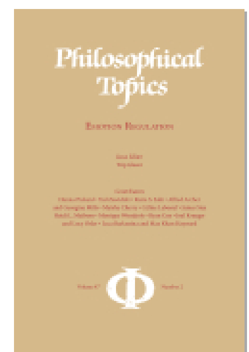
Stop Telling Me What to Feel! A Clinical Theory of Emotions and What's Wrong with the Moralization of Feelings



Hanna Pickard

Philosophical Topics, Volume 47, Number 2, Fall 2019, pp. 1-25 (Article)

Published by University of Arkansas Press

For additional information about this article
https://muse.jhu.edu/article/774353

[Access provided at 4 Dec 2020 14:16 GMT from University of Dayton Libraries]

PHILOSOPHICAL TOPICS
VOL. 47, NO. 2, FALL 2019

Stop Telling Me What to Feel! A Clinical Theory of Emotions and What's Wrong with the Moralization of Feelings

Hanna Pickard
Johns Hopkins University

ABSTRACT. "Don't be jealous of your sister." "Don't be angry with your father." "You should be more forgiving." "You ought to feel terrible for what you've done." "You ought to feel ashamed of yourself!" It is common practice within our society to morally reprimand people for their emotions, thereby expressing a kind of *moralism*: the idea that there are morally right and morally wrong ways to feel. Drawing on an alternative way of engaging with emotions derived from my experience working clinically with people with personality disorders, I argue against the value of this common practice and the moralization of emotions that underpins it. Stop telling people what to feel!

"Hate evil." "Love good." "Cease from anger and forsake wrath." "I say to you, whoever is angry with his brother will be subject to judgment!" The language of these biblical commands and warnings may sound alien to secular ears, but the legislation of how to feel is all too familiar. "Don't be jealous of your sister." "Don't be angry with your father." "You should be more forgiving." "You ought to feel terrible for what you've done." "You ought to feel ashamed of yourself!" It is common practice within our society to morally reprimand people for their emotions.

My aim in this paper is twofold: to question the value of this practice, and to argue against a key assumption underpinning it, namely, that there are morally right and morally wrong ways to feel. I shall call this claim "moralism about the emotions" and the opposing picture to emerge from the various considerations I put forward "anti-moralism about the emotions."

The paper's twofold aims are related, but distinct. A practice may be problematic not because of its underpinnings, but because of its consequences. Consider as an example the practice of truth-telling. In general, this is a valuable practice. But not all truths need to be spoken: there may be occasions when telling the truth hurts another person too much to be worth it. When this is so, the practice should be avoided, but not because what would be said is false. Similarly, our practice of morally reprimanding people for their emotions could be problematic because of its consequences, even if it is true that people often fail to feel as they morally should. If moralism about the emotions is false, this provides strong prima facie grounds for changing our practice of morally reprimanding people's emotions; but it is not the case that, if the practice is problematic and of limited value because of its consequences, then moralism about the emotions is false.

The paper has three parts. In the first, I clarify both the nature of our practice of moral reprimand, and the nature of moralism about the emotions. In the second, I question the value of our practice, by describing an alternative way of engaging with emotions, derived from my experience working clinically with people with personality disorders; and I offer a simple, short argument in favor of anti-moralism about the emotions. In the third, I offer a more complicated, lengthy assessment of what arguments there are in favor of moralism. My conclusion: Stop telling people what to feel! For, even if moralism about the emotions is not conclusively false, it lacks sufficient evidence in favor of its truth to make the consequences of our practice of moral reprimand worth it.

PART I: CLARIFYING THE QUARRY

"Don't be jealous of your sister." "Don't be angry with your father." "You should be more forgiving." "You ought to feel terrible for what you've done." "You ought to feel ashamed of yourself!" Although these sentences may be used to morally reprimand another's emotion or lack thereof, this is not the only end they can serve. Among other things, they may instead be used to (i) morally reprimand another's *action*, not their emotion; (ii) *rationally* correct another's emotion, to more accurately fit the world; and (iii) warn against the *expression* of emotion. To help fix on the practice of moral reprimand, I consider these alternative ends in order.

Consider the sentence: "You ought to feel terrible for what you've done." Although it purports to tell a person how to feel, it may in fact mean little more than: "*What you've done is wrong*." The speaker's aim may simply be to morally condemn *the action*—not tell the hearer how they should feel about the wrong

they have done. Such uses are perfectly legitimate, but they are not my focus. My focus is on occasions when the speaker's aim in uttering such a sentence *is* precisely to tell the hearer how they should feel about what they've done. "You know what you've done is wrong, and yet you carry on like it's water off a duck's back. You should be feeling terrible!" This utterance is a demand for an emotion like *guilt* in response to the recognition of one's own wrongdoing.

Consider next the sentence: "Don't be jealous of your sister." Although there is lively debate within the philosophy of emotions about how to account for their intentional content, there is near universal agreement that many (although perhaps not all) emotions *have* intentional content: emotions are affective states that represent objects as having particular evaluative properties, which these objects may or may not in fact have.[1] When an emotion represents its object as having a particular evaluative property which in fact it does not have, it misrepresents the object; and we may try to convince the person that this is so. Suppose, very roughly, that jealousy involves the representation of another person as possessing something valuable which is or ought to be yours. "Don't be jealous of your sister," we might therefore say, *because it's not actually the case that she's your mother's favorite*—your mother loves you too, just as much, so there is *no reason* to be jealous. This sibling's jealousy misrepresents their sister as possessing something valuable—more than her share of their mother's love—which in fact she does not have. To borrow a term introduced by Justin D'Arms and Daniel Jacobsen (2000) to capture this idea, in such cases, the sibling's jealousy does not *fit* the world. But (at least without further and significant argument) an emotion's lack of representational *fit* is not a *moral* fault.[2]

Utterances that aim at rationally correcting another's emotion may be voiced with a sharp edge—perhaps especially when the error appears to manifest what we typically think of as a character flaw or vice. For example, perhaps this jealous sibling is prone to *selfishness* and so predisposed to think their sister is getting more than her fair share of love. Selfishness is not, of course, the only explanation of a disposition to feel jealous for no good reason: a person who is insecure and struggles to maintain a sense of their own self-worth may be similarly predisposed. I discuss the relation between moralism about the emotions and virtue ethics toward the end of part 3. But, putting the connection between emotions and vices to one side for now, utterances that aim at rational correction can frequently be distinguished from moral reprimand by manner as much as by aim. Rather than criticize, we may try to guide, persuade, enjoin, suggest reasons for and against, and indeed even empathize with: "I am so sorry you feel jealous, but,

1. See e.g. Deonna and Teroni (2012). For a dissenting voice, see Shargel (2015); despite arguing that emotions in fact lack particular intentional objects, Shargel does not disagree that we often *take* them to.

2. D'Arms and Jacobsen (2000) argue the opposite: they assume that we *can* be morally faulted for our emotions, but argue that the moral wrongness of an emotion does not entail a lack of representational fit.

you know, you really shouldn't. Your mother loves you so much!" Once again, the use of imperatives or deontic modals to rationally correct another's emotion is perfectly legitimate, but not my focus. My focus is on occasions when the aim *is* moral reprimand. To return to the example, perhaps this mother in fact loves the sister better: the world is *exactly* as the jealous sibling represents it to be. Yet this sibling is nonetheless scolded for their jealousy because the speaker takes the view that *it is wrong to be jealous*, e.g. not the way a good child ought to feel—entirely apart from whether or not the feeling *fits*.

Consider finally the sentence: "Don't be angry with your father." In some contexts, this sentence might be used not to morally reprimand the child for their anger, but to warn against its expression: for example, perhaps this father has a temper and is prone to violence when his children talk back. In such cases, the sentence is used as a kind of warning: the speaker may think that the child has every reason to be angry, *but it is not advisable to show it*. "Don't be angry with your father" really means: "Don't *let him know* you're angry!" Of course, a salient characteristic of emotions is that, when you have them, they can be hard to hide. So, perhaps, if it would be better overall for your father not to know that you're angry, then the simplest way to achieve this would be not to be angry in the first place: you have an instrumental reason not to be angry (see below). It is usually even more difficult to forbear from *having* an emotion than from expressing it, simply because it would be advisable (see part 2). Nonetheless, it is common for sentences instructing another *not to feel* a certain way to be used as a warning against *expressing* the feeling. But, if this is the aim, then the utterance is not a form of moral reprimand: it is not an expression of *moralism about the emotions*.

