

The Reasonable and the Moral

1. Introduction: The Status of Reasonable Contractualism

In *What We Owe to Each Other*¹ T.M. Scanlon defends the view that an act is wrong if and because “its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced general agreement” (153). This theory of what it is to wrong another is novel in at least two important respects. First, in contrast to other thinkers who also deem wrongness to be a function of unreasonableness, for example, Samuel Clarke and John Balguy,² Scanlon conceives of unreasonableness as a feature of a hypothetical contract. Whereas others conceive of wrongness in terms of *acts* that are themselves unreasonable, Scanlon conceives of wrongness in terms of acts disallowed by *agreements* that could be rejected only on unreasonable grounds. Second, unlike other hypothetical contractualists such as John Rawls,³ Scanlon takes contractualism to be more than a mere device of representation. Others deem the imaginary consensus among a group of parties simply to *indicate* the features making an act wrong, whereas Scanlon takes the hypothetical agreement to *constitute* the property of wrongness itself.

In this essay, I consider whether these developments are justified. My thesis is that they are not. I defend this thesis by critically exploring the explanatory power of Scanlon’s contractualism. A theory of other-regarding moral requirements is often evaluated in terms of three considerations: the degree to which it entails only considered judgments about duties to others, entails all considered judgments about duties to others, and explains such judgments. I focus on the latter criterion, questioning the

¹T.M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998). Parenthetical page citations refer to this text.

²Samuel Clarke, *A Discourse of Natural Religion*, repr. in D.D. Raphael (ed.), *British Moralists 1650-1800*, Volume 1 (Indianapolis: Hackett, 1991), pp. 190-209; John Balguy, *The Foundation of Moral Goodness*, repr. in Raphael, *British Moralists*, Vol. 1, pp. 397-408.

³John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), esp. pp. 24-28.

degree to which Scanlon's theory explains firm intuitions about wrongness.⁴ I develop an alternative account of the property in virtue of which actions are wrong, an account that retains the notion of unreasonableness but rejects the contractualist framework. Specifically, I maintain (roughly) that the property of treating another unreasonably better explains what makes an act wrong than does the property of it being prohibited by principles that contractors with an ideal motivation could not reasonably reject.

I begin by noting that Scanlon's theory has *prima facie* difficulty entailing that we have direct duties to nonrational beings (section 2) and that there are degrees of wrongness (section 3). It turns out that contractualism can in fact entail these intuitions, but, I suggest, only at the cost of a relatively implausible explanation of them. A more plausible explanation, on the face of it, appeals to the idea that certain actions are themselves unreasonable. To flesh out this proposal, I construct a new theory of unreasonable action, one that is analogous to commonsensical judgments of unreasonable emotion (section 4). And I then argue that the theory of wrongness *qua* unreasonable action does a better job of accounting for several important aspects of wrongness than Scanlon's theory (section 5). I conclude by indicating what I consider to be immensely useful about Scanlon's book and by pointing to some avenues for future research into the connection between the reasonable and the moral (section 6).

⁴I believe that the contractualist formula, as it stands, fails to pick out only duties to others. Recall that Scanlon's parties aim for agreement about rules "for the general regulation of behavior," meaning that they seek principles to govern a person's voluntary decisions (p. 269) when and only when they bear on life with others (p. 154). Parties seeking principles "for the general regulation of behavior," so understood, select rules that have a scope intuitively extending beyond the domain of what we owe to each other. For example, the parties to Scanlon's contract could not reasonably reject a principle requiring cooperation in prisoner dilemma situations. Scanlon's contractors, in addition, must accept on pain of unreasonableness the principle of the hypothetical imperative, at least when applied to interaction with others; they could not reasonably reject a rule requiring individuals to take the necessary and sufficient means to a collective end (or to forsake the end). Furthermore, those seeking an agreement that could not be reasonably rejected must accept maxims requiring individuals to think for themselves and to listen to others when interacting.

These examples have a common theme: they are principles that could not be reasonably rejected because of a general interest in promoting ends in cooperative situations. Although there can sometimes be immoral features about behavior violating these unrejectable principles, the point is that it is not necessarily immoral to violate them. For instance, breaking a rule prescribing cooperation in prisoner dilemma situations is not always intuitively wrong. Therefore, Scanlon must narrow the class of principles on which his parties seek to find agreement. Remaining within the contractualist framework, I suggest that Scanlon simply construe his parties as seeking principles not "for the general regulation of behavior" but instead for "the domain of what we owe to each other." This is the most direct way to solve the problem, and is, I suspect, the only way to do so.

2. Wronging the Nonrational

Scanlon's theory has difficulty accommodating intuitions about our duties to nonrational beings. This sort of objection is often raised against contractualist theories. I want to explore it because Scanlon has some novel ways of responding to it. Critical discussion of his responses points the way to what I believe is a more plausible account of wrongness.

The straightforward concern about any contractualist theory of morality is that it fails either to entail or to explain adequately the wrongness of mistreating beings who cannot make agreements. In a contract theory, moral status is represented in terms of being party to an agreement. Since one can make an agreement only if one is capable of certain complex forms of rational behavior, it appears that a contract theory implies that nonrational beings are not worthy of moral consideration. But this is counterintuitive; it is a serious moral wrong to treat a cat or a mentally retarded baby as a soccer ball.

Scanlon discusses three different ways that his contractualist theory can blunt the force of this objection. First, he notes that a party to his contract can reasonably reject a principle on the ground that it prevents him from pursuing what Scanlon calls "impersonal value" (pp. 218-23). A personal value is a consideration bearing on the life of an individual that provides a positive reason for action. An impersonal value also grounds a positive reason for action, but not insofar as it affects a person's life. Examples of impersonal values are protecting a rare species or saving the redwoods. Scanlon's theory implies that impersonal values cannot ground a reasonable rejection of a principle, but he argues that they may (even must) figure indirectly into a reasonable rejection. Specifically, a person can reasonably reject a principle for the reason that its observance would not sufficiently accommodate the personal value of being able to pursue impersonal value. It is therefore open to parties to reject principles that would severely restrict their ability to realize the impersonal value of caring for individual animals or retarded humans.

This aspect of Scanlon's theory might constitute part of an acceptable response, but clearly will not do as a complete answer (and Scanlon does not intend it to be). All it does is make space to reject a principle that would prohibit someone from imposing a light burden on another person (e.g., breaking a promise to meet for coffee) for the sake of benefiting a non-person (e.g., taking a wounded animal to a shelter). It does not support the judgment that it would be wrong to impose a serious burden on a non-person for a trivial benefit to a person. Someone's being permitted to help a nonrational being does not imply that anyone is required not to harm such a being. Therefore, something more is needed to buttress Scanlon's view.

A second line of response is to maintain that it is wrong to impose serious burdens upon nonrational creatures for trivial benefits but to suggest that this sense of “wrong” is not the one in which an agent has failed to give another what she is owed (pp. 178, 181-84).⁵ Scanlon distinguishes between broader and narrower concepts of wrongness. The broad concept of wrongness is, roughly, the notion of violating norms so as to warrant serious criticism, say, blame or guilt. One subset of this idea of wrongness is the narrow concept, which is the idea of wronging an individual. Wrongness in the narrow sense is not merely doing wrong, but wronging another, not according her the treatment to which she is entitled. Scanlon invokes this distinction to suggest that while mistreating nonrational beings is wrong in the broad sense, it is perhaps not wrong in the narrow sense in which he is interested. Hence, his contractualist theory cannot be faulted for failing to account for duties to nonrational beings, since his theory is meant to capture only what we owe to each other.

