have thoroughly misleading epistemic feelings. Nevertheless, agents

[25] See, e.g., Arango-Muñoz (2014) and Carruthers (2017).

[26] See Carruthers (2017, 70–72).

whose epistemic feelings are moderately reliable will be able to increase the reliability of their belief-forming habits still further by treating them as defeaters.

What is the cumulative effect of treating both negative metaemotions and epistemic feelings as defeaters? So long as, firstly, both are more likely to accompany an emotion if it is unfitting than if it is fitting and, secondly, some unfitting emotions accompanied by epistemic feelings aren't also accompanied by negative metaemotions, treating both as defeaters will provide a further boost in reliability. To illustrate, let's develop our case further.

AMINA'S METAEMOTIONS AND EPISTEMIC FEELINGS: Amina's psychological profile is as described in AMINA'S METAEMOTIONS, with the addition that she has moderately reliable epistemic feelings: 80% of her unfitting first-order emotions are accompanied by negative epistemic feelings, while 20% of her negative epistemic feelings are misleading (i.e., accompany fitting first-order emotions). There's a lot of overlap between Amina's epistemic feelings and her metaemotions, but the correlation isn't perfect: 1 in 5 of her negative epistemic feelings accompanies an emotion towards which she doesn't experience a negative metaemotion. Amina treats both negative metaemotions and negative epistemic feelings as defeaters. The upshot is that, for every thousand ethical emotions she experiences, 800 are fitting, of which 40 are accompanied by negative metaemotions, 40 are accompanied by negative epistemic feelings, and 48 are accompanied by either a negative metaemotion or a negative epistemic feeling. Meanwhile, of her 200 unfitting ethical emotions, 160 are accompanied by negative metaemotions, 160 are accompanied by negative epistemic feelings, and 192 are accompanied by either a negative metaemotion or a negative epistemic feeling. Consequently, Amina forms 752 true ethical beliefs and 8 false ones: her belief-forming habit is 98.9% reliable.

By attending to her negative epistemic feelings, Amina winnows out still more unfitting emotions. This makes her remaining pool of undefeated emotions still more reliable (see fig. 3).

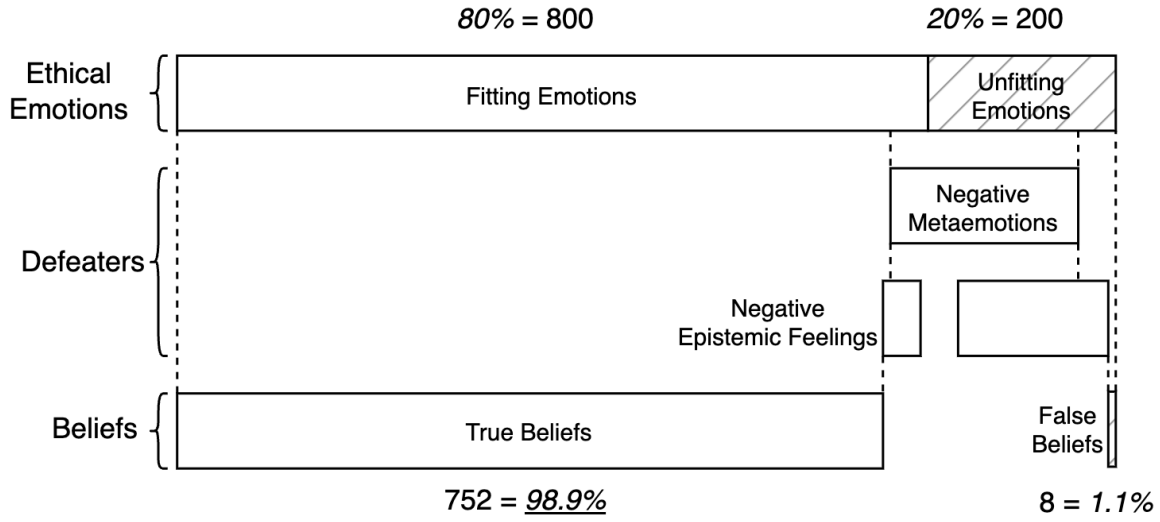

Figure 3. Amina's habit of forming emotion-based ethical beliefs, while treating negative metaemotions and epistemic feelings as defeaters.