Moralism about the emotions is the idea that there are morally right and morally wrong ways to feel. For a clear example, consider the following quotation from Robert Adams (1985):

> How then can we be blamed for anger? My answer is that the thesis that we are ethically accountable only for our voluntary actions and omissions must be rejected. There are involuntary sins, and unjust anger is only one of them. Among the others are jealousy, hatred, and other sorts of malice; contempt for people, and the lack of healthy concern for their welfare . . . Obviously there are types of voluntary action that are characteristic of anger, for example, and that are morally objectionable. But that is not all that is wrong about unjust anger. If I am furiously angry at you without proportionate cause, but refrain from hitting you, insulting you, yelling at you, or speaking ill of you, no doubt my self-restraint is a credit to me; but the anger remains a moral fault. (pp. 3–4)

Adams qualifies his claim that anger is a moral fault with the caveat that this is so when it is "unjust" or "without proportionate cause." Unjust anger suggests a misrepresentation of its object; while anger without proportionate cause suggests the object has done something worthy of anger but that nonetheless the anger is outsized relative to the wrong done. I noted above that (without further argument) a lack of representational fit is not a moral fault; I will say more about this in parts 2

and 3. The point for now is that Adams does not similarly qualify jealousy, hatred, malice, contempt, or lack of concern for another's welfare. In this, he is arguably in step with our common practice, which may morally reprimand when there is no representational error or (we might add) lack of proportionality, as in the above example of the child who is scolded for jealousy, despite its worldly fit.

Moralism about the emotions is a specific mode of assessment: it is *intrinsic* as opposed to instrumental; and it is *deontic* as opposed to value-based. Consider first the contrast between deontic and value-based assessment. As Michael Smith notes, the contrast is a generalization of the distinction between *the right* and *the good* (Smith 2007). Deontic concepts include the obligatory, the forbidden, and the permissible; value-based concepts include the desirable and the undesirable, the good and the bad, the better and the worse. Although the correct account of the nature of and connections between these concepts is controversial, it is nonetheless widely agreed that the distinction depends on the possibility of attributions of moral responsibility. When a person fails to do something they are morally obliged to do, or does something that morality forbids them from doing, they have done something morally wrong; and the question of whether or not they are responsible for what they have done immediately arises. Of course, they may not be responsible: perhaps they have an excuse. The point is that the mere fact of deontic assessment raises the possibility. Value-based assessment does not. The mere fact that something is judged to be undesirable or bad or worse than something else does not immediately raise the question of anyone's moral responsibility (indeed, these concepts apply in nonmoral domains, such as e.g. the aesthetic and the epistemic). For example, pain is undesirable and bad; pleasure is better. But the fact that a person is in pain and that pain is negatively evaluated does not ipso facto make this person a candidate to be held morally responsible for their pain; mutatis mutandis with respect to moral praise and pleasure.

In the above quotation, Adams endorses the idea that emotions are objects of *deontic* assessment, as evidenced by his talk of "blame" and "fault" and "accountability." Indeed, it is deontic assessment that is concretized in our practice of moral reprimand of emotions, which not only liberally employs the language of *oughts* and *shoulds*, but can itself be understood *as* a practice of holding people morally responsible—precisely by calling them out for their emotions and thereby demanding that they answer to the charge of *wrong feeling*. For it is *wrongness* that immediately raises the specter of moral responsibility and with it the appropriateness of reproach; while *badness*, at least in the first instance, calls simply for regret.

Consider next the contrast between intrinsic and instrumental assessment. Moralism is a mode of intrinsic assessment: it assesses the emotion irrespective of its consequences. But we nonetheless do assess emotions instrumentally in relation to many sorts of consequence. For example, emotions can incline us toward actions. Recall the father whose temper leads to violence: his anger inclines him to do something morally objectionable (as Adams puts it). Indeed, all the emotions that Adams lists have this potential: jealousy, hatred, malice, and contempt may all, on

occasion, incline us to act badly (for discussion see part 2). On the other hand, some emotions may incline us to do good, e.g. love and compassion. Insofar as emotions aid and abet actions and outcomes that are morally right or wrong, or good or bad, we can assess them as *instrumentally* welcome or otherwise, relative to these actions and outcomes. But this instrumental assessment does not confer any *independent* moral status on the emotions themselves: their value or disvalue depends entirely on the consequences to which they contribute. If this is not immediately obvious, consider that e.g. *pain* can sometimes (and understandably) make us act badly. When this is so, it is unwelcome not only because it is a form of suffering, but in addition because of its influence on how the sufferer behaves toward others. But this does not mean there is anything intrinsically morally wrong with being in pain. Indeed, in such cases, pain may well function as an *excuse* for how a person behaves—not a further possible cause for reproach.

I have used the contrast between deontic and value-based modes of assessment alongside the contrast between intrinsic and instrumental modes of assessment to clarify the nature of moralism about the emotions: moralism is the deontic claim that there are intrinsically morally right and morally wrong ways to feel. But, having drawn these contrasts, one might wonder whether emotions can also be evaluated as *intrinsically good or bad*. For example, Amia Srinivasan (2018) defends the *intrinsic value* of *anger* at injustice:

> Imagine a person who does everything, as it were, by the ethical book—forming all the correct moral beliefs and acting in accordance with all her moral duties—but who is left entirely cold by injustice, feeling nothing in response to those moral wrongs of which she is perfectly aware. I don't want to say that such a person has done anything wrong. But I do think it is natural to say that there is something missing in her; indeed, that it would be better, ceteris paribus, if she were capable of feeling anger towards the injustice she knows to exist. (p. 132)

Srinivasan sees anger as a distinctive way of *appreciating* injustice and, for that reason, of intrinsic value—akin to how aesthetic appreciation is a distinctive way of knowing that an object is beautiful. She suggests that to deny this intrinsic value is to embrace a broad and, for all that has been said, baseless, skepticism. We should no more doubt that emotions have intrinsic value than we should doubt this of aesthetic or indeed epistemic goods, such as e.g. truth, justification, and knowledge.

Srinivasan is right to emphasize that, for many people, anger is an important way of appreciating injustice—and one which our society all too often fails to validate. Equally, it may well be true that affective and aesthetic modes of appreciation, alongside epistemic goods, all have intrinsic value. If so, then, on the assumption that the more value there is in the world, the better, then the more modes of appreciation and epistemic goods we all of us have, the better. In this sense, it would indeed be *better* if this hypothetical person felt anger. But it is a further claim that, because she does not, something is *missing* in her. This suggests that there is a norm or standard—whether that is based in human biology or in ethics (see

part 3)—from which she deviates in virtue of failing to feel anger. In other words, although Srinivasan is clear that the woman does nothing wrong, she nonetheless implies that something is wrong *with her*: there is something about her as a person that is *flawed*. This further claim, which goes beyond the more trivial point that the more value there is, the better, evokes the kind of moral assessment of persons and their characters constitutive of virtue ethics.[3] As noted above, I shall discuss the relationship between moralism about the emotions and virtue ethics toward the end of part 3. For now, the point is simply to distinguish moralism both from the idea that emotions can be assessed instrumentally in relation to the actions and outcomes they aid and abet, and from the idea that they have intrinsic value—on the model of aesthetic appreciation and epistemic goods. Both these ideas are orthogonal to moralism and compatible with anti-moralism: the denial of the deontic claim that that there are intrinsically morally right and morally wrong ways to feel.

Lastly, a point of terminology and an analogy. The term "moralism" is sometimes colloquially used to mean that too much is made of matters which are, nonetheless, genuinely of moral concern. For example, a person may be judged "moralistic" if they treat minor moral infractions too harshly, or they give moral reasons too much weight when they conflict with e.g. prudential reasons. This is not how I mean to use the term in relation to the emotions. By "moralism" I mean what we might think of as a prior claim, namely, that a domain is the appropriate object of deontic moral assessment *in the first place*, and by "anti-moralism" I mean the rejection of such assessment as appropriate to the domain in question.