This line obviously enables Scanlon to claim that people are required not to harm nonhuman animals and nonrational humans. The problem is that there are strong reasons for thinking that one can indeed wrong such individuals. Scanlon himself mentions one reason, namely, that torturing an animal seems to be something “for which we should feel guilty *to* the animal itself, just as we can feel guilt to a human being” (182). I presume that feeling “guilty *to*” an animal means feeling bad for the way one treated that individual. One criticizes oneself for having imposed an unnecessary burden on an innocent creature, which differs from, say, reproaching oneself for not having lived up to a code of honor.

There are additional considerations for thinking that one can wrong a nonrational being. For one, it appears as though torturing an unowned animal or a severely disabled human being requires compensating that particular being. One must pay the medical bills to rectify any harm that one did to him. These individuals can be owed compensation, which suggests that they were initially owed better treatment, too. Furthermore, the reason not to torture animals for fun is quite strong, indeed stronger than, say, the duty not to make a false promise in order to get a small sum of money. Scanlon considers the fact that an action wrongs another individual to constitute “sufficient reason not to do it (almost?) no matter what other considerations there might be in its favor” (148; see also p. 334). If indeed the narrow sorts of right and wrong “always or even almost always take precedence over other values” (148), then we have good reason to deem at least some duties to nonrational beings to fall within the narrow class; for many duties to nonrational beings, particularly those not to harm them seriously for trivial benefits, are just as overriding in their normative force as

⁵See also pp. 6-7, 171-77, 342-49, 357.

duties to rational beings.

For all these reasons, it is worth considering whether Scanlon's theory could entail that it can be wrong in the narrow sense to treat nonrational beings in certain ways. Scanlon's third line of response, the idea of trusteeship, aims to show this. On this interpretation of Scanlon's theory, a trustee, who represents the claims of beings unable to contract, would be party to the agreement and could reasonably reject a principle on the ground that it does not sufficiently accommodate the claims of his clients. This maneuver solves the problems facing the earlier two lines of response. It enables Scanlon to say that there are requirements not to harm nonrational beings, requirements the violation of which is "wrong" in the narrow sense of failing to give another what is owed.

Although this version of Scanlon's contractualism can entail the duties that many intuitively think we have with respect to nonrational beings, its explanation of these duties is dubious. A theory of morality not only ought to entail considered moral judgments, but also ought to provide a compelling explanation of them. Scanlon's fundamental account of why it is wrong to be cruel to nonrational beings is that a trustee representing their claims could reasonably reject principles that permitted such treatment. If one has taken the step of admitting that nonhuman animals and retarded humans can be owed certain treatment, then one is unlikely to find this plausible. It seems too indirect to constitute the basic explanation. Instead, the reason why the trustee could reject the principle permitting the act is the natural candidate for the fundamental explanation of why the act is wrong.

To press the point, consider the start of an alternative explanation of why it is wrong to torture an animal for fun: such an action is unreasonable. It is more plausible to explain the wrongness of unnecessary cruelty in terms of *an unreasonable balance of interests* than in terms of *a trustee being able to reject on reasonable grounds a principle that balances interests in this way*. The difference here is between appealing to the unreasonableness of the way an agent treats a nonrational being, on one hand, and invoking the reasonableness of a trustee rejecting principles affecting his client, on the other. Below I will develop this alternative theory in some detail and argue that it provides a better explanation not only of wronging the nonrational but also of wrongness generally. For now, however, I wish merely to note that this sketch of a rival account highlights the respect in which Scanlon's theory provides an implausible explanation of our duties to nonrational individuals, assuming there are such.

3. Degrees of Wrongness

Wrongness admits of degrees. In the first place, some actions are more wrong than others. There is a ranking of wrong acts, so that, for example,

killing one's spouse for the insurance money is morally worse than breaking a promise to meet a student in order to play pinball. Not only do we judge that some acts are more wrong than others, but we also seem to judge the rough extent to which one act is more wrong than another. For instance, the degree to which killing for money (A) is worse than breaking a promise for amusement (B) is greater than the degree to which breaking a promise for amusement (B) is worse than forgetting to pay for one's share of office coffee (C). The space between A and B on the scale is larger than the space between B and C. Although one is hard-pressed to say exactly how much space there is between these acts, surely wrongness has degrees of at least the two sorts just noted.

It would therefore be a major strike against a theory if it were not able to account well for the ways in which wrongness admits of degrees. Rivals to Scanlon's contract theory are able to entail with relative ease that there are degrees of wrongness. The divine command theorist can do so in terms of degrees of God's disapproval or rankings of God's preferences. The Kantian can conceive of degrees of wrongness in terms of degrees of disrespect for agency, with killing for money degrading a person or attacking her rational capacities to a greater extent than breaking a promise. And the utilitarian can say that killing innocents is morally worse than breaking a promise, since the former tends to have worse consequences than the latter.

On the face of it, Scanlon's theory cannot entail that there are degrees of wrongness. The standard version of Scanlon's formula says that an act is wrong insofar as it would be forbidden by a principle that could not be reasonably rejected (by those seeking principles that no one can reasonably reject). The problem is that there are no degrees to which a principle could not be reasonably rejected. Being unable to reasonably reject a principle consists of there being no good reason (of a sort) to reject it. And there are simply no degrees to the condition of there being no good reason to reject a principle. Since there are equally no reasonable grounds for rejecting principles that forbid killing for money and breaking a promise for amusement, Scanlon's theory apparently cannot account for any difference in the degree to which these acts are wrong.

The best way for Scanlon to respond to this worry, I believe, is to appeal to a version of his theory that is logically implied by the standard one. Not being able to reasonably reject a principle forbidding an act means being able to reasonably reject every principle permitting that act (pp. 4, 195). If a party has no reasonable grounds for rejecting a principle that prohibits killing for money, then she has reasonable grounds for rejecting any principle that allows killing for money. The same goes for a principle regarding breaking promises for amusement. Now, while there are no degrees of there being no reasonable grounds, there do seem to be degrees of reasonable grounds. Scanlon can therefore seek to account for the degree

of an act's wrongness in terms of the degree to which the parties to his contract could reasonably reject principles that permit that act. So, not only are there sufficient reasonable grounds to reject any principles that permit either murder for money or breaking a promise for amusement, but also, Scanlon might say, there are stronger reasonable grounds to reject any principles permitting the murder than there are to reject any principles permitting the promise-breaking.

Scanlon's theory successfully entails that there are differences in the degrees of wrongness. But we must also consider the ability of his theory to explain them. Think about this alternative strategy for explaining degrees of wrongness: some acts are more unreasonable than others. What makes killing one's spouse for the insurance money morally worse than breaking a promise so that one can play pinball is, roughly, that the former treats a person more unreasonably than the latter. This rival explanation conceives of degrees of wrongness in terms of a property of actions, whereas Scanlon's explanation accounts for degrees of wrongness in terms of a property of the ability of ideal contractors to reject principles permitting actions. My suggestion is that the unreasonableness of the act itself explains more naturally and deeply the degree of an act's wrongness than the reasonableness of rejecting principles that permit this act.