With these fallible but useful indicators of unfittingness layered one on top of the other, our agent's reliability creeps up still further. This cumulative effect of fallible defences has been summed up nicely in another context with the image of "slices of Swiss cheese being layered on top of one another, until there [are] no holes you [can] see through."[27] It isn't a problem that each slice has lots of holes in it; as we add more layers, we close the gaps.

5.3 Unintelligible Moods

The next kind of cue is provided by *moods that aren't rationally intelligible*. Sometimes we experience our moods as intelligible responses to how things are going.[28] Lixin is backpacking in a foreign country, and he has experienced all kinds of mishaps—he's been mugged; people

[27] See Lewis (2021, 76). The original context is pandemic mitigation.

[28] See Mitchell (2019b, 127–32). Here, I have in mind what Mitchell calls the "normative sense" of intelligibility, as opposed to "merely causal" intelligibility.

have tricked him out of money; strangers have been rude and dismissive. This puts him in a certain mood: he feels ill at ease. But as well as feeling ill at ease, Lixin is aware of various features of his environment in virtue of which it makes sense for him to feel ill at ease, namely the aforementioned mishaps. Consequently, Lixin experiences his mood as an intelligible response to the hostility of his present environment. Contrast this with cases in which one's mood is *not* experienced as rationally intelligible. When I just wake up feeling unaccountably irritable or find myself growing irritable due to the incessant whine of my neighbour's lawnmower, there is nothing I'm aware of in virtue of which it makes sense for me to approach my environment with a diffuse sense of irritation. Lixin's mood is rationally intelligible; mine is rationally unintelligible.

In many cases, a rationally unintelligible mood will serve as a fallible but useful indicator that one's emotions are less likely than usual to be fitting. So long as the pattern of mishaps that have made Lixin feel ill at ease are manifestations of a genuine tendency in his environment (rather than just a run of bad luck), the shift in emotional dispositions brought on by his mood will increase his tendency to experience fitting emotions. For instance, he will be more likely to view a hostel roommate's ambiguous behaviour with suspicion, right at a time when he finds himself in an environment where suspicion is more likely to be fitting. In contrast, my rationally unintelligible mood, brought on by the sound of the lawnmower or some other ethically irrelevant factor, will simply put me out of tune with the ethical landscape. I'm more disposed to experience anger, but there's nothing about my environment that makes genuine offences more likely. In other words, the unintelligible mood is one in which I'm at an increased risk of having an emotion that results from an extraneous factor. It follows that, so long as one's sense of the intelligibility of one's mood is moderately reliable, the presence of a rationally unintelligible mood will serve as a useful though fallible signal that one's emotions are more likely than usual to be unfitting.

It's part of folk wisdom about the limitations of emotion that certain moods put our emotions temporarily out of tune, and thus render us ill-equipped for making ethical judgments or other important evaluative decisions. So, it's fairly plausible that many of us are already in the habit of opting to "sleep on it" or "mull things over" when we find ourselves in moods that don't make rational sense from the inside. On the other hand, some of us might need to practice the attentional skill of noticing such moods in order to make full use of this kind of cue. Either way, unintelligible moods add to the growing list of cues which it's feasible—and advisable—to treat as defeaters. I won't run any more numbers, but as long as, firstly, Amina's sense of being in an unintelligible mood is more likely to accompany an unfitting than a fitting emotion and, secondly, it accompanies some unfitting emotions which aren't accompanied by negative metaemotions or epistemic feelings, she will be able to increase the reliability of her belief-forming habit still further by treating it as a defeater.

5.4 Conflicting Beliefs

Another important kind of cue is provided by *clashes with existing beliefs*. One example of this would be feeling angry about some piece of conduct while seeing no relevant difference between this and another piece of conduct you believe to be innocuous.[29] Another example would be feeling angry about some piece of conduct while believing that your emotions aren't reliable in this context because the agent belongs to a group you suspect you're emotionally biased against.

[29] Here, you would be drawing on the arguably conceptual truth that two actions can't differ in ethical status unless they differ in ethically relevant nonethical properties. Proponents of the moral sense view typically allow that "non-substantive (conceptual or analytic)" ethical truths like this can be known by means other than emotion (Milona 2016, 898–99; Kauppinen 2013, 373).