An analogy can help to explain. Modern liberal society has made great advances battling moralism, thus understood, in a variety of domains, such as e.g. sexual orientation. For example, homosexuality used to be widely viewed as morally wrong, as evidenced by the criminalization of same-sex sexual acts, the illegality of same-sex marriage, as well of course as the pervasiveness of discriminatory and stigmatizing social practices. The recent reversal of these laws and the rejection of these practices are moral advances. But they are not moral advances because, say, criminalization "made too much" of the relatively minor but nonetheless genuinely moral infraction that homosexuality constitutes. They are moral advances insofar as they express the correct view *that homosexuality is not wrong.* This is not, however, because it is *right* to be homosexual, but because sexual orientation is not an appropriate object of such moral assessment. It is *neither right nor wrong* to be lesbian or gay or straight or bi or queer or asexual—it just is how it is—and it is a mistake to think otherwise

3. Note that this character-style assessment is in no way necessary to Srinivasan's core aim in her paper, which is to articulate the additional *affective* injustice that victims of injustice can suffer, when they find themselves feeling anger which, to put it in the terms used above, is representationally fitting but instrumentally counterproductive given the oppressive circumstances in which they live. I discuss her argument not to dissent from her ultimate conclusion, but because it is such a striking example of the allure of moralizing the emotions.

and to legislate, both literally and metaphorically, over this domain. I aim to explore the possibility that we should be anti-moralistic about the emotions in the same way we have rightly become anti-moralistic about sexual orientation. Just as we have stopped telling people whom to love, we should stop telling people how to feel.

PART II: TREATING PEOPLE WITH PERSONALITY DISORDERS

Between 2007 and 2017, I worked part-time in a National Health Service (NHS) specialist service for people with personality disorders and complex needs. A personality is something each of us has: it is constituted by the set of traits that make us the kind of person we are—the ways we are inclined to think, feel, and act, in response to particular circumstances, as well as more generally. These traits can serve us better or worse; rare is the person who would not change any part of their personality, if only they could. Personality *disorders* occur when these personality-constituting traits cause a person to experience clinically significant distress and impairment in social, occupational, and other important contexts—the ways they are inclined to think, feel, and act cause them serious harm, directly or via the effects they have on relationships, work, and life more generally. The DSM-5 (American Psychiatric Association 2013) distinguishes ten kinds of personality disorder, grouped into three clusters known colloquially as “mad,” “bad,” and “sad.” Diagnosis with any kind of personality disorder requires meeting a complicated, polythetic set of criteria, spanning not only behavior but beliefs, desires, and, most importantly for my purposes here, emotions. Indeed, many of the emotions characteristic of personality disorders (especially the “bad” cluster) are precisely those which commonly *invite* moral reprimand. For example:

I. Extreme, inappropriate anger, which may lead to self-harm and/or aggression and/or violence toward others

II. Contempt and lack of empathy toward others, which may lead to exploitation and a willingness to “step on another’s toes”

III. Intense jealousy of romantic partners, which may lead to harassment and/or coercion

IV. Intense envy of others, which may lead to mistreatment and/or abuse

V. Lack of guilt or remorse for serious criminal and/or immoral conduct

The primary goal of all clinical work is to help people improve and recover from their disorder; in the case of personality disorders, this involves therapy aimed at changing the relevant set of traits. This is valuable in itself because of the good to the patient. But, given that personality disorders often involve harm to others (as illustrated by (i) to (v)), the moral stakes in clinical success are especially high.

How do you treat people with personality disorders? In other words, how do you effectively address these emotions and associated actions in the clinic, thereby helping people with personality disorders improve and recover while at the same time reducing the likelihood of harm to others? The answer is that what you most certainly *don't* do is morally reprimand people for how they feel. This is, in truth, a platitude. One hardly needs to look to the clinic to know that e.g. being told you shouldn't be angry when you already are is like throwing fuel on a fire. We all know that our practice of morally reprimanding people for their emotions rarely serves to *change* these emotions—indeed, it may have the opposite effect, serving to further augment and entrench the feeling. What attention to clinical work with patients with personality disorders shows is how much is at stake—and what you might do instead if you do want to construct a practice that harbors the possibility of change.

With respect to actions that cause harm, effective clinical treatment employs a stance I have elsewhere called *responsibility without blame* (Pickard 2011, 2013a; www.responsibilitywithoutblame.org). Within the specialist service where I worked, people with personality disorders were treated as responsible for their actions and held to account; but, crucially, this was done in a way that maintained care and compassion, and avoided both judgments condemning the person's character and the hostile, negative emotions—such as e.g. anger, indignation, resentment, contempt, disgust, and scorn—that are commonly part and parcel of how we blame. Divorcing practices of responsibility and accountability from *affective blame* enables these practices to promote behavioral change. Rather than ascriptions of responsibility functioning as a necessary condition for the backward-looking aim of censure and punishment for past wrongdoing, in the clinic they become a tool for the forward-looking aim of improvement and recovery—supporting and enabling the person to do things differently in the future.

But people can only change that over which they have choice and (at least a degree of) control, e.g. voluntary actions. This is a constraint on the *practical effectiveness* (not just the moral permissibility) of responsibility and accountability practices promoting forward-looking goals: they must target behavioral patterns over which the person has choice and control. But emotions, unlike actions, are not directly subject to choice and control. So, by contrast with actions, patients are not typically held responsible or to account for *having* emotions—even when these emotions lead to actions which cause harm and for which they *are* held responsible and to account.

I say that patients are not "typically" held responsible or to account for having emotions because, of course, we are sometimes able to exercise *indirect* choice and control over emotions, e.g. by avoiding or seeking out situations we know may trigger them, or by critically reflecting on where they come from in us, or by undertaking therapy. For this reason, if we know that some of our emotions typically lead us to hurt other people, we may have a moral obligation to do all we can to stop ourselves from experiencing them—and we may be *derivatively* responsible

for having an emotion if we fail to fulfill this duty. But it is important that this responsibility is derivative. Unlike actions, emotions are fundamentally passive: things that happen to us, not things we do. This passivity is evident not only in the fact that we cannot choose our emotions at will, but also in the fact that we cannot rely on our emotions being responsive to reason. To illustrate this point, consider the difference between doxastic and emotional deliberation. If we don't already have a belief, we often deliberate about what to believe. But if we don't already have an emotion, we don't usually deliberate about *what to feel*, for the simple reason that such deliberation is, at least in most cases, *idle*. It will not reliably produce an emotion—unlike deliberation about what to believe. Now, if we *already* feel a certain way, we can of course deliberate about whether our feeling in fact fits the world: whether there is good reason to feel it. If not, then, if we are lucky, the feeling may dissipate. But, as we all know from our own experience, and the philosophy of emotions has started to recognize more fully (Brady 2008; D'Arms and Jacobson 2003; Pickard 2013b; Roberts 2003), emotions are often highly irrational and recalcitrant: they regularly persist in the face of knowledge that they do not accurately represent the world (again, contrast belief). Irrational and recalcitrant emotions are normal: they are not aberrations.

Consider the following four features of emotions, all part of our folk theory of mind and routinely used to explain and predict feelings. First, emotions depend on moods as much as they depend on the world they represent. Grumpy, your child's disobedience makes you angry. Cheerful, it fills you with joy in their willfulness. Second, emotions capture attention and skew the salience of new stimuli, a phenomenon known as "emotional congruence." The child who is scared of the dark is not helped when their parents turn on the nightlight, for their fear simply transfers to the monster-like pile of clothes on the floor. In other words, emotions *find* objects to fit them; they do not merely fit objects already found. Third, and relatedly, emotions are all too often *idées fixes*, as Peter Goldie (2004) has emphasized. Rather than being guided by reason, they find reasons to justify why they represent the world right. Consider, for instance, Leontes in Shakespeare's *A Winter's Tale*, who, jealous of his wife, Hermione, rejects every reason why his jealousy might not be fitting while amassing reason after reason why it is; for a more familiar example (at least in my case), think of how easy it is to find oneself shifting ground to justify one's anger when arguing with one's partner. Fourth, many of our deepest, most important emotions are utterly irrational and recalcitrant, and *known* by us as such. They persist despite the fact that we believe them to lack any and all rational grounds. Obvious examples include phobias, aversions, and emotions which carry the residue of past experience and, especially, trauma. Richard Wollheim suffered a lifelong debilitating aversion to wet newspaper, which in his autobiography he calls "the most persistent thread in my life, stronger, more unchanging, than any taste or interest, more demanding than any intellectual challenge" (2004, p. 248). Wollheim knows perfectly well that there is nothing objectively dangerous or disgusting about wet newspaper. But he traces

his aversion to a childhood experience of his brother flicking saliva-sodden pellets of newspaper at him out of malice—defiling not only Wollheim himself, but also, as it happens, a picture of the recently deceased Queen, that Wollhein was enjoying looking at while sitting comfortably in his nanny's lap. Freud would of course have a heyday with this confession; but, once the aversion is understood as a response to this particular childhood experience as opposed to the nature of wet newspaper in and of itself, it becomes perfectly intelligible—indeed, arguably, it is not just intelligible but *proportionate* once it is understood to contain the residue of this quasi-traumatic past formative experience. This is because our folk theory of mind takes emotions to be formed and sustained as much through narrative arcs and symbolic associations as through rational engagement with the world; and, when they are thus formed and sustained, we understand that they reveal more about the self and its history than about their actual worldly objects.