4. A Theory of Unreasonable Action

Sections 2 and 3 have a common structure. They both begin by noting that Scanlon's theory on the face of it cannot entail certain intuitions about wrongness. They then suggest ways of understanding Scanlon's theory so that it can entail these intuitions, but in the end question the ability of his theory to explain them. In sum, the concern is that Scanlon's theory can entail commonsensical judgments about morality only at the cost of an implausible explanation of them. In this section I begin to develop in detail the objection that there is a more fundamental explanatory property than reasonable rejection among ideal contractors, namely, the unreasonableness of actions not inherent to a contractual framework. By drawing an analogy with unreasonable emotions, I here develop a theory of wrongness *qua* unreasonable action. In the following section, I argue that this theory explains key features of wrongness better than Scanlon's theory.

The objection I am making against Scanlon's theory differs from some related objections that have been made. First, it is different from the objection that one can adequately explain the wrongness of actions merely by noting the diverse properties that make them wrong under various conditions. For example, Simon Blackburn suggests that appealing to things such as misery, pain, or death provides a sufficient explanation of immoral-

ality, so that reasonable contractualism does not add anything.⁶ Contra Blackburn, I think it would in fact be revealing to find a unifying element among the diverse properties in virtue of which they constitute wrongness. Identifying a common denominator among wrong actions, viz., providing a theory of immorality, would have explanatory power. My objection differs from Blackburn's in that I maintain that there is a theory that does a better job of accounting for the nature of immorality than Scanlon's.

My objection also differs from the claim that a judgment that a principle could be reasonably rejected must be grounded upon a more basic judgment that the principle is immoral. Judith Jarvis Thomson suggests this kind of objection when she says that "it is the patent wrongfulness of the conduct that explains why there would be general agreement to disallow it."⁷ Unlike Thomson, I do not say that one can reasonably reject a principle only by making the prior judgment that an act permitted by the principle falls under the description "wrong." I agree with Scanlon that wrongness can be understood in terms of a more fundamental notion of unreasonableness. I just deny that Scanlon's sort of unreasonableness is the correct notion, for I claim that actions qua unreasonable are more fundamental than the unreasonableness of accepting or rejecting principles permitting actions.

Furthermore, the objection I make here differs from the claim that contractual consensus is reducible to individual choice. Elsewhere I have argued that the procedure of collective agreement in all moral hypothetical contracts (not just those with a veil of ignorance) is superfluous, since the parties are conceived as reasoning in the same way.⁸ That objection maintains that Scanlon's theory is reducible to (roughly) the notion that wrongness is to be identified with actions that would be forbidden by principles that a single ideal agent could not reasonably reject. The present objection contends that wrongness is to be identified with unreasonable actions, not with actions that would be forbidden by principles that could not be reasonably rejected, whether by an individual or a group.

The theory that I develop maintains that actions are wrong just insofar as they treat others unreasonably. This theory develops some ideas of classic philosophers such as Clarke and Balguy, who conceive of wrongness in terms of actions being themselves unreasonable. For example, according to Clarke,

Some things are in their own nature good and reasonable and fit to be done; such as keeping faith, and performing equitable compacts, and the like ... *Iniquity* is the very same in *action*,

⁶Simon Blackburn, "Am I Right?" *New York Times Book Review*, 21 February 1999, p. 24.

⁷Judith Jarvis Thomson, *The Realm of Rights* (Cambridge, Mass.: Harvard University Press, 1990), p. 30 n.

⁸Thaddeus Metz, "Reason, Politics, and Contractualism," *Inquiry* 18 (1998): 61-72.

as *falsity or contradiction in theory*; and the same cause which makes the one *absurd* makes the other *unreasonable* ... It would be as impossible, that a man, contrary to the eternal reason of things, should desire to gain some small profit to himself, by doing violence and damage to his neighbour; as that he should be willing to be deprived of necessaries himself, to satisfy the unreasonable covetousness or ambition of another.⁹

Scanlon avoids calling actions themselves “reasonable” or “unreasonable”; he instead uses these terms to characterize “deciding which principles to accept” (192) or “objections to possible moral principles” (194). If Scanlon were inclined to call actions themselves “reasonable” or “unreasonable,” he would have to contend that actions are unreasonable because parties to his hypothetical social contract could reasonably reject any principles permitting them. But I maintain that the explanation goes the other way around: the fact that an act treats another unreasonably explains why ideal contractors could reasonably reject any principles permitting the act and, more generally, better explains what makes actions wrong. For example, the basic property in virtue of which a false promise for money is wrong is, as a first approximation, that it is an unreasonable act, not that parties seeking principles that cannot be reasonably rejected could reasonably reject any set of principles that would permit the act. Or at least I want to say that, insofar as the latter property has any explanatory weight, it is in virtue of tracking the former property.

Although I presume readers will find this line of reasoning *prima facie* plausible, it would be useful to develop in some detail a theory of unreasonable action, which I do in the remainder of this section. Then we can truly see, in the next section, whether Scanlon’s account is an epiphenomenon of it and its general ability to explain what makes actions wrong. I proceed here by noting intuitive respects in which emotions are unreasonable and then developing parallels with unreasonable actions.

* * *

Emotions can be reasonable or unreasonable at least when the agent can modify them in light of reasons and when they are directed toward states of affairs (which I call “objects”). That is, the word “emotions” here picks out a subset of judgment-sensitive attitudes that excludes mere feelings such as hunger or physical pain and that includes an element of “aboutness.” For example, I am interested in reactions such as anger, fear, worry, embarrassment, resentment, annoyance, delight, pride, and hope (or I am interested in these reactions insofar as they are directed toward objects and insofar as reflection can change their existence or intensity).

When we call an emotion “unreasonable” (or “inappropriate” or “irrational”), we are referring at least partially to a kind of disproportionality

⁹Clarke, *A Discourse of Natural Religion*, pp. 196, 208.

between the emotion and its object.¹⁰ This is a familiar idea.¹¹ Consider, first, that the objects of emotions come in degrees. One may have an emotional response to the occurrence of ugliness, danger, or pain, on one hand, and of beauty, benefit, and pleasure, on the other. Such states of affairs come in degrees. Think next about the fact that emotions admit of degrees of intensity. To see this, distinguish between negative and positive emotions. Negative emotions are ones that include a negative attitude, something like a wish that the object toward which the emotion is directed not obtain.¹² The more intense the negative emotion, the stronger the wish that the object not exist. For example, the more angry, regretful, or annoyed one is, the more one wishes that whatever one is angry, regretful, or annoyed about not exist. Positive emotions are, correspondingly, ones that have a positive attitude, a wish that the state of affairs toward which the emotion is directed obtain. And, again, the more intense the positive emotion, the stronger the wish that the object exist. The more one is hopeful, delighted, or proud, the more one wishes that whatever one is hopeful, delighted, or proud about obtains.