Once again, it's commonsense to treat such clashes as defeaters. I suspect that virtually all of us already have a good nose for conflicts between our ethical emotions and our background ethical beliefs.[30] Moreover, by the time we reach adulthood, most of us are walking around with a large stock of mainly accurate ethical beliefs about particular cases, plus some mainly accurate generalizations about the *prima facie* ethical properties of various types of action and situation. As long as one's background ethical beliefs are moderately accurate, an emotion will be more likely to conflict with them if it is unfitting than if it is fitting. So, for most of us, the ability to treat these clashes as defeaters will enhance the reliability of our emotion-based beliefs. (Of course, there are no guarantees here. An obvious exception would be anyone who has been brought up to believe a pervasively defective moral ideology; here as elsewhere, bad ideology can impede one's ability to acquire knowledge.)

We might wonder where an agent could have got this background web of true ethical beliefs against which to check her emotions. On a modest version of the moral sense view—according to which emotion is just one of several routes we have to noninferential ethical knowledge—one suggestion would be that these true ethical beliefs trace back to some non-emotional source, such as rational intuition. More interestingly, on an ambitious version of the moral sense view—according to which the whole superstructure of ethical knowledge rests on emotion—an agent's background ethical beliefs would stem from past emotions, plus chains of reasoning and/or testimony tracing back to emotion. On this latter picture, a kind of bootstrapping would be possible, in which an initial batch of moderately unreliable ethical beliefs enable one to filter out some of one's unfitting emotions, leading to a less unreliable second batch of ethical beliefs, and so on in a virtuous circle. As Richard Boyd once argued,

[30] For a general account of how we maintain coherence in thought, see Thagard (2000, 15–40). For applications of this to emotion-based ethical judgment, see Thagard (2006, 135–56) and Milona (2016, 903–5). See also Campbell and Kumar (2012).

"In [science], improvements in knowledge can be expected to produce improvements in method […], and there is no reason to exclude this possibility in the moral case" (1988, 182).

5.5 Conversational Cues

Lastly, let me note another rich source of defeaters which is normally available to us when we're forming ethical judgments: feedback from others. We often make ethical judgments and decisions in contexts of social interaction, and we're highly attuned to our interlocutors' reactions to what we are saying and to the emotions we express. Consequently, in addition to the introspectable cues described up to this point, we very often have access to a range of social cues too. These social cues include the emotions your interlocutor expresses (verbally and nonverbally) towards the situation under consideration; the metaemotions your interlocutor expresses towards your emotions; the epistemic feelings your interlocutor expresses towards the thoughts you put forward; your interlocutor's sense of whether your present mood is rationally intelligible; and the interlocutor's beliefs about the fittingness of your emotions and the tenability of the ethical judgments you're considering. So long as your interlocutor's emotions, metaemotions, etc. are moderately reliable, all of these signals will be valuable defeaters.

Having access to your interlocutor's responses is valuable even if your interlocutor's emotions, metaemotions, etc. are no more reliable than your own. Your interlocutor simply has a different perspective on the situation from you, both literally and figuratively, so there will be cases in which something about the situation leads you to experience an undefeated unfitting emotion but doesn't have the same effect on your interlocutor. Suppose that for any given unfitting emotion, you and your interlocutor each have an independent 95% chance of experiencing one of the previously mentioned defeaters. In that case, the chance that *at least one of you* experiences such a defeater rises to 99.75%. In this situation, if you treat your

interlocutor's conflicting emotions, metaemotions, etc. as defeaters as well as your own, you will hardly form any false emotion-based beliefs.

The dynamics of such emotional-ethical exchanges are a ripe topic for future research.[31] This is a situation in which we make epistemic use of non-linguistic as well as linguistic information from others. Moreover, if treating your interlocutor's reactions as defeaters is necessary for reaching the threshold of reliability, then we have an interesting new case in which the acquisition of knowledge (not just its testimonial transmission) is an inherently social endeavour.[32] For our purposes, the takeaway is simply that it is very plausible that cues from one's interlocutor make it dramatically more feasible to elevate the reliability of one's belief-forming habit to the threshold required for knowledge.