Summing up: emotions are neither directly responsive to the will nor reliably responsive to reason. Clinical practice therefore draws a sharp distinction between emotions and any associated actions—which, unlike emotions, are both responsive to will and to reason. In the clinic, how you feel is one thing. What you *do* with that feeling—whether, in particular, you choose to act on it, and, in general, how you manage it—is another. In other words, in the clinic it is taken for granted that e.g. *just because you're angry and have an impulse to act aggressively or violently doesn't mean you have to act on it*—similarly for the other emotions in (i) to (v) above. There is a sharp and clean divide between actions and emotions, and, as a matter of *effective* clinical practice (putting aside the issue of moral permissibility), people are responsible and accountable only for that over which they have choice and control: their actions, not their emotions.

However, the emotions in question also need to be addressed. They cause people terrible distress. As illustration, consider this first-person report from a woman with a diagnosis of borderline personality disorder:

> "Being a borderline feels like eternal hell. Nothing less. Pain, anger, confusion, hurt, never knowing how I'm gonna feel from one minute to the next. Hurting because I hurt those who I love. Feeling misunderstood. Analyzing everything. Nothing gives me pleasure. Once in a great while I will be 'too happy' and then anxious because of that. Then I self-medicate with alcohol. Then I physically hurt myself. Then I feel guilty because of that. Shame. Wanting to die but not being able to kill myself because I'd feel too much guilt for those I'd hurt, and then feeling angry about that so I cut myself or O.D. to make all the feelings go away. Stress!" (Anonymous, quoted by Dakota, August 7, 2007, http://recoveryourlife.com/forum.)

In line with the final thought in this quotation, harmful or wrongful actions are often understood in the clinic as ways of getting relief from distressing emotions. When this is so, the key is to help people find alternative ways of managing these emotions, both to offer relief and to avert the action. Crucially, this requires an anti-moralistic therapeutic space. Moral reprimand for emotions is no help.

But, if you take seriously the thought that there are no intrinsically morally right or wrong ways to feel—emotions just are what they are—this creates a safe and nonjudgmental context to experience, explore, and put into words emotions that would otherwise be met with moral reprimand.

There are at least three ways this helps people. First and most straightforwardly, putting feelings into words can bring immediate relief from emotional distress. Talking about emotions is a healthy, positive way of coping with them. Second, the process of experiencing, exploring, and talking about emotions may change them over time by offering a form of re-processing and re-consolidation (Lane et al. 2015). But, even if not, the improved reflection engendered by this process enables people to stop and think about their feelings, rather than acting without thinking to relieve them. This is especially important in treatment of personality disorders because the emotions in question quite often, like Richard Wollheim's aversion, contain the residue of past experience and trauma (Pickard 2015). They are experienced as "extreme" and "intense" by the person (and others);[4] but this is when their fit and proportionality is assessed only in relation to their present object in the here-and-now. Narrative exploration of the impact of a person's past on their present can serve to render these emotions more intelligible and proportionate in the eyes of the person as well others; this improved understanding is a powerful tool, in turn, for emotional regulation and management. Third, the emotional *validation* or *toleration* that an anti-moralistic therapeutic space affords gives people an experience of acceptance, understanding, and empathy. It thereby reduces isolation, increases hope, and is a panacea to emotional distress, by offering recognition and support. Part of why this is important is that most people with personality disorders are just like the rest of us: they judge and morally reprimand *themselves* for their emotions. This does not help them change any more than being reprimanded by others does. Anti-moralism offers relief from external forms of criticism and punishment, which can model a different way for patients to treat themselves—often a foundation to enabling change and recovery.

Our common practice of morally reprimanding people for their emotions is typically futile, insofar as we have forward-looking ends. Indeed, it may make things worse, by entrenching the emotion, augmenting distress, ostracizing the person, and therefore increasing the likelihood of these emotions leading to harm. In contrast, anti-moralism as embodied by clinical practice is poised to do moral good, for the patient and for others whom the patient might otherwise harm. Therefore, with respect to the emotions characteristic of personality disorders, moral reprimand has significant moral costs, while the clinical alternative has significant moral benefits.

My discussion has focused on the specialized context of a personality disorder clinic. But there is every reason to think its conclusions generalize to more everyday contexts. First and foremost, there is no sharp diagnostic line to be drawn

4. Compare Adams's "unjust" and "disproportionate" anger.

between people with personality disorders and the rest of us: we are all on the same continuum. Personality disorders are diagnosed through a structured clinical interview requiring even more than the usual degree of clinical judgment; the criteria for diagnosis are polythetic; and, in ways most importantly, the threshold for diagnosis with the disorder is both arbitrary and contingent on circumstances: in the right environmental niche, personalities that would lead to distress and impairment in other contexts may instead be rewarded with success. Hence, what works to help people with personality disorders manage their emotions is likely to help the rest of us manage ours. Second, rates of serious moral offense, let alone more minor infractions, are pervasive in our society—people with personality disorders are far from the only perpetrators. On the assumption that these offenses and infractions may be driven by emotions not only when the perpetrators have a diagnosable personality disorder but also when they do not, we have reason to deploy anti-moralism and associated clinical techniques quite generally to address these emotions. Third, once the clinical practice of anti-moralism and associated techniques is brought to light, it becomes evident that it exists in various pockets outside of the clinic already; perhaps most strikingly, in current parenting and child education practices,[5] but also arguably in many examples of friendship. All in all, although people with personality disorders have emotions that can strike

5. For a lovely example (notwithstanding its gendered conclusion) consider the children's television presenter Fred Rogers's song "What Do You Do with the Mad That You Feel?" which frankly could be a page out of a clinical textbook: https://www.youtube.com/watch?v=COJIi26hrZk; indeed, Mr. Rogers worked closely with Margaret McFarland, a child psychologist, to develop his show *Mr. Rogers's Neighborhood.*

> What do you do with the mad that you feel?
> When you feel so mad you could bite?
> When the whole wide world seems oh so wrong?
> And nothing you do seems very right?
> What do you do? Do you punch a bag?
> Do you pound some clay or some dough?
> Do you round up friends for a game of tag,
> And see how fast you go?
> It's great to be able to stop,
> When you plan a thing that's wrong.
> And be able to do something else instead,
> And *think* this song.
> I can stop when I want to, can stop when I wish,
> Can stop stop stop anytime
> And what a good feeling to feel like this,
> And know that the feeling is really mine.
> Know that there's something deep inside
> That helps us become what we can
> For a girl can be someday a woman
> And a boy can be someday a man.

them and us as intense and extreme, they are hardly unique: this is something we can all experience from time to time, adults and children alike. Anti-moralism and associated clinical techniques are therefore likely to do as much good outside of the clinic as within it.

Although my aim in this section has been to argue against our practice of moral reprimand of emotions, my discussion contains the resources for a short, simple argument in favor of anti-moralism:

> I. Deontic claims about what is morally right and morally wrong entail the possibility of attributions of moral responsibility.
>
> II. For the possibility of attributions of moral responsibility (and the accountability practices thereby prima facie licensed) to be legitimate, the person in question must have choice and control over that for which they are judged morally responsible.[6]
>
> III. It is not the case that emotions are subject to choice and control: emotions are neither directly responsive to the will nor reliably responsive to reason.
>
> IV. Therefore, deontic claims about rightness and wrongness do not apply to the emotions.

Of course, an advocate of moralism about the emotions may deny one or more of the premises of this argument; e.g. in the quotation above, Robert Adams denies (ii).[7] I shall not try to defend each premise against all possible objections. What I hope to have established in this section is more modest: to have delineated the consequences of our practice of moral reprimand of emotions and the value of the clinical anti-moralistic alternative; and to have offered an intuitive argument for the truth of anti-moralism, as evidence of its cogency.