To grasp the relevant sort of disproportionality between emotions and their objects, consider some examples. It is unreasonable to be afraid *at all* of a backrub, assuming that there are no special circumstances.¹³ Fear is a negative emotion, but a backrub has no disvalue that would warrant any negative emotion, let alone the particular one of fear. Fear is instead an apt

¹⁰I am not purporting to present a complete theory of when or why an emotion is unreasonable. My goal is merely to tease out one central strand of unreasonable emotion.

¹¹Adam Smith, for one, articulates the idea well: "In the suitableness or unsuitableness, in the proportion or disproportion which the affection seems to bear to the cause or object which excites it, consists the propriety or impropriety, the decency or ungracefulness of the consequent action ... When we blame in another man the excesses of love, of grief, of resentment, we not only consider the ruinous effects which they tend to produce, but the little occasion which was given for them. The merit of his favourite, we say, is not so great, his misfortune is not so dreadful, his provocation is not so extraordinary, as to justify so violent a passion." See Adam Smith, *The Theory of the Moral Sentiments*, repr. in D.D. Raphael (ed.), *British Moralists 1650-1800*, Volume 2 (Indianapolis: Hackett, 1991), pp. 208, 209. I have little that is novel to say about this notion of disproportionality; my contribution lies in the use to which it is put, namely, invoking a relatively common account of unreasonable emotion to illuminate the more obscure idea of unreasonable action.

¹²The analogy I draw between unreasonable emotions and unreasonable actions does not depend on any of the several accounts of emotional valence in the literature. To fix ideas, I follow the account found in Robert Gordon, *The Structure of Emotions* (New York: Cambridge University Press, 1987), pp. 27-32.

¹³There is, say, no risk of paralysis from an inexperienced masseuse. Note, too, that the notion of unreasonableness here is a substantive one, fixed independently of at least the individual's (justified) set of beliefs. If a person incorrectly (but perhaps justifiably) believed that her masseuse were inexperienced, there may be a sense in which she would not be "unreasonable" to fear the backrub. I do not address this subjective sense of "unreasonable," instead maintaining that such a reaction would be unreasonable given the nature of the object.

response to the particular disvalue of danger. So, it is not unreasonable to be afraid of an electric shock. And while it would not be unreasonable to be extremely afraid of the electric chair, it would be unreasonable to be *extremely* afraid of the kind of spark that results from touching another person when static electricity is in the air. Now think about the positive emotion of pride. It is unreasonable to take *any* pride in the fact of having defaced a great work of art for the fun of it.¹⁴ There is nothing there to ground any positive emotion, let alone the particular response of pride. In contrast, it would not be unreasonable to take substantial pride in having created a great work of art, though it would be unreasonable to take *substantial* pride in having sought and found a good deal on a camera.

These intuitions suggest the following principle: an emotion is unreasonable at least insofar as it either (a) is not the right kind given the nature of the object toward which the emotion is directed, or (b) is the right kind but is too intense given the object's degree. On one hand, the object might not be worthy of the particular emotional response at all. In fearing a back-rub or taking pride in senseless destruction, one has emotional reactions toward objects that do not warrant these reactions in any respect. One wishes that a good state of affairs not obtain, and one wishes that a bad state of affairs obtain, respectively. On the other, the object might be worthy of the particular emotional response, but the response is too strong relative to the degree of the object. Being extremely afraid of minor pain and taking great pride in tracking down a good bargain are cases in which the emotions are the right kind for their objects but the emotions are too intense given the amount of dis/value. One strongly wishes that a small bad not obtain, and one strongly wishes that a small good obtain, respectively.

Some might instead suggest a more encompassing principle than the one I have proposed: an emotion is unreasonable insofar as it fails to track the kind and degree of its object. This broader principle would imply that it is unreasonable for one's emotional response to be *very weak* relative to the amount of dis/value warranting the response. For example, it would imply that one would be unreasonable not to be extremely afraid of the electric chair or unreasonable not to take substantial pride in having created a great work of art. The principle I have suggested, that an unreasonable emotion is one that is not warranted at all by the kind of dis/value or is too strong given the amount of dis/value, does not have this implication. I rest content with the more narrow formulation, since my intuitions about when emotions are unreasonable are less firm in the broader cases just

¹⁴Some readers will be inclined to say that breaking social conventions is something in which one could reasonably take some pride. However, it is presumably the courageous character that it would take to do so, and not the act of destroying a great work of art, that would make pride a reasonable response here.

mentioned and since the narrow version is sufficient to ground the analogy I want to make between unreasonable emotions and unreasonable actions.

Note that this partial account of what makes an emotion unreasonable does not appeal to the consequences of having the emotion. This does not imply that the effects of emotions never figure into the overall justification for them. For example, a feeling of guilt that is proportionate to its object could be on the whole unwarranted if it would lead to depression and thereby inhibit a person from taking care of her children. What I am concerned to highlight here is that an emotion is not unreasonable *merely* insofar as it causes harm or fails to produce something desirable. There is a *deontological* aspect to what makes the judgment-sensitive attitude of emotion unreasonable, which points the way toward a deontological account of the unreasonableness of another sort of judgment-sensitive attitude, namely, intentional action.

* * *

The above account cashes out the unreasonableness of emotions in terms of a sort of *disproportionality*. Now, the classic philosophers who think of wrongness as unreasonableness conceive of unreasonableness as equivalent to a *lack of fit*. For example, Balguy says,

Between *bounty* and *gratitude* there is a plain congruity of *moral fitness*; and between *bounty* and *ingratitude* a plain incongruity, or *unfitness*. Therefore gratitude is *reasonable*, and ingratitude *unreasonable*. Therefore the one ought to be *observed*, and the other *detested*.... Is not ingratitude to a benefactor very unfitting? Is it not therefore very unreasonable? Ought it not therefore to be abhorred and avoided by every body?"¹⁵

Balguy grounds the immorality of ingratitude upon what is unreasonable, which he in turn grounds upon what is unfit. Notoriously, neither Balguy nor anyone else who conceives of wrongness qua unreasonableness has articulated this unfitness in a revealing way. That is what I aim to do in the rest of this section, taking a cue from the disproportionality in unreasonable emotions.

Let me start by addressing the most suggestive remark that Balguy makes regarding disproportionality:

There is ... a wide difference between the nature of rational creatures, and that of brutes; and between the nature of brutes, and that of inanimate things. They require therefore respectively a suitable treatment. To treat men in the same way we treat brutes, and to treat brutes in the same way we do stocks and stones, is manifestly as disagreeable and dissonant to the natures of things, as it would be to attempt the forming of an angle with two parallel lines.¹⁶

Balguy proposes that actions may be unreasonable in virtue of failing to fit

¹⁵Balguy, *The Foundation of Moral Goodness*, pp. 405, 406. See also pp. 400-401, 406-7, as well as Clarke, *A Discourse of Natural Religion*, pp. 192-93, 196, 200-202, 208-9.

¹⁶Balguy, *The Foundation of Moral Goodness*, pp. 400-401.

certain beings. A plausible way to understand this lack of fit is, roughly, in terms of failure to be contoured to a being's intrinsic value. As a first approximation, an unreasonable action is one that responds disproportionately to the degree to which a being is good for its own sake. To make this account more precise, I first spell out the ways that beings may admit of degrees of intrinsic value and the ways that interactions with these beings may also come in degrees. Then I will be in a position to specify an account of unreasonable action qua disproportionality that parallels the account of unreasonable emotion.