5.6 How to Fall Short of Emotion-Based Ethical Knowledge

Let me close this section by highlighting the three ways agents can fall short of acquiring ethical knowledge through emotional experience. Firstly, an agent's emotions might be so thoroughly misaligned with the ethical landscape that no amount of caution will make her belief-forming habit reliable; this will be the case for anyone who has internalized a thoroughly defective ethical outlook to such an extent that her emotions are pervasively unfitting. Secondly, an agent might have moderately reliable emotions but have too few accessible cues, or cues too poorly correlated with unfittingness; for an agent whose unfitting

[31] The resulting socialized version of the moral sense view aligns with the emotion-centric methodology championed by feminist advocates of consciousness raising in the 1960s. As Kathie Sarachild put it in her address to the First National Women's Liberation Conference in 1968: "We assume that our feelings are telling us something from which we can learn... that our feelings mean something worth analysing... that our feelings are saying something political. [...] In our groups, let's share our feelings and pool them. [...] Our feelings will lead us to ideas and then to actions" (1969, 78).

[32] See Levy's (2007) notion of "radically socialized knowledge."

emotions are insufficiently covered by metaemotions, epistemic feelings, etc., there is no attentional skill she could learn that would transform her habit of forming emotion-based ethical beliefs into a capacity for ethical knowledge. Thirdly, even if an agent's unfitting emotions are adequately covered by accessible cues, she might lack the attentional skill of noticing these and treating them as defeaters; until this agent learns the kind of cognitive skills exhibited by Amina, her emotion-based ethical beliefs will not amount to knowledge.

So, as I've sought to make clear throughout this section, there is no *a priori* guarantee that the emotion-based ethical beliefs of any given agent amount to ethical knowledge. However, by setting out this list, I hope to have provided sufficient reason for concluding that it is within the power of unexceptional people to form reliable ethical beliefs on the basis of emotion, by cultivating the skill of noticing and responding to defeaters.

7. Conclusion

Most, perhaps all, human beings have unreliable ethical emotions. Nevertheless, moderately virtuous agents with the skill of spotting defeaters can form reliable ethical beliefs on the basis of emotion. This means that we can acquire emotion-based ethical knowledge despite the imperfections of our emotions, so long as we learn to navigate those imperfections through responsiveness to defeaters. This conclusion has important practical implications. Intuitive judgments based on emotion deserve a central place in our endeavours to figure out how we should live. However, in order to make good use of our emotions, we must cultivate the attentional skill of spotting defeaters. I've suggested an initial list of defeaters we should learn to look out for, but it would be interesting and worthwhile for other researchers to take up the task of extending and refining this list. It's tempting to imagine that, by means of interdisciplinary work, it should be feasible to perfect the list of defeaters; to develop strategies

for teaching people to notice their presence; and to test the claim that this approach brings reliable emotion-based judgment within reach.[33]

My argument has important philosophical implications too. The unreliability of emotion is one of the main reasons many philosophers are resistant to the moral sense view. By offering a solution to this problem, I hope to have made this already promising view even more attractive. As the objections to the moral sense view fall away, we move closer towards a satisfying account of ethical knowledge that fits with the visceral, imperfect realities of life as a human being.[34]

References

Alanen, Lilli. 2013. 'Emotions in the Seventeenth and Eighteenth Century'. In *Sourcebook for the History of the Philosophy of Mind*, edited by Simo Knuuttila and Juha Sihvola, 499–533. Dordrecht: Springer.

Allman, John, and Jim Woodward. 2008. 'What Are Moral Intuitions and Why Should We Care about Them? A Neurobiological Perspective'. *Philosophical Issues* 18: 164–85.

Arango-Muñoz, Santiago. 2014. 'The Nature of Epistemic Feelings'. *Philosophical Psychology* 27 (2): 193–211. https://doi.org/10.1080/09515089.2012.732002.

Aristotle. 2000. *Nicomachean Ethics*. Edited by Roger Crisp. Cambridge: Cambridge University Press.

Banaji, Mahzarin, and Larisa Heiphetz. 2010. 'Attitudes'. In *Handbook of Social Psychology*, edited by Susan Fiske, Daniel Gilbert, and Gardner Lindzey, 348–88. New York: Wiley.