PART III: MORALISM ABOUT THE EMOTIONS

The moral costs I have charged to our common practice of moral reprimand of emotions are consequentialist: it does not serve to *change* these emotions and it may indeed have the opposite effect, augmenting or entrenching the feeling and increasing the likelihood of ensuing harm. However, if moralism about the emo-

---

6. Note that the justification for this claim about the legitimacy of an attribution of moral responsibility and any corresponding accountability practices could be either *backward-looking* or *forward-looking*: a question of moral permissibility or a question of practical effectiveness.

7. "How then can we be blamed for anger? My answer is that the thesis that we are ethically accountable only for our voluntary actions and omissions must be rejected" (Adams 1985, p. 3). Note that a commitment to (ii) does not entail that we can *never* be responsible, in any sense, for having emotions. As suggested above, given that we are sometimes able to exercise *indirect* choice and control over emotions, e.g. by avoiding or seeking situations we know may trigger them, etc., we can be *derivatively* responsible for having emotions when they are the foreseeable outcomes of our actions and omissions.

tions is correct—that is, if there really are intrinsically morally right and morally wrong ways to feel—then one might hold there is moral value in moral reprimand, no matter the consequences. The value is *to censure* that which is morally wrong, thereby upholding the moral code that governs emotions.

Is moralism about the emotions correct? There are two parts to moralism. There is the *positive* claim that there are certain emotions it is morally right to feel: if you ought to have a particular emotion, but you fail to have it (either because you have a different emotion or because you feel nothing), then you do not feel as you ought. There is also the *negative* claim that there are certain emotions it is morally wrong to feel: if you ought *not* to have a certain emotion, but you do have it, then you do not feel as you ought. I shall consider these claims in order.

How might the moralist argue for the positive claim, that there are certain emotions that, morally, one ought to have? Within the literature on neo-sentimentalism and value theory, people often speak of certain facts "calling for" or sometimes "demanding" or "requiring" a specific emotional response. Inspired by Amia Srinivasan's discussion of anger as a way of appreciating injustice (see part 1), we might suggest that injustice *calls for* anger; similarly, perhaps one's own wrongdoing calls for guilt or shame; or another's suffering demands compassion.[8] Relative to the circumstances, these emotions constitute the *morally right* ways to feel. So, if someone does not respond to these facts in these ways, they do not feel as they morally should.

Why would this be true? I take the intuition driving this idea to be that, when confronted by a fact that is morally significant, one ought to *recognize* it as such. This seems prima facie plausible. For instance, if one did not recognize it, then one couldn't reliably *act rightly* in response. But it is a further and as yet undefended claim that it needs to be recognized *by feeling* a certain way. As Mark Johnston (2001) puts it, "a cool appraisal" may be all that's called for.

Why would the recognitional requirement need to be fulfilled by feeling in particular? Christine Tappolet (2016) has argued that a natural way to interpret neo-sentimentalism is as a claim about the possession conditions for evaluative concepts.[9] Just as having the experience of red may be a necessary condition to possess the phenomenal concept of red, so too having the experiences of anger and

---

8. Michael Moore (2010) takes this view of guilt, shame, and compassion. For example, he suggests that, if one is the perpetrator of extreme moral wrongdoing, "one ought to feel so guilty one wants to die" (p. 145). Moore holds that we evaluate not only people's characters but also their emotions as intrinsically virtuous or vicious: he remarks that "the emotions in such a case are the object of moral evaluation" (p. 134).

9. See too David Wiggins (1987). Note that this neo-sentimentalist claim must be distinguished from the claim that there is an *epistemic* connection between emotions and moral judgment—that, insofar as emotions can reveal morally significant facts to us, they can ground moral judgment. Johnston (2001), Moore (2010), and Srinivasan (2018) all emphasize the importance of this epistemic feature of emotions; for a more pessimistic discussion of the epistemic value of emotions, see Goldie (2004).

guilt, say, may be necessary conditions to possess the concept of wrongdoing.[10] If this interpretation of neo-sentimentalism is accepted, then we have the following argument for the moralist's positive claim:

> I. If confronted by a morally significant fact, one ought to recognize it as such.
>
> II. To recognize it, one must possess the relevant moral concepts.
>
> III. It is a necessary condition to possess certain moral concepts that one experience certain corresponding emotions.
>
> IV. Therefore, one must sometimes respond to morally significant facts with the corresponding emotions, on pain of failing to have the moral concepts required to recognize them.

Take injustice as an example. Recognition of injustice requires possessing the concept of injustice. Let the moralist concede that we may not always need to respond to injustice with anger: a "cool appraisal" will sometimes do. But if we grant the neo-sentimentalist idea about concept possession, then, if a person *never* responded with anger, say, they could not have the concept of injustice. Hence, on some occasions where recognition of injustice is called for, it is necessary to respond with anger, on pain of failing to have the concept of injustice and the ability to recognize it at all.

The problem with this argument is that it contains a hidden assumption about the scope of the recognitional demand in premise (i). A rat scurries by a mugging on the city streets. Ought it to recognize the wrongdoing confronting it, by feeling or by cool appraisal? Certainly not. Rats are social and intelligent animals, but they have not evolved to have the capacity for moral thought and action. The assumption contained in this argument is that we as a species have evolved to have certain emotional capacities in response to certain morally relevant circumstances; and the argument moralizes deviation from that species-specific *biological* norm.

The roots of the idea that humans have evolved to have a set of basic emotions is often attributed to Charles Darwin's *Expressions of the Emotions in Man and Animals* (1998). However, it is far from clear that Darwin in fact held this view. Darwin discusses over sixty emotions without arguing for a basic subset. More significantly, Sue Campbell (1997) has argued that Darwin's theory of the universality of certain species-specific *expressions* of emotion in fact undermines the plausibility of a theory of universal species-specific *emotions*: once expressions are postulated as inherited habits that may once have been "serviceable" but no longer are, emotions themselves, understood as affective and intentional psychological states, become explanatorily idle. Other theorists, most notably Paul Ekman, have more

10. To my knowledge, the only person to have explicitly argued against this view is Gil Harman (2009). Harman offers as a counterexample himself (and some unnamed friends of his) who apparently do not feel guilt but are moral people nonetheless, in possession of the full panoply of moral concepts.

clearly argued for a set of universal species-specific basic emotions, including at minimum anger, fear, sadness, enjoyment, disgust, and surprise, and possibly in addition contempt, shame, guilt, embarrassment, and awe (1992). Part of Ekman's argument for this claim derives from his now famous studies, demonstrating a reliable pancultural human capacity to *recognize* facial expressions; another part stems from studies of actual expressions, both deliberate and spontaneous (1989). I do not wish to take a stand on the question of whether or not there exists a set of species-specific basic emotions, and if so, which ones. But I do think it is important that these studies do not establish what proportion of the human population *has* some or all of these emotions—let alone which emotions we have evolved to have in which circumstances.

Consider something as simple as being confronted with a man who threatens you with violence. What do you feel? In my case, *fear*. I experience the man as *dangerous*, which indeed he is. But I know some people—and have worked clinically with some men—who feel, not fear, but *anger*. How *dare* he threaten them? And I have known some men and women alike who feel neither fear nor anger but, although I would not go so far as to say joy, *thrill*, or *excitement*. Note that an angry or excited response does not fail to see the man as dangerous. Far from it—that is part of what makes the person angry or excited. Fear and anger (if not thrill or excitement) are universal basic emotions if any are; and I suspect which (if either) you feel in such a context partially (but of course not entirely) tracks gender—for cultural as much if not more than biological reasons. The point is that even a simple case like this, which was surely common enough in our early ancestral environment, does not suggest that humans have evolved to have a *particular* basic emotion when confronted with danger. There is *diversity* in how we respond. Indeed, a similar kind of point can be made with respect to witnessing or experiencing injustice. Many of us indeed feel *anger*, but some of us may feel *sadness*, while others of us, especially perhaps those who have frequently been victims of injustice, may feel *fear*, while others again—like Srinivasan's hypothesized woman—may feel nothing at all. This variety is no mystery: we can explain it by appeal to factors such as culture, context, personality, genetics, and past experience. But, although it is no mystery, it is extremely important: by analogy with diversity in sexual orientation, diversity in emotional response should make us pause before jumping to conclusions that there are *biologically* (let alone morally) right and wrong ways to feel. It is simply not obvious that, in virtue of being human, specific emotions are "called for" by specific facts—morally significant or otherwise.