Suppose that certain beings have intrinsic value, where one can rank these beings according to the degree of their value. Common intuitions support the judgment that there are natures that are good apart from their usefulness and that can be ordered according to the degree of their inherent goodness. For example, if one must choose between driving over a cat or an innocent person, one should surely run over the cat. I presume one should run over the cat not merely because one could go to jail, get sued, cause more pain, or make an uglier scene by striking the intelligent human. Part of the explanation is that people are intrinsically worth more than cats. Similar remarks apply, I believe, to a version of this thought experiment applied to cats and stones. If one must choose between running over a cat or a pebble, one should choose the pebble. To be sure, part of the initial explanation here is the pain one might cause the cat, but factoring this element out (imagine the cat will die instantly), one still finds grounds for saving the cat; ordinary cats are worth more than ordinary stones.

Familiar accounts of what make persons and animals valuable may be invoked here. Normal adult human beings have a superior inherent worth by virtue of their highly developed rational faculties. And animals are intrinsically valuable for, say, having conscious lives, that is, having beliefs, desires, emotions, feelings, and memories and engaging in goal-directed activity to satisfy interests. Is there a feature common to rational beings and certain nonrational beings that explains not only what makes them both intrinsically valuable, but also what makes the former more valuable than the latter? Worth considering (in a different context) are the proposals that it is the degree of complexity or the degree of autonomy.¹⁷

Now consider some ways to interact with intrinsic values and how they come in degrees. Distinguish between positive and negative interactions. Negative treatments of an intrinsic value are, roughly, ones that seek to undermine it or trade it off. Negative responses come in degrees, depending on the sort of impairment, causality, and intentionality involved. First, a negative response toward a being is more intense the more harsh the injury. Contributing to a person's death is a greater negative action than, say,

¹⁷For both ideas, see G.W.F. Hegel, *The Philosophy of Fine Art*, Volume 1, trans. F.P.B. Osmaston (New York: Hacker Art Books, 1975), esp. pp. 139-207.

stealing a candy bar from her. Second, a negative response toward a being is more intense the more one directly sacrifices it. Killing someone, for example, is a more intense negative treatment of her than is letting her die. Third, a negative response toward a being is more intense the more one purposefully undermines it. To continue with the example, intending to kill a person is a more intense negative treatment of her than is killing her without having foreseen this result. Positive interactions with an intrinsic value, in contrast, are ones that seek to sustain it. Treating a being positively includes refraining from negative actions and includes performing actions that protect, develop, and promote the being. And, correspondingly, a positive treatment of a being is more intense the more supportive, direct, and purposeful the action is.

Now I suggest that an action is unreasonable at least insofar as it responds to a highly valuable being as though she either (a) has none of the intrinsic value she has or (b) has less value than she does. On the one hand, an action may treat another's particularly valuable features as having *no* intrinsic value. This is to respond to the being negatively, by trading her off for something else. It is to respond to the entity as though she were "merely a means," viz., as though she has only an instrumental value. Killing one's spouse for money, breaking a promise for the sake of playing pinball, and lying to get money are good examples here. These actions destroy or stunt the exercise of a person's capacity to live according to her rational decisions for the sake of someone or something else. On the other hand, an unreasonable action may treat another as having intrinsic value, but as having *less* than she actually does. The central case would be acting on a maxim of never helping other people even when it would cost oneself little to do so. Here, one does not treat the person negatively, viz., reduce the other to a mere means, but it is unreasonable since one is not acting in a way that treats her as "an end-in-itself," as having the highest value in the world. That is, her value is such as to warrant a certain degree of positive responses but they are not forthcoming.

As with the case of unreasonable emotion, one might be tempted to propose a more encompassing account of unreasonable action: an action treats another unreasonably in failing to treat a very valuable being as having the kind and amount of value that she has. Such a principle would imply that it would be unreasonable to treat an entity as *more* valuable than it is, for example, to be extremely servile toward a dog. There may be a sense of "unreasonable" in which this is true. But the sense relevant to wronging another individual does not apply to such idolatry.

In sum, both unreasonable emotions and unreasonable actions (insofar as they bear on wronging others) are well captured by a sort of disproportionality between a judgment-sensitive attitude and an axiological condition. An emotion is unreasonable at least insofar as it either (a) is not the

right kind given the dis/valuable object toward which the emotion is directed, or (b) is the right kind but is too intense given the object's dis/value. And an action is unreasonable at least insofar as it either (a) treats another highly valuable being as having no intrinsic value, or (b) treats another highly valuable being as having less intrinsic value than it has. Let us compare the two accounts in respect of both conditions.

Regarding the (a) condition, an unreasonable emotion is not the right kind given the dis/valuable object toward which the emotion is directed. For example, in the case in which the emotion's object is something good, it is unreasonable to respond negatively to it. To respond negatively to something good is to wish that a good object not obtain. One is, in effect, *responding to it as though it were not good*. This parallels the (a) condition of unreasonable action in which one *treats an intrinsically valuable being as having no such value*. Here, one responds negatively to a being with superior worth. One undermines it or trades it off, which is not the right kind of reaction toward a being with a superlative worth.

Regarding the (b) condition, an unreasonable emotion is the right kind but is too intense given the object's dis/value. So, where the emotion's object is something good in a small respect, it is unreasonable to have an overwhelmingly positive emotional reaction to it. One strongly wishes that a good object obtain, which is, as it were, to *respond to it as more good than it really is*. This condition corresponds, although inversely, to the (b) condition of unreasonable action whereby one *treats an intrinsically valuable being as less good than it really is*. Here, one responds positively, but one does not respond positively enough. One does not interfere with the being, but one does not sufficiently go out of one's way for its sake. The treatment is the right kind, but it is not intense enough given the being's exceptional value.

Now let us return to Balguy's particular judgments of unreasonableness and see how the theory of disproportionality developed here underwrites them. Recall Balguy's claim that it is unfitting to treat a human as an animal (and an animal as a stone). To treat a human as an animal straightforwardly means trading off or failing to support a human's valuable characteristic, that is, its ability to make rational decisions, in favor of a characteristic shared with animals. A key example would be coercing a competent person in order to prevent her from causing herself pain that she foresees. The theory entails that this would treat a person unreasonably, since it would undermine the exercise of a person's rational faculties for the sake of something else.

As it stands, the theory does not yet fully account for Balguy's comments about ingratitude being unfitting and unreasonable. If one reacts injuriously to one's benefactor, then the action could be deemed unreasonable by virtue of condition (a). But what if a beneficiary is merely indiffer-

ent? To deal with this problem, and to suggest a general strategy for developing the theory of unreasonable action, I return to the case of unreasonable emotion.