Baron, Andrew Scott, Yarrow Dunham, Mahzarin Banaji, and Susan Carey. 2014. 'Constraints on the Acquisition of Social Category Concepts'. *Journal of Cognition and Development* 15 (2): 238–68. https://doi.org/10.1080/15248372.2012.742902.

Blackburn, Simon. 1996. 'Securing the Nots: Moral Epistemology for the Quasi-Realist'. In *Moral Knowledge? New Readings in Moral Epistemology*, edited by Mark Timmons and Walter Sinnott-Armstrong, 82–100. Oxford: Oxford University Press.

Boyd, Richard. 1988. 'How to Be a Moral Realist'. In *Essays on Moral Realism*, edited by Geoffrey Sayre-McCord, 181–228. Ithica, NY: Cornell University Press.

Brady, Michael S. 2014. *Emotional Insight: The Epistemic Role of Emotional Experience*. Oxford: Oxford University Press.

Brogaard, Berit, and Elijah Chudnoff. 2016. 'Against Emotional Dogmatism'. *Philosophical Issues* 26 (1): 59–77. https://doi.org/10.1111/phis.12076.

Brownstein, Michael. 2018. *The Implicit Mind: Cognitive Architecture, the Self and Ethics*. Oxford: Oxford University Press.

[33] See Sinnott-Armstrong (2011) for suggestions regarding how to test the reliability of moral judgment in the absence of consensus about what the ethical truths are.

[34] [Acknowledgments]

Brownstein, Michael, and Alex Madva. 2012a. 'The Normativity of Automaticity'. *Mind & Language* 27 (4): 410–34. https://doi.org/10.1111/j.1468-0017.2012.01450.x.
———. 2012b. 'Ethical Automaticity'. *Philosophy of the Social Sciences* 42 (1): 68–98. https://doi.org/10.1177/0048393111426402.
Campbell, Richmond, and Victor Kumar. 2012. 'Moral Reasoning on the Ground'. *Ethics* 122 (2): 273–312. https://doi.org/10.1086/663980.
Carruthers, Peter. 2017. 'Are Epistemic Emotions Metacognitive?' *Philosophical Psychology* 30 (1–2): 58–78. https://doi.org/10.1080/09515089.2016.1262536.
Carter, J. Adam. 2020. 'Epistemic Perceptualism, Skill and the Regress Problem'. *Philosophical Studies* 177 (5): 1229–54. https://doi.org/10.1007/s11098-019-01243-x.
Confucius. 2003. *Analects, with Selections from Traditional Commentaries*. Translated by Edward Slingerland. Indianapolis, IN: Hackett.
Cowan, Robert. 2016. 'Epistemic Perceptualism and Neo-Sentimentalist Objections'. *Canadian Journal of Philosophy* 46 (1): 59–81.
———. 2018. 'Epistemic Sentimentalism and Epistemic Reason-Responsiveness'. In *Evaluative Perception*, edited by Anna Bergqvist and Robert Cowan, 219–36. Oxford: Oxford University Press.
Cuneo, Terence. 2006. 'Signs of Value: Reid on the Evidential Role of Feelings in Moral Judgement'. *British Journal for the History of Philosophy* 14 (1): 69–91. https://doi.org/10.1080/09608780500449164.
Danziger, Shai, Jonathan Levav, and Liora Avnaim-Pesso. 2011. 'Extraneous Factors in Judicial Decisions.' *Proceedings of the National Academy of Sciences of the United States of America* 108 (17): 6889–92.
Deonna, Julien A. 2006. 'Emotion, Perception and Perspective'. *Dialectica* 60 (1): 29–46. https://doi.org/10.1111/j.1746-8361.2005.01031.x.
Deonna, Julien A., and Fabrice Teroni. 2012. *The Emotions: A Philosophical Introduction*. London: Routledge.
Döring, Sabine. 2007. 'Seeing What to Do: Affective Perception and Rational Motivation'. *Dialectica* 61: 363–94.
Fjær, Eivind Grip. 2015. 'Moral Emotions the Day After Drinking'. *Contemporary Drug Problems* 42 (4): 299–313. https://doi.org/10.1177/0091450915604988.
Fricker, Elizabeth. 1994. 'Against Gullibility'. In *Knowing from Words: Western and Indian Philosophical Analysis of Understanding and Testimony*, edited by Bimal Matilal and Arindam Chakrabarti, 125–61. Dordrecht: Kluwer.
Furtak, Rick. 2018. *Knowing Emotions: Truthfulness and Recognition in Affective Experience*. Oxford: Oxford University Press.
Gendler, Tamar. 2011. 'On the Epistemic Costs of Implicit Bias'. *Philosophical Studies* 156 (1): 33–63.
Gold, Ann E., Kenneth M. MacLeod, Brian M. Frier, and Ian J. Deary. 1995. 'Changes in Mood During Acute Hypoglycemia in Healthy Participants'. *Journal of Personality and Social Psychology* 68 (3): 498–504. https://doi.org/10.1037/0022-3514.68.3.498.
Goldie, Peter. 2004. 'Emotion, Reason, and Virtue'. In *Emotion, Evolution, and Rationality*, edited by Dylan Evans and Pierre Cruse. Oxford: Oxford University Press. https://doi.org/10.1093/acprof.
Greene, Joshua. 2013. *Moral Tribes: Emotion, Reason, and the Gap between Us and Them*. New York: Penguin.
Haidt, Jonathan. 2012. *The Righteous Mind*. London: Penguin.
Harrison, Eilidh. 2021. 'The Prospects of Emotional Dogmatism'. *Philosophical Studies* 178 (8): 2535–55.