How might the moralist reply to this objection? Perhaps what is universal is that we as a species have evolved to have a *range* of basic emotions. Which emotion someone feels in specific circumstances may vary. But if someone never felt anger, or sadness, or fear, or excitement, in *any* relevant context, something would be biologically wrong with them (*missing*, as Srinivasan says). And, the retort might continue, this *biological* deviation results in *moral* deviation: granting the neo-sentimentalist

idea about possession conditions, such a person could not then *have* the moral concepts paired with whichever emotions were missing from their psychology, and so they could not recognize the relevant moral facts.[11] Therefore, there would be something *morally wrong* with them because of this deviation from species-specific biological norms.

The first problem with this retort is that the fact that we as a species have evolved to have a range of emotions, if true, does not establish that we have evolved to each have that range. Compare: we as a species have evolved to have a range of eye colors, but we have not evolved to each have all colors in that range. The second problem is that there is no valid inference from biological norm deviation to moral norm deviation without additional premises: biology and ethics are separate domains, each with their own distinctive norms. Within the philosophy of medicine, Chrisopher Boorse (1975, 1977) has long and forcefully argued that deviation from species-specific biological norms has no intrinsic disvalue. The reason is that we always can (and should) ask: what exactly is supposed to be wrong with being biologically atypical or dysfunctional? Applying Boorse's point to the case at hand: I do not deny that it would be *better* if each and every human was able to recognize all possible morally significant facts they are confronted with, on the assumption that the more people who can recognize and reliably act in the face of morally significant facts, the better. If, as the neo-sentimentalist holds, this requires each of us to have a full range of basic emotions, then it would be better if we did. But this is a broad, instrumental, and value-based claim, not a deontic one: the fact that it would be morally better if we each did does not mean that it is morally wrong if we don't.

Recognizing emotional diversity and the gap between biological and moral deviation provides a backdrop for the problem with premise (i). If ought implies can, then, if a person can't recognize certain moral facts, it is not the case that they ought to. The recognitional demand must be relativized to those who possess the recognitional capacities required to meet it; in this case, ex hypothesi according to the neo-sentimentalist, those capable of having the relevant emotions. But once relativized, the requirement does not have to be met by *having* the emotions. Hence the argument for the positive claim of moralism faces a dilemma. Either the person is not in possession of the relevant moral concept (as rats and perhaps some humans are not), in which case there is no recognitional demand. Or they are in possession of the relevant moral concept, in which case the demand for recognition need not be satisfied specifically by having an emotion.

11. Psychopathy is commonly used to support this kind of argument. It is often assumed that people with psychopathy cannot feel empathy, which impairs them from acquiring moral concepts. However, the fact that some people who lack moral concepts cannot feel empathy, if true, does not establish that all people who have moral concepts can feel empathy. For the record, it is in fact unclear that people with psychopathy lack moral concepts (Glenn 2017) and cannot feel empathy: although they may not often feel it spontaneously, there is evidence that they are able to empathize when they try (Meffert et al. 2013).

Consider now the negative claim of moralism, that there are certain emotions that, morally, one ought not to have. To be absolutely clear, this is *not* simply the claim that there are emotions that are rationally ungrounded and do not fit the world. As all our irrational and recalcitrant emotions amply demonstrate, this is uncontroversial (see part 2). Nor is it the claim that it is *instrumentally* unwelcome to have certain emotions, e.g. because they may incline us toward wrong actions or bad outcomes. Again, this is uncontroversial, and it is why people may e.g. try to avoid triggers if they have anger management problems, say, or critically reflect on why they are angry or where it comes from in them, or undertake therapy (see part 2). The negative part of moralism is that it is *intrinsically* morally wrong to feel certain ways. Why would this be?

Angela Smith (2011) has argued that we can make sense of "wronging another" (as she puts it) simply by having entirely private, undisclosed attitudes, including certain very specific emotions, such as e.g. contempt for others based on race or gender. She holds that having such an emotion constitutes a moral transgression in itself, irrespective of whether the emotion is expressed or acted on or causes any actual harm to anyone. Smith suggests that these emotions violate the norms governing the relationship of "mutual recognition" that Tim Scanlon (1998) has argued we stand in to one another: contempt for others based on race or gender is not just rationally ungrounded—because the intentional content of the emotion presents racial and gender status as indicating lesser moral status, which is not the case—but the *morally wrong* way to feel. Smith reasons that merely *having* such emotions breaches the claim to moral regard that others have on us, simply in virtue of being in relationship with one another as moral agents.[12]

Racist and sexist contempt present a serious challenge to anti-moralism about the emotions. The idea that there is nothing *intrinsically morally wrong* with having these or similar emotions is counterintuitive, to say the least.[13] But intuition is not argument. What ought we to think about these emotions?

To consider this question, I want to begin by putting the relationship of "mutual recognition" aside. One thing that *is* clear is that the norms of garden-variety relationships include emotions within their remit. Smith herself discusses the norms of friendship. These clearly involve *caring* about a person. If you had a friend who successfully walked the walk and talked the talk of friendship on the outside, but, unbeknownst to you or anyone else, despised you on the inside, they would not be a good friend. Should you learn of their feelings, you would have a right to object that, *qua friend*, that they did not feel as friends ought. However, not all relationships involve norms for emotions. Consider the relationship you have with your boss. If they, entirely unbeknownst to you or anyone else, privately despise you,

12. For arguments in support of a similar claim in relation to *beliefs*, see Basu (2019), Marušić and White (2018), Nolfi (2018), Paul and Morton (2018), and Schroeder (2018). For a dissenting view, see Begby (2018).

13. See Macalaster Bell (2013) for discussion of the nature of contempt and defense of the view that, in the right circumstances, it can be rationally fitting and instrumentally valuable.

you may not like it, but you have no right to object to it *qua employee*, so long as they have *treated* you fairly and respectfully throughout your employment. Some garden-variety relationships include norms not only in relation to our actions but also in relation to our emotions—but some do not.

Now consider the relationship of "mutual recognition." This is perhaps the most abstract and encompassing of all relationships. It must hold between us and strangers as much as between us and intimates—it covers each and every moral agent in the moral community. Given its breadth, it is a substantial obligation on us if it is to include emotions within its remit. Recall that, unlike actions and beliefs, emotions are neither directly responsive to the will, nor reliably responsive to reason. With this in mind, when assessing whether or not emotions like racist and sexist contempt are the morally wrong things to feel—when they are entirely private, in no way expressed or acted on or the cause of any actual harm to anyone—there are two kinds of case that must be distinguished.

There is the case of the person who has grown up, as we all have, in our racist and sexist world, and feels contempt *despite themselves*. They believe the representation of people of color and women embodied by their involuntary feeling of contempt is rationally ungrounded: they recognize the world is not as their emotion presents it as being. They would never consider expressing it or acting on it. Indeed, they may well hate themselves for feeling it. They do all they can to challenge the feeling and exorcise it from their psychology: it is entirely egodystonic. But their contempt is recalcitrant. Much as they work to rid themselves of the feeling, it doesn't ever go away.

Then there is the case of the person whose beliefs about people of color and women and disposition to expression and action accord perfectly with their feeling of contempt. They *assent* to the feeling: it is entirely ego-syntonic. As we are imagining the example, they do not in fact express or act on their contempt: no actual harm is done. But they are in principle if not in practice perfectly willing.

These cases represent extreme and hence simplified ends of a continuum. There will, of course, be myriad, complicated cases that lie somewhere in between and blend elements of both; what to say about these cases will be comparably complicated. But I think it is less complicated what to say about the extreme cases, which then provides a basis from which to approach any others.

It is bad for the person who feels recalcitrant contempt that they feel it, and no doubt they are as likely to morally reprimand themselves for how they feel as we are likely to morally reprimand them. But I do not see why they violate the relationship of "mutual recognition" with other moral agents by feeling, despite themselves, as they do. For they *recognize* the moral facts very clearly: race and gender are not indications of lesser moral status or grounds for disregard or disrespect; and they believe and act entirely in accordance with this fact.