Note that one can have second-order emotions, that is, emotions about emotions. For example, one can have fear of feeling guilty and hope of being pleased. A plausible way to account for the unreasonableness of second-order emotions is in terms of the unreasonableness of the first-order ones. If the object of a second-order emotion is an unreasonable first-order emotion, then the second-order emotion ought to be negative. And if the object of a second-order emotion is a reasonable first-order emotion, then the second-order emotion ought to be positive. For example, if the guilt of which one is afraid is unreasonable (there is nothing to feel guilty about), then fearing the guilt is reasonable. And if the guilt would be reasonable, then fearing the guilt would be (*pro tanto*) unreasonable.

A parallel feature of unreasonable action promises to ground the unreasonableness of ingratitude. Just as one can have emotions about emotions, one can act consequent to actions (not just in response to entities). The unreasonableness of actions performed consequent to other actions plausibly depends on the unreasonableness of the actions to which they are a reaction. If the action one is reacting to is unreasonable, then it warrants a negative reaction. Conversely, if the action one is reacting to is reasonable, then it warrants a positive reaction.

So, let us enrich the account of unreasonable action as follows: an action is unreasonable at least insofar as it (a) responds negatively to a highly valuable being (where this response is not consequent to an action she has performed), (b) responds in an insufficiently positive way to a highly valuable being (where this response is not consequent to an action she has performed), or (c) responds in an insufficiently positive way to a being consequent to her having performed an action that was more positive than was required for her to avoid the unreasonableness of conditions (a) or (b). Condition (c) accounts for Balguy's remarks about ingratitude being unreasonable. When someone has gone out of her way to support you, when she has done more for you than is owed, then you are unreasonable for not being willing to help her in return.

This recursive strategy also promises to solve another problem facing the theory of unreasonable action, that it apparently entails that it would be unreasonable ever to punish the guilty. Many believe it is not only permissible to punish the guilty, but also wrong *not* to punish the guilty (in proportion to their guilt). Since punishing a person is a negative response to an entity with a superlative value, it appears that the theory of unreasonable action must deem any punishment to be unreasonable and hence wrong. However, the distinction between actions that are consequent to other acts and those that are not can help here. Just as a reasonable action can warrant

a positive reaction such as gratitude, an unreasonable action can warrant a negative reaction such as punishment. If an action is unreasonable, then it is reasonable to respond negatively to the being who performed the action because she performed that action. So, if we prefer the strong view that proportionate punishment of the guilty is required, we can add the condition that an action is unreasonable at least insofar as it: (d) responds in an insufficiently negative way to a being consequent to her having performed an action that is unreasonable by conditions (a), (b), or (c).

This account of unreasonable action is no doubt sketchy. There are several respects in which the theory of unreasonable action outlined here requires more development, some of which I note in the conclusion. However, I believe that the theory as it stands can be shown to have more explanatory force than Scanlon's theory.

5. Unreasonable Action, Reasonable Rejection, and Explanatory Force

Having sketched out an account of unreasonable action, I want to consider its ability to explain what makes an act wrong. Here I compare the theory of wrongness qua unreasonable action with Scanlon's theory of wrongness qua reasonable rejection. First, I return to the issues of wronging the nonrational and degrees of wrongness, showing that the alternative theory accounts for them better than contractualism. Then, I provide grounds for thinking that reasonable rejection explains wrongness only to the extent that it tracks the property of unreasonable action. Next, I critically address the reasons why Scanlon is inclined to set reasonableness in the context of a hypothetical agreement. Finally, I contend that a norm of unreasonable-ness underlies several parts of morality in addition to wronging others and that the contractualist procedure clearly does not.

* * *

Recall Scanlon's explanation of why it is wrong, say, to torture nonrational beings for the fun of it. Scanlon must say that, if this is wrong in the narrow sense in which he is interested, this is because a trustee for the nonrational being could reasonably reject any principle permitting the act. I noted in section 2 that, on the face of it, a better explanation of why such an action is wrong has to do with the action itself, not the relation between the action and what a representative for a nonrational being could not agree to in a social contract. Now I can flesh out the alternative explanation in more detail. An action treats another unreasonably insofar as it fails to treat a highly valuable being as having at least the intrinsic value that it does. Supposing that nonhuman animals and nonrational humans have a relatively high intrinsic worth, an agent who tortures them merely for the fun of it treats them as having value only insofar as they serve the interests

or ends of the agent. That is, in negatively responding to valuable beings, he treats them as merely having extrinsic value. I propose that this is a more revealing explanation of why the act is wrong than the bare idea that a nonrational being's trustee could reasonably reject a principle licensing this act.

Indeed, one major thing going for the theory of wrongness qua unreasonable action is that, compared to prominent moral theories, it best explains duties to nonrational beings. Not only does it account for direct duties to nonrational beings, but it also does not place rational and nonrational beings on an equal footing. Egoist, self-ownership, Kantian, and contractualist views cannot fully account for our duties to nonrational beings, necessarily offering indirect and implausible accounts of why it is wrong to mistreat them. And utilitarian and respect for life theories cannot fully account for the greater importance of rational beings, necessarily offering indirect and implausible accounts of why rational beings should take precedence in cases of serious conflict. The theory that one must treat beings as having at least the intrinsic value they do, where rational beings and (some) nonrational beings have different degrees of intrinsic value, promises to do the trick.

Now consider the way Scanlon accounts for degrees of wrongness. As we saw in section 3, Scanlon must say that an action is wrong in proportion to the degree of reasonable grounds for rejecting any principle that would permit the act. Murder is more wrong than breaking a promise, in that while both acts are permitted by principles that could be reasonably rejected, the former act is allowed by rules that are more reasonably rejectable than rules allowing the latter act. In contrast, the theory of unreasonable action says that murder is a more unreasonable act than breaking a promise. While Scanlon is silent about what generally makes one principle more reasonably rejectable than another, the theory developed in the previous section specifies what makes one action more unreasonable than another. Degrees of unreasonable action are degrees to which an action fails to treat a being as having at least the intrinsic value it does. Killing for money treats someone more unreasonably than breaking a promise in order to play pinball, since the former is a more intense negative response to a superlative value than the latter. That is, murder undermines the capacity for rational agency to a greater degree than breaking a promise.

I submit that the theory of wrongness qua unreasonable action accounts for wronging the nonrational and for degrees of wrongness better than Scanlon's contractualist theory. Another respect in which unreasonable action has greater explanatory force than reasonable rejection is that the former grounds the latter. That is, I now want to argue that to the extent that reasonable rejection has *any* explanatory power, it is because reasonable rejection tracks the property of unreasonable action.

* * *

The fact that an act treats others unreasonably explains well why ideal contractors could reasonably reject any principles permitting the act. After all, any time a contractor could reasonably reject a principle, it is natural to ask, "Why?" When a contractor rejects a principle permitting an act, he does so just insofar as it permits the act (p. 195). It is fair to ask, then, "What is it about the act in virtue of which a principle permitting the act is susceptible of reasonable rejection?" Features of the act outside the contract are surely what ground any judgments inside the contract about these acts. And, I will now suggest, the theory of unreasonable action does a very good job of capturing the salient features of those acts that Scanlon believes would be disallowed by principles that could not be reasonably rejected.

Consider some central actions forbidden by principles that could not be reasonably rejected: not according a guaranteed minimum of water to others if one has acquired (without force or fraud) control over a territory's sole source of water (p. 192); letting a person suffer serious harm so that millions of people can watch the World Cup match without interruption (p. 235); punishing a person who could not regulate her conduct according to legal standards (p. 264); leading a person to expect that one will perform an act that one does not in fact intend to perform, in order to get him to do something one wants (p. 298).

