https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

"BUT IT WOULD BE WRONG"*

By Stephen Darwall

I. Introduction

At a critical juncture in the Senate Watergate hearings in 1973, H. R. Haldeman testified about a taped conversation he had heard between President Richard Nixon and the White House lawyer, John Dean. Nixon and Dean were discussing the possibility of raising hush money to meet the demands of one of the Watergate burglars, E. Howard Hunt, so that Hunt wouldn't spill the beans about the break-in at the Democratic National Committee headquarters. The president asked how much money would be required, and Dean answered, "Probably a million dollars—but the problem is that it is hard to raise." Nixon replied, "There is no problem in raising a million dollars, we can do that." Then, according to Haldeman, Nixon added the following, soon to be famous, phrase: "But it would be wrong."[1] We now know that Nixon did not actually say this, but it seems clear what Haldeman's intention was in saying he did.[2] Haldeman wished to portray Nixon as implying that if paying off Hunt would be wrong, then there was good reason, decisive reason in fact, not to pursue that course of action, and that it should therefore not be considered further.

Here arises a recently discussed philosophical problem.[3] Does the fact that an act would be wrong provide any reason *itself*, let alone a decisive

* For helpful comments, I am very much indebted to the other contributors to this volume, and to audiences at Brown University, the Western Canadian Philosophical Association meetings at the University of Alberta, the Scots Philosophical Club meetings at the University of Stirling, the University of Maryland, the University of Leiden, Washington University, the University of Oslo, Bielefeld University, the Moral Philosophy Working Group at Yale University, and the Middle Atlantic Reading Group in Ethics. I am also indebted to Ellen Frankel Paul for editorial advice.

[1] "Counterattack and Counterpoint," *Time* (August 13, 1973).

[2] Here is what the transcript of the March 21, 1973, tape says: "PRESIDENT: How much money do you need? DEAN: I would say these people are going to cost, uh, a million dollars over the next, uh,—two years. (Pause) PRESIDENT: We could get that. DEAN: Uh, huh. PRESIDENT: You, on the money, if you need the money, I mean, uh, you could get the money. Let's say— DEAN: Well, I think that we're going— PRESIDENT: What I mean is, you could, you could get a million dollars. And you could get it in cash. I, I know where it could be gotten. DEAN: Uh, huh." (Watergate Trial Conversations, Nixon Presidential Library and Museum, http://nixon.archives.gov/forresearchers/find/tapes/watergate/trial/transcripts.php.)

[3] See T. M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998), 11; Jonathan Dancy, "Should We Pass the Buck?" in *Philosophy, the Good, the True, and the Beautiful*, ed. Anthony O'Hear, supp. vol. 47 (Cambridge: Cambridge University Press, 2000): 159–73; Philip Stratton-Lake, "Scanlon's Contractualism and the Redundancy Objection," *Analysis* 63, no. 1 (2003): 70–76; Michael J. Zimmerman, "The Good and the Right," *Utilitas* 19, no. 3 (2007): 326–53; and T. M. Scanlon, "Wrongness and Reasons: A

doi:10.1017/S0265052509990185
© 2010 Social Philosophy & Policy Foundation. Printed in the USA.

reason, not to perform it? The question is not whether there is adequate reason to avoid wrongdoing. Recent discussion has tended to take it for granted that there is. The question, rather, is whether the fact that an action would be wrong is itself a reason of any kind, whether especially strong or conclusive or of any strength whatsoever. Various considerations have led some philosophers to think that it is not.[4] Briefly, these considerations include the following.

An action is wrong only if there are reasons that make it wrong—*wrong-making features* of performing the action or "*grounds*" of its wrongness.[5] It seems, then, that it should be these features of the action, the features that make it wrong, that are the reasons not to perform the wrongful act, not the fact that the action would be wrong itself. Such a further reason would seem superfluous, and might indeed be thought to involve a kind of double counting.

Moreover, it is sometimes pointed out that a morally good person is moved directly by wrong-making features of the wrongful act, rather than by the fact that the act would be wrong; the former and not the latter are *the good person's reasons* for acting as she does.[6] Someone who is moved primarily by overall moral verdicts, such as an action's being wrong, rather than by the grounds of the verdict, has been said to have a kind of "moral fetish," rather than being truly morally good.[7] Thus, even Haldeman's fantasy endowed Nixon with something less than moral goodness. It was as if the fact, say, that paying Hunt off would subvert justice left Nixon entirely unmoved until he realized that that made the act morally wrong.

Of course, from the fact that a morally good person is typically moved to act by the morally relevant features of an action, rather than by the moral status these features bestow upon an action, it hardly follows that

Re-examination," *Oxford Studies in Metaethics,* vol. 2 (Oxford: Oxford University Press, 2007), 5–20.

[4] These include Dancy, Stratton-Lake, and Zimmerman (see previous note).

[5] The latter term derives from Dancy, "Should We Pass the Buck?" When I speak of "good-making" or "right-" or "wrong-making" properties in what follows, I will have this sense in mind, namely, *grounds* or *normative reasons* for something's value, rightness, or wrongness. We might call this a *normative* sense, as opposed to a *metaphysical* sense of good-making or right- or wrong-making features (i.e., the features in which something's being good, right, or wrong consists metaphysically). I shall be arguing that an act's wrong-making features in the normative sense (i.e., the grounds of or normative reasons for something's being wrong) do not exhaust the reasons not to perform the wrongful act—that the fact that the act is wrong is itself a reason not to perform it. I should not be understood as claiming that if there is some complex fact or facts in which this latter fact consists metaphysically, the fact that an act is wrong is a reason that is additional to *these* facts, since, on the hypothesis in question, that is what the fact of an act's being wrong would itself consist in. I have been helped here by discussion with, among others, Janice Dowell and Peter Schulte.

[6] Cf. Michael Smith on the *de dicto* desire to avoid wrongdoing as opposed to *de re* desires to avoid actions of wrong-making kinds: Michael Smith, *The Moral Problem* (Oxford: Blackwell, 1994), 75–76.

[7] Ibid.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

moral wrongness is not itself a reason for acting. Even if it is not a reason for a morally good person, it might still be a reason for those who are no better than morally mediocre. Or it might be a reason even for the morally good, though not a reason they typically act on.

Yet a further line of thought in favor of holding that moral wrongness is not itself a reason comes by analogy with T. M. Scanlon's "buck-passing" theory of value.[8] According