Herz, Rachel. 2009. 'Aromatherapy Facts and Fictions: A Scientific Analysis of Olfactory Effects on Mood, Physiology and Behavior'. *International Journal of Neuroscience* 119 (2): 263–90. https://doi.org/10.1080/00207450802333953.

Howard, Scott. 2017. 'Metaemotional Intentionality'. *Pacific Philosophical Quarterly* 98 (3): 406–28. https://doi.org/10.1111/papq.12093.

Hutton, James. 2022. 'Moral Experience: Perception or Emotion?' *Ethics* 132 (3): 570–97. https://doi.org/10.1086/718079.

Inbar, Yoel, David Pizarro, and Paul Bloom. 2012. 'Disgusting Smells Cause Decreased Liking of Gay Men'. *Emotion* 12 (1): 23–27. https://doi.org/10.1037/a0023984.

Jacobson, Daniel. 2005. 'Seeing by Feeling: Virtues, Skills, and Moral Perception'. *Ethical Theory and Moral Practice* 8 (4): 387–409. https://doi.org/10.1007/s10677-005-8837-1.

Johnston, M. 2001. 'The Authority of Affect'. *Philosophy and Phenomenological Research* 63 (1): 181–214. https://doi.org/10.1111/j.1933-1592.2001.tb00097.x.

Kahneman, Daniel, Olivier Sibony, and Cass Sunstein. 2021. *Noise: A Flaw in Human Judgment*. London: William Collins.

Kauppinen, Antti. 2013. 'A Humean Theory of Moral Intuition'. *Canadian Journal of Philosophy* 43 (3): 360–81.

———. 2022. 'Moral Sentimentalism'. In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.

Kontaris, Ioannis, Brett East, and Donald Wilson. 2020. 'Behavioral and Neurobiological Convergence of Odor, Mood and Emotion: A Review'. *Frontiers in Behavioral Science* 14. https://doi.org/10.3389/fnbeh.2020.00035.

Kurth, Charlie. 2019. 'What Sentimentalists Should Say about Emotion'. *Behavioral and Brain Sciences* 42: e158. https://doi.org/10.1017/S0140525X18002601.

Lacewing, Michael. 2005. 'Emotional Self-Awareness and Ethical Deliberation'. *Ratio* 18 (1): 65–81. https://doi.org/10.1111/j.1467-9329.2005.00271.x.

Landy, Justin, and Geoffrey Goodwin. 2015. 'Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence'. *Perspectives on Psychological Science* 10 (4): 518–36.

Levy, Neil. 2007. 'Radically Socialized Knowledge and Conspiracy Theories'. *Episteme* 4 (2): 181–92. https://doi.org/10.3366/epi.2007.4.2.181.