Note that egoists and utilitarians would not consider some of these actions necessarily to be wrong, that is, Scanlon recognizes the existence of restrictions on interfering with others. Note, too, that self-ownership theorists would not consider some of these actions necessarily to be wrong; Scanlon believes it can be wrong not to aid others, even when one neither has created expectations that would restrict one to aid nor has the prospect of net long-term benefit from aiding. In short, the output of Scanlon's theory fundamentally includes both deontological constraints and positive duties to help others.

The actions required by the theory of unreasonable action have the same structure. Constraints arise from the idea that it is unreasonable to treat an intrinsically valuable being as having no intrinsic value. Responding negatively to a being in order to prevent more of the same negative responses to others would treat the being merely as a means. And positive duties to others spring from the claim that it is unreasonable not to respond positively to beings with a superlative inherent worth. In fact, representative instances of the acts that Scanlon believes could be reasonably rejected by his parties are cases in which beings with the capacity for rational decision-making are either negatively responded to or do not receive the positive responses that are proportionate to their high value. Kantian arguments about respecting the capacity for choice, with which I assume

the reader is familiar, seem to do the work.

In sum, I have noted that features of actions themselves must explain why contractors can reasonably reject principles that permit them. And I have suggested that the theory of unreasonable action captures the features common to those actions forbidden by principles that Scanlon believes his contractors could not reasonably reject. It is therefore plausible to think that the property of unreasonable action grounds judgments of when a principle permitting an act can be reasonably rejected.

* * *

What prevents Scanlon from adopting the sort of theory I have sketched here? Why does he conceive of wrongness in terms of acts that could not be the object of a reasonable agreement rather than acts that treat others unreasonably? Why think, with Scanlon, that wrongness is fundamentally a function of hypothetical agreement? The text suggests three major arguments for retaining contractualism.

First, Scanlon doubts that there is any unifying element to acts in virtue of which principles allowing them could be reasonably rejected (pp. 218, 246). If that were true, then contractualism would probably be the most unifying element available.

However, I have worked to show that there are good reasons to be optimistic about finding unity at the level of actions. The theory of unreasonable action provides an explicit (albeit still somewhat imprecise) account of unreasonableness so far as it is a property of actions, one that has a deep parallel with the unreasonableness of emotions (and of additional judgment-sensitive attitudes, as I discuss below).

A second reason Scanlon is inclined to conceive of wrongness in fundamentally contractualist terms is that he believes that hypothetical agreement best captures intuitions about moral motivation. What is it that leads people to think that duties to others typically provide overriding reason for action? Scanlon seeks an answer to this question that (1) is phenomenologically accurate and (2) appeals to something intimately connected to the content of morality that can be recognized as important from outside a moral perspective. Here is Scanlon's answer:

When I reflect on the reason that the wrongness of an action seems to supply not to do it, the best description of this reason I can come up with has to do with the relation to others that such acts would put me in: the sense that others could reasonably object to what I do ... (155)

We have strong reason not to perform wrong acts purportedly because of an ideal of acting only on principles that are justifiable to others.

Now, there are respects in which justifiability to others does not track our experience of morality as having strong normative force. For instance, the reason one senses that one has for not treating certain nonrational be-

ings like soccer balls involves no thought about principles that a trustee for them could reasonably object to. I, for one, instead sense that such behavior treats another being unreasonably, viz., responds disproportionately to its intrinsic value. Cats and severely retarded humans do not warrant the negative treatment of getting kicked around; soccer balls do. Similar remarks apply to persons; the reason it seems so important to help other rational beings, when this can be done at little cost to oneself, is that something with the highest value in the world calls for going out of one's way to protect it. And this sort of account of moral motivation, I submit, satisfies the requirement that it be closely related to morality yet be capable of appreciation from a nonmoral point of view.¹⁸

The third major reason why Scanlon might prefer reasonable rejection to unreasonable action is that the latter carries some metaphysical baggage that he considers too heavy. Scanlon has some metaethical theses about intrinsic value that probably inform his normative theory of wrongness. The theory of unreasonable action conceives of wrongness in terms of responses that are disproportionate to intrinsic value, where intrinsic value is constituted by certain objects in the natural world, for example, rational agency. Now, Scanlon denies that goodness is identical to any set of natu-

¹⁸Scanlon critically discusses a related account of moral motivation, contending that it is best understood in contractualist terms. Specifically, he claims that what it is to treat rational life with respect must be grasped in terms of action conforming to principles that spring from the use of our rational powers, viz., principles that could not be reasonably rejected by those seeking reasonable agreement (pp. 8, 13, 103-7, 159, 168-71, 183, 268, 271-72, 354).

This argument is much like one that John Rawls once made in *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971). Rawls used to hold that it is hard to see what respecting persons could involve apart from acting in accordance with principles that emerge from the original position (pp. 179-80, 586). However, Rawls soon reversed his view of the relationship between Kantian respect and the original position. Rawls's considered view is that his contract is best understood as following from a more fundamental principle of equal respect. See, e.g., John Rawls, "A Well-Ordered Society," in Peter Laslett and James Fishkin (eds.), *Philosophy, Politics, and Society*, Fifth Series (New Haven: Yale University Press, 1979), chap. 1. The reason for the reversal is that the veil of ignorance and the parties' aims of ensuring an adequate package of primary goods need defense from the objections of egoists, utilitarians, and self-ownership theorists. The principle of respect not only can be understood independent of the original position, but also figures into the best justification for it.

Similar remarks apply to Scanlon's contract. The theory of unreasonable action is obviously akin to the principle of respect for persons. As I see it, the principle of respect for persons is an instance of the more general theory of unreasonable action, which regulates the treatment of more beings than just persons. (And the theory of unreasonable action, in turn, can be derived from the genus of unreasonableness, the fundamental norm regulating more judgment-sensitive attitudes than just intentional actions.) I have worked to show not only that actions can be deemed unreasonable independent of a contractual framework, but also that the unreasonableness of actions explains why principles permitting certain acts can be reasonably rejected. Hence, if one is a fan of respect for persons, one has no reason to construe wrongness fundamentally in terms of hypothetical agreement.

ral properties (pp. 96-97). Since the theory of unreasonable action is naturally interpreted to imply value naturalism, and since Scanlon rejects value naturalism, he would reject the theory of unreasonable action.

Scanlon denies value naturalism because of Moore's open question argument. In a nutshell, this argument says that two terms denote the same property if and only if they connote the same idea, that "goodness" does not connote the same idea as any term denoting a natural property, and, therefore, that "goodness" does not denote any natural property. But the open question argument rests upon a descriptivist semantic theory that has been severely criticized. Many have argued that two terms can have the same referent without having the same sense. For a common example, "water" does not connote the same idea as " H_2O ," yet these words denote the same thing. Similarly, even though an evaluative term such as "good" may not have the same sense as any term picking out a natural property (e.g., "rational being"), these terms could co-refer. Since the open question argument relies on a questionable semantics, Scanlon's reason for avoiding value naturalism is weak.