On the other hand, the person who embraces their contempt does seem at least potentially to violate the relationship of "mutual recognition" with other moral agents. *But this is because they embrace it*. They assent to the feeling that race and

sex are indications of lesser moral status and grounds for disregard and disrespect; it accords with their beliefs and they are perfectly willing in principle to express and act on it. In other words, it is not the feeling that is morally wrong, *but what they do with it*. This clinical distinction between emotions and actions—which can include mental actions, such as e.g. assent—is the basis for distinguishing these two extreme cases, and judging the person who embraces their contempt differently from the person who does not. Hence, even in these hard cases, I think it is far from clear that moralism about the emotions—given a proper appreciation of the nature of emotions—is true. The feeling of contempt based on race or gender is always ugly, no doubt; but in itself, *it* is not the appropriate object of moral assessment—in contrast to how it is actively taken up by the subject, woven into their thinking and their acting.[14]

In part 1, I promised to return to the relationship between moralism about the emotions and virtue ethics. I am now in a position to do so. Following Aristotle (1980), virtue ethics grounds morality in the ideal of the virtuous person, who is famously required to feel, relative to the circumstances, the right way, to the right degree, for the right amount of time, on pain of failing to be virtuous—a picture that appears to embrace moralism. Suppose that, in contrast to what I suggested in part 1, deontic assessment is defined not in relation to the possibility of attributions of moral responsibility, but in relation to the virtuous person and their character: what they are disposed to think, feel, and do. I shall take it as given that most virtue ethicists would agree that the virtuous person is not disposed to feel racist or sexist contempt, whether recalcitrant or otherwise. Then, according to virtue ethics, such contempt is the morally wrong way to feel. To use Adams's evocative phrase from the quotation in part 1, it is an "involuntary sin."

Thus understood, anti-moralism about the emotions is inconsistent with virtue ethics; and the arguments I have offered in this part against moralism about the emotions tell against virtue ethics. Insofar as virtue ethics depends on the idea of *human nature*, which includes a set of basic emotions in response to specific

14. Iris Murdoch (1970) famously describes a related case: a mother-in-law M, who behaves beautifully to her daughter-in-law D, but nonetheless harbors hostility toward D within. M occupies a position between the person who experiences recalcitrant contempt and the person who embraces their contempt. Like the person who experiences recalcitrant contempt, she tries to rid herself of her hostility by "looking again" at D; but, as Murdoch tells the story, unlike them, she succeeds—she comes to see D in a different light, so that her feelings, thoughts, and actions toward D are both positive and in harmony with each other. Murdoch is concerned to emphasize that D's attempt to rid herself of her inner hostility is morally valuable notwithstanding the fact that it makes no difference to her external behavior. This can make it seem as if she straightforwardly embraces moralism about the emotions. However, similarly to my discussion above, Murdoch explicitly locates M's moral achievement less in the outcome (namely, an overall positive disposition to D) and more in her *inner, mental activity* of reflecting on her hostility and striving to see D in a new light. Suppose M tries as hard as she can but fails to fully eradicate the emotional component of her hostility to D: recalcitrant glimmers of jealousy and contempt continue to erupt occasionally, despite her sincere and ongoing attempts to rid herself of these feelings. If so, would Murdoch see M as less morally good than how she is originally described? Perhaps. But would she be right to?

circumstances as an essential part of ethics and a life of flourishing, this clash is no surprise. For, I tried to suggest not only that there is much more human emotional diversity than this picture permits, but that it is a mistake to moralize whatever biological norms there in fact are. Nonetheless, there is a way to reconcile anti-moralism about the emotions with at least the structure of virtue ethics. This is to reject the focus on emotions, and replace it with an emphasis on thought and action as the fundamental components of virtuous character.[15] This would allow a sharp distinction between the person who feels recalcitrant contempt and the person who embraces their contempt, precisely because the emotion of the latter, and not the former, is woven into their thinking and acting. The person who embraces their contempt is not virtuous, because they are racist and sexist in how they think and act: the fundamentals of moral character. In contrast, the person whose contempt is recalcitrant *is* virtuous—no matter their contempt—because it is not woven into their thinking and acting, and so not fundamental to their moral character. This is not to deny that their contempt *says something* about who they are. For example, it reveals that they have grown up in a racist and sexist world, which they carry within them despite themselves. Equally, perhaps we might like them better, and they might find it easier to like themselves, if they did not feel this way. But they are virtuous nonetheless—insofar as they do not deviate from the standard set by a virtue ethics that emphasizes thought and action as opposed to emotion.[16]

I initially developed the argument that even emotions like racist and sexist contempt are not ipso facto morally wrong through consideration of the way that a garden-variety relationship like friendship can include norms for emotions. But friendship is, in most ways at least, a *voluntary* relationship. The friend who walks the walk and talks the talk of friendship, but despises you, not only fails to feel as a friend ought: they also fail *to end the friendship* when, given that they despise you, they should. But some relationships that, in our culture at least, include norms for emotions cannot be ended. For example, mothers are supposed to love their children (and vice versa, of course); but this is not a relationship that either party can end. To be sure, they can walk away from each other; but the fact that they stand in this relationship endures rupture, abandonment, and rejection.[17]

Suppose that, as with Srinivasan's example of the woman who cannot feel anger at injustice, there is a mother who finds that she does not love her child. However,

15. Thus reconfigured, the virtuous person and their ethical function are interestingly similar to Michael Smith's conception of an ideal rational agent who functions like an *advisor*, providing nonideal agents with guidance about what to do in the circumstances (Smith, 2020).

16. To return to the example of the jealous sibling in part 1 whose jealousy is due not to low self-esteem but to selfishness: what deserves moral reproach is the assent to, embrace of, and willingness to act on, the idea that one should get more than one's share; not the jealousy per se.

17. I do not base this claim on biological parenthood: it is as true for adoptive mothers and stepmothers. Once you have become a parent, whatever the causal route by which you became one, you cannot stop being a parent even if you stop *parenting* the child.

to the very best of her ability, she treats the child right—she does everything by the ethical book.[18] I do not pretend that this is the kind of case where the failure to feel can be kept entirely private; let us acknowledge that, despite the fact that the mother treats the child right to the very best of her ability, the child feels the lack of love, and this is a terrible thing for the child. In all likelihood, the mother also feels absolutely devastated by this fact and punishes herself for it. In my experience, although it is far from common, women who were victims of extreme childhood abuse may sometimes struggle to feel love for their own children; so too may women who did not choose to have a child, for example, because they were forced or raped.

Any mother who does not love their child violates the emotional norms governing motherhood in our culture. But we are not therefore required to hold that such a mother fails to feel as they ought, *qua mother*. Rather than accept this conclusion, we may instead choose to challenge the emotional norms governing our cultural conception of motherhood: to reject the prescriptive force of "mothers love their children" even while we concur with its descriptive truth in the majority of instances.[19] To be absolutely clear, I am not denying the tremendous emotional and ethical importance of the relationship between mothers and their children—and the value to all parties of love. As I suggested in part 1, emotions can have intrinsic value. I am rather suggesting that, when we find an example of a relationship whose norms include emotions within its remit, it may be open to us to refuse to accept these norms as requirements on the relationship, rather than agree that their violation constitutes a wrong relative to the relationship.

I have argued that there are moral costs to our common practice of moral reprimand of emotions and described a clinically-derived alternative way of theorizing about and engaging with emotions; and I have offered a short, simple argument in favor of anti-moralism. The possibility of offsetting the moral costs of our practice of moral reprimand therefore depends on successfully arguing against anti-moralism and in favor of moralism. For, if moralism is true, then moral reprimand would at least have the moral value of *censuring* that which is indeed morally wrong, thereby upholding the moral code that governs emotions. But is there such a code? Despite the intuition that there is, I have argued that there is not. But I have not, of course, conclusively refuted moralism. Rather, I hope to have established that, as things stand, moralism lacks sufficient evidence in its favor to make the practice of moral reprimand worth it. We should be on our guard against the moral *creep* from wrong actions to wrong feelings—and with it the

18. Matthew Liao (2015) has argued that children have a right to be loved; I am not persuaded by this claim, but, if it is true, then a mother who does not love her child may be required to find the child a parent who will love them in order to count as treating the child right.

19. The possibility of challenging the emotional norms governing our cultural conception of motherhood is all the more pressing when one considers that the emotional norms governing our cultural conception of fatherhood are, by comparison, lax. For a groundbreaking qualitative study that gives voice to mothers who regret having had children and critically interrogates the "feeling rules" governing women in motherhood, see Donath (2017).

all-too-ubiquitous tendency to stand in judgment[20]—in absence of a convincing argument. In the meantime, surely we have enough to be getting on with if we reprimand people only for their actions and not for their emotions. Stop telling people what to feel!