Lewis, Michael. 2021. *The Premonition: A Pandemic Story*. London: Penguin.

Machery, Edouard. 2022. 'Anomalies in Implicit Attitudes Research'. *WIREs Cognitive Science* 13 (1): e1569. https://doi.org/10.1002/wcs.1569.

Madva, Alex, and Michael Brownstein. 2018. 'Stereotypes, Prejudice, and the Taxonomy of the Implicit Social Mind'. *Noûs* 52 (3): 611–44. https://doi.org/10.1111/nous.12182.

Mathews, Kenneth E., and Lance Kirkpatrick Canon. 1975. 'Environmental Noise Level as a Determinant of Helping Behavior'. *Journal of Personality and Social Psychology* 32 (4): 571–77. https://doi.org/10.1037/0022-3514.32.4.571.

May, Joshua. 2014. 'Does Disgust Influence Moral Judgment?' *Australasian Journal of Philosophy* 92 (1): 125–41. https://doi.org/10.1080/00048402.2013.797476.

———. 2018. *Regard for Reason in the Moral Mind*. Oxford: Oxford University Press.

McBrayer, Justin. 2010. 'Moral Perception and the Causal Objection'. *Ratio* 23 (3): 291–307. https://doi.org/10.1111/j.1467-9329.2010.00468.x.

McDowell, John. 1985. 'Values and Secondary Qualities'. In *Morality and Objectivity*, edited by Ted Honderich, 110–29. London: Routledge & Kegan Paul.

McGregor, Rafe. 2015. 'Making Sense of Moral Perception'. *Ethical Theory and Moral Practice* 18 (4): 745–58. https://doi.org/10.1007/s10677-015-9601-9.

McKinney, Adele, and Kieran Coyle. 2006. 'Alcohol Hangover Effects on Measures of Affect the Morning after a Normal Night's Drinking'. *Alcohol and Alcoholism* 41 (1): 54–60. https://doi.org/10.1093/alcalc/agh226.

Mengzi. (300BC) 2008. *Mengzi: With Selections from Traditional Commentaries*. Translated by Bryan W. Van Norden. Indianapolis: Hackett.

Milona, Michael. 2016. 'Taking the Perceptual Analogy Seriously'. *Ethical Theory and Moral Practice* 19 (4): 897–915. https://doi.org/10.1007/s10677-016-9716-7.

———. 2017. 'Intellect versus Affect: Finding Leverage in an Old Debate'. *Philosophical Studies* 174 (9): 2251–76. https://doi.org/10.1007/s11098-016-0797-x.

Milona, Michael, and Hichem Naar. 2020. 'Sentimental Perceptualism and the Challenge from Cognitive Bases'. *Philosophical Studies* 177 (10): 3071–96. https://doi.org/10.1007/s11098-019-01360-7.

Milton, Ian James, Elizabeth Sillence, and Melanie Mitchell. 2019. 'Exploring the Emotional Experiences of Alcohol Hangover Syndrome in Healthy UK-Based Adults'. *Drugs: Education, Prevention and Policy* 27 (3): 248–60. https://doi.org/10.1080/09687637.2019.1654431.

Mitchell, Jonathan. 2017. 'The Epistemology of Emotional Experience'. *Dialectica* 71 (1): 57–84.

———. 2019a. 'Understanding Meta-Emotions: Prospects for a Perceptualist Account'. *Canadian Journal of Philosophy*, 1–19. https://doi.org/10.1017/can.2019.47.

———. 2019b. 'The Intentionality and Intelligibility of Moods'. *European Journal of Philosophy* 27 (1): 118–35. https://doi.org/10.1111/ejop.12385.

———. 2021. *Emotion as Feeling Towards Value: A Theory of Emotional Experience*. Oxford: Oxford University Press.

Montague, Michelle. 2017. 'A Contemporary View of Brentano's Theory of Emotion'. *The Monist* 100 (1): 64–87.

Pelser, Adam. 2014. 'Emotion, Evaluative Perception, and Epistemic Justification'. In *Emotion and Value*, edited by Sabine Roeser and Cain Todd, 107–23. Oxford: Oxford University Press.