Scanlon accepts that natural properties (alone) provide reasons for judgment-sensitive attitudes (pp. 97-98). When a contractor reasonably rejects a principle permitting the act, he does so in light of the natural properties of the act. But if what we call "basic reasons" are certain features of the natural world, why deny that what we call "intrinsic values" are also aspects of nature? When the theory of unreasonable action instructs us (among other things) not to treat the most valuable being as having no intrinsic value, this means that we are not to respond negatively to a rational being. There is nothing more metaphysically suspect here than a contractor having a good reason to reject a certain principle because it permits the coercion or deception of an agent.

* * *

There is one more consideration I would like to offer on behalf of the theory of wrongness qua unreasonable action, namely, that it coheres with other parts of morality in ways that Scanlon's theory does not. There are at least two other moral domains that one can naturally consider to be functions of unreasonableness but that one cannot plausibly conceive in terms of a hypothetical social contract.

First, the idea of unreasonable action could easily be broadened to include not just the idea of treating others unreasonably but also the notion of treating oneself unreasonably. That is, the theory articulated here could ground an attractive account of self-regarding duties. Just as one should not fail to treat other people as having less than the superior intrinsic value they do, so should one not fail to treat oneself as having any less worth than one has. Familiar Kantian notions about there being moral require-

ments to develop (respond positively to) and not to impair (not respond negatively to) one's rational faculties fit nicely with the theory that doing oneself wrong is a species of unreasonableness.

Second, many of the moral virtues seem to be judgment-sensitive attitudes subject to a norm of unreasonableness. Consider some commonsensical views about motivational virtue and vice: a virtuous agent does not want to do what is wrong and does right acts for their own sake, whereas a base agent wants to do what is wrong and has an ulterior motive when he acts rightly. These intuitions suggest the following rough principle: motivational vice consists of disproportionality between a person's conative powers and the un/reasonableness of actions. There are positive and negative conative states, with a preference that a state of affairs obtain being positive and a preference that it not obtain being negative. These conative states also come in degrees, for example, a desire or motivating reason for action is strong when it is higher-order and intrinsic. And there are of course grades of unreasonable action, too. Hence, at least some motivational vice is plausibly conation that responds in an insufficiently negative way to the degree of unreasonable action and in an insufficiently positive way to the degree of reasonable action.

In addition to motivational virtue and vice, at least some emotional virtue and vice is naturally construed in terms of disproportionate reactions to un/reasonable actions. Certain emotions are commonsensically evaluated in terms of their proportionality to the wrongness of the actions to which they are consequent. For instance, the more unreasonable the action one has performed, the more intense one's guilt ought to be.

Space unfortunately precludes a more sophisticated development of these claims, but the basic ideas should be clear. It appears that an attractive theory of self-regarding duties could be cashed out in terms of a kind of proportionality between certain actions and one's own valuable nature. And given a first-order account of dutiful conduct in terms of avoiding unreasonable action, it would be promising to construct an account of moral virtues in terms of avoiding a second-order form of unreasonableness, that is, avoiding disproportionate emotional or motivational reactions to the un/reasonableness of actions. In contrast, it is not at all clear that self-regarding duties or moral virtues can be plausibly viewed as the outcome of an ideal agreement among a group of agents. Since these aspects of morality do not fundamentally concern relationships with others, the idea of one's behavior being justifiable to other agents is quite out of place. Scanlon notes that his contractualist theory of wrongness entails that what falls under the heading of "moral" is quite diverse and does not admit of any unifying principle;¹⁹ the contractualist theory of wrongness is weak-

¹⁹As Scanlon says. "It would be a mark against views like contractualism, then, if it should turn out that they are unable to account for the degree of connection that there seems,

ened to the extent that we find a principle that does unify other-regarding duties, self-regarding duties, motivational virtues, and emotional virtues.

6. Conclusion

This essay began by noting that the theory developed in Scanlon's *What We Owe to Each Other* differs from the views of others who also conceive of wrongness in terms of the unreasonable or a social contract. Instead of deeming acts themselves to be unreasonable, Scanlon thinks of unreasonableness as a feature of the rejection or acceptance of principles governing acts. And rather than deeming hypothetical agreement merely to represent that which makes an act wrong, Scanlon believes that hypothetical agreement constitutes it. I have worked to show that these elements of Scanlon's theory should be rejected; for there is a related theory that has more explanatory force, namely, the view that wrongness is action that treats others unreasonably. Drawing on intuitive ideas about the ways in which emotions are unreasonable, I have articulated an account of unreasonable action in terms of behavior that, roughly, is disproportionate to a being's high intrinsic value. I have argued that this theory of unreasonable action explains wronging the nonrational and degrees of wrongness better than Scanlon's theory, provided reasons for thinking that reasonable rejection has explanatory power only in virtue of tracking the property of unreasonable action, and suggested that a general norm of unreasonableness plausibly grounds additional moral domains such as duties to oneself and character virtues. I have also responded to the considerations that I suspect would prevent Scanlon from adopting my not-too-unfriendly amendment to his view.

What makes *What We Owe to Each Other* an outstanding book on moral philosophy is not, in my opinion, the general contractualist form of Scanlon's theory, but rather, for one, the specific analyses of actions disallowed by principles that cannot be reasonably rejected. Scanlon's discussions of rules that govern, for example, lying, punishing, rescuing, and promise-breaking are enormously revealing. They are the rich kernel of truth in a contractualist shell that I believe can be removed without untoward results. I view these discussions as articulating instances of unreasonable action, but one can appreciate their force even if one does not deem them to be unified by any principle whatsoever.

Several critical questions can be posed regarding the theory of unreasonable action, as it stands: how might the theory of unreasonable action accommodate the existence of pro tanto duties? to what extent could it

intuitively, to be between the reasons backing our ideas of right and wrong and those behind other moral values." See T.M. Scanlon, "Moral Theory: Understanding and Disagreement," *Philosophy and Phenomenological Research* 55 (1995): 343-56, p. 349.

provide guidance about what to do when pro tanto duties conflict? can the theory account for the idea that there are unchosen obligations to particular others, say, duties to aid one's fellow citizens or one's siblings? is there a way for it to explain the wrongness not just of injury but also of insult? should the theory of unreasonable action include the idea of disproportionate responses not just to entities and actions, but also to relationships or interests? can the theory avoid the objections that Francis Hutcheson and David Hume make to Clarke and Balguy?²⁰ If I am correct that the theory of unreasonable action, as developed so far, has strong explanatory force, then answers to these questions about the reasonable and the moral should be sought in future work.²¹

Thaddeus Metz

Department of Philosophy
University of Missouri, St. Louis
thadmetz@umsl.edu

²⁰See Francis Hutcheson, *Illustrations Upon the Moral Sense*, repr. in Raphael, *British Moralists*, Vol. 1, pp. 305-15; David Hume, *A Treatise of Human Nature*, repr. in Raphael, *British Moralists*, Vol. 2, pp. 8-19.

²¹For comments that have helped me to sharpen and clarify my ideas, I thank Adila Hassim, Jon McGinnis, Michael Pendlebury, David Phillips, Piers Rawling, Jim Stone, and Eric Wiland.