ACKNOWLEDGMENTS

I would like to thank Hilary Bok, Steve Darwall, Daniela Dover, Johann Frick, Trip Glazer, Liz Harman, Grace Helton, Ted Hinchman, Sarah Hirschfield, Doug Husak, Mark Johnston, Harvey Lederman, Doug Maclean, Jessica Moss, Ekédi Mpondo-Dika, Shaun Nichols, Sarah Paul, Ian Phillips, Connie Rosati, Nishi Shah, Walter Sinnott-Armstrong, Michael Smith, Sigrún Svavarsdottir, Elanor Taylor, Kevin Toh, David Velleman, Andrea Westlund, Susan Wolf, and especially Monique Wonderly, for invaluable comments on this paper.

REFERENCES

Adams, R. M. 1985. "Involuntary Sins." *Philosophical Review* 94(1): 3–31.
American Psychiatric Association (APA). 2013. *Diagnostic and Statistical Manual of Mental Disorders*, 5th edition. Washington, DC: American Psychiatric Association.
Aristotle. 1980. *Aristotle's Nicomachean Ethics*. Translated by W. D. Ross, revised by J. D. Ackrill and J. O. Urmson. Oxford: Oxford University Press.
Basu, R. 2019. "The Wrongs of Racist Beliefs." *Philosophical Studies* 176: 2497–515.
Begby, E. 2018. "Doxastic Morality: A Moderately Skeptical Perspective." *Philosophical Topics* 46(1): 155–72.
Bell, M. 2013. *Hard Feelings: The Moral Psychology of Contempt*. Oxford: Oxford University Press.
Boorse, C. 1975. "On the Distinction between Disease and Illness." *Philosophy & Public Affairs* 5(1): 49–68.
Boorse, C. 1977. "Health as a Theoretical Concept." *Philosophy of Science* 44(4): 542–73.
Brady, M. S. 2008. "The Irrationality of Recalcitrant Emotions." *Philosophical Studies* 145(3): 413–30.
Campbell, S. 1997. "Emotion as an Explanatory Principle in Early Evolutionary Theory." *Studies in the History and Philosophy of Science* 28(3): 453–73.
D'Arms, J., and D. Jacobson. 2000. "The Moralistic Fallacy: On the 'Appropriateness' of Emotions." *Philosophy and Phenomenological Research* 16(1): 65–90.
D'Arms, J., and D. Jacobson. 2003. "The Significance of Recalcitrant Emotion." In *Philosophy and the Emotions*, edited by A. Hatzimoysis, pp. 127–45. Cambridge: Cambridge University Press.
Darwin, C. 1998. *The Expression of the Emotions in Man and Animals, with an Introduction, Afterword, and Commentary by Paul Ekman*. London: HarperCollins Publishers.
Deonna, J. A., and F. Teroni. 2012. *The Emotions: A Philosophical Introduction*. Abingdon, GB: Routledge.
Donath, O. 2017. *Regretting Motherhood: A Study*. Berkeley, CA: North Atlantic Books.
Dover, D. 2019. "Criticism as Conversation." *Philosophical Perspectives* 33(1): 26-61.
Ekman, P. 1992. "An Argument for Basic Emotions." *Cognition and Emotion* 6(3/4): 169–200.

20. See Watson (2013) for an argument that judgmentalism is a moral fault and Dover (2019) for an argument that moral philosophy has been far too focused on our practice of standing in judgment as a model for all forms of ethical engagement.

Ekman, P. 1989. “The Argument and Evidence about Universals in Facial Expressions of Emotion.” In *Handbook of Social Psychophysiology*, edited by H. Wagner and A. Manstead, pp. 143–64. New York: John Wiley and Sons.
Glenn, A. L. 2017. “Personality Disorders and Character.” In *Moral Psychology, Volume 5, Virtue and Happiness*, edited by W. Sinnott-Armstrong and C. Miller, pp. 497–520. Cambridge, MA: MIT Press.
Goldie, P. 2004. “Emotion, Feeling, and Knowledge of the World.” In *Thinking about Feeling: Contemporary Philosophers on Emotions*, edited by Robert C. Solomon, pp. 91–106. Oxford: Oxford University Press.
Harman, G. 2009. “Guilt-Free Morality.” *Oxford Studies in Metaethics* 4:203–14.
Johnston, M. 2001. “The Authority of Affect.” *Philosophy and Phenomenological Research* 63(1): 181–214.
Lane, R. D., L. Ryan, L. Nadel, and L. Greenberg. 2015. “Memory Reconsolidation, Emotional Arousal and the Process of Change in Psychotherapy: New Insights from Brain Science.” *Behavioral and Brain Sciences* 38:1–19.
Liao, M. 2015. *The Right to Be Loved*. Oxford: Oxford University Press.
Marušić, B., and S. White. 2018. “How Can Beliefs Wrong?—A Strawsonian Epistemology.” *Philosophical Topics* 46(1): 97–114.
Meffert, H., V. Gazzola, J. A. den Boer, A. A. J. Bartels, and C. Keysers. 2013. “Reduced Spontaneous but Relatively Normal Deliberate Vicarious Representations in Psychopathy.” *Brain* 136:2550–62.
Moore, M. 2010. *Placing Blame: A Theory of the Criminal Law*. Oxford: Oxford University Press.
Murdoch, I. 1970. *The Sovereignty of Good*. London: Routledge.
Nolfi, K. 2018. “Moral Agency in Believing.” *Philosophical Topics* 46(1): 53–73.
Paul, S., and J. Morton. 2018. “Believing in Others.” *Philosophical Topics* 46(1): 75–96.
Pickard, H. 2011. “Responsibility without Blame: Empathy and the Effective Treatment of Personality Disorder.” *Philosophy, Psychiatry, Psychology* 18:209–23.
Pickard, H. 2013a. “Responsibility without Blame: Philosophical Reflections on Clinical Practice.” In *The Oxford Handbook of Philosophy and Psychiatry*, edited by K. W. M. Fulford et al., pp. 1134–52. Oxford: Oxford University Press.
Pickard, H. 2013b. “Irrational Blame.” *Analysis* 73(4): 613–26.
Pickard, H. 2015. “Stories of Recovery: The Role of Narrative and Hope in Overcoming PTSD and PD.” In *Oxford Handbook of Psychiatric Ethics, Vol. 2*, edited by J. Z. Sadler, K. W. M. Fulford, and C. W. van Staden, pp. 1315–27. Oxford: Oxford University Press.
Roberts, R. C. 2003. *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press.
Scanlon, T. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
Schroeder, M. 2018. “When Beliefs Wrong.” *Philosophical Topics* 46(1): 115–27.
Shargel, D. 2015. “Emotions without Objects.” *Biology & Philosophy* 30: 831–44.
Smith, Angela M. 2011. “Guilty Thoughts.” In *Morality and the Emotions*, edited by Carla Bagnoli, pp. 235–56. Oxford: Oxford University Press.
Smith, M. 2007. “Meta-ethics.” In *The Oxford Handbook of Contemporary Philosophy*, edited by F. Jackson and M. Smith, pp. 3–30. Oxford: Oxford University Press.
Smith, M. 2020. “The Modal Conception of Ideal Rational Agents: Objectively Ideal not Merely Subjectively Ideal, Advisors not Exemplars, Agentially Concerned not Agentially Indifferent, Social not Solitary, Self-and-Other-Regarding not Wholly Self-Regarding.” In *Explorations in Ethics*, edited by D. Kaspar, pp. 59–79. New York: Palgrave Macmillan.
Srinivasan, A. 2018. “The Aptness of Anger.” *Journal of Political Philosophy* 26(2): 123–44.
Tappolet, C. 2016. *Emotions, Values, and Agency*. Oxford: Oxford University Press.
Watson, G. 2013. “Standing in Judgment.” In *Blame: Its Nature and Norms*, edited by D. J. Coates and N. A. Tognazzini, pp. 282–302. Oxford: Oxford University Press.
Wiggins, D. 1987. “A Sensible Subjectivism?” In his *Needs, Values, Truth*. Oxford: Clarendon Press. 185–214.
Wollheim, R. 2004. *Germs: A Memoir of Childhood*. London: Waywiser Press.