Penning, Renske, Adele McKinney, and Joris Verster. 2012. 'Alcohol Hangover Symptoms and Their Contribution to the Overall Hangover Severity'. *Alcohol and Alcoholism* 47 (3): 248–52. https://doi.org/10.1093/alcalc/ags029.

Poellner, Peter. 2016. 'Phenomenology and the Perceptual Model of Emotion'. *Proceedings of the Aristotelian Society* 141 (3): 261–88.

Railton, Peter. 2014. 'The Affective Dog and Its Rational Tale: Intuition and Attunement'. *Ethics* 124 (4): 813–59. https://doi.org/10.1086/675876.

Roberts, Robert. 1988. 'What an Emotion Is: A Sketch'. *The Philosophical Review* 97 (2): 183–209.

———. 2013. *Emotions in the Moral Life*. Cambridge: Cambridge University Press.

Roeser, Sabine. 2011. *Moral Emotions and Intuitions*. Basingstoke: Palgrave Macmillan.

Sarachild, Kathie. 1969. 'A Program for Feminist "Consciousness Raising"'. In *Notes from the Second Year: Women's Liberation. Major Writings of the Radical Feminists*, edited by Shulamith Firestone, 78–80. New York, NY: Radical Feminism.

Schnall, Simone, Jonathan Haidt, Gerald Clore, and Alexander Jordan. 2008. 'Disgust as Embodied Moral Judgment.' *Personality & Social Psychology Bulletin* 34 (8): 1096–1109.

Seidel, Angelika, and Jesse Prinz. 2013. 'Sound Morality: Irritating and Icky Noises Amplify Judgments in Divergent Moral Domains'. *Cognition* 127 (1): 1–5. https://doi.org/10.1016/j.cognition.2012.11.004.

Sias, James. 2014. 'Ethical Intuitionism and the Emotions: Toward an Empirically Adequate Moral Sense Theory'. *The Journal of Value Inquiry* 48 (3): 533–49. https://doi.org/10.1007/s10790-014-9427-z.

Sinnott-Armstrong, Walter. 1991. 'Moral Experience and Justification'. *Southern Journal of Philosophy* 29 (S1): 89–96. https://doi.org/10.1111/j.2041-6962.1991.tb00614.x.

———. 2011. 'Emotion and Reliability in Moral Psychology'. *Emotion Review* 3 (3): 288–89. https://doi.org/10.1177/1754073911402382.

Smith, Andrew, Helen Whitney, Marie Thomas, Kate Perry, and Pip Brockman. 1997. 'Effects of Caffeine and Noise on Mood, Performance and Cardiovascular Functioning'. *Human Psychopharmacology: Clinical and Experimental* 12 (1): 27–33.

Sosa, Ernest. 1991. *Knowledge in Perspective*. Cambridge: Cambridge University Press.

———. 2017. *Epistemology*. Princeton: Princeton University Press.

Szigeti, András. 2013. 'No Need to Get Emotional? Emotions and Heuristics'. *Ethical Theory and Moral Practice* 16 (4): 845–62. https://doi.org/10.1007/s10677-012-9386-z.

Tappolet, Christine. 2016. *Emotions, Values, and Agency*. Oxford: Oxford University Press.

Thagard, Paul. 2000. *Coherence in Thought and Action*. Cambridge, MA: MIT Press.

———. 2006. *Hot Thought: Mechanisms and Applications of Emotional Cognition*. Cambridge, MA: MIT Press.

Tolhurst, William. 1990. 'On the Epistemic Value of Moral Experience'. *Southern Journal of Philosophy* 29 (Supplement): 67–87.

Wang, Yangming. (1572) 2009. 'A Record for Practice'. In *Readings from the Lu-Wang School of Neo-Confucianism*, translated by Philip Ivanhoe, 131–59. Indianapolis: Hackett.

Wiggins, David. 1977. 'Truth, Invention, and the Meaning of Life'. *Proceedings of the British Academy* 62: 331–78.

Woodward, James. 2016. 'Emotion versus Cognition in Moral Decision-Making: A Dubious Dichotomy'. In *Moral Brains: The Neuroscience of Morality*, edited by S. Matthew Liao, 87–116. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199357666.003.0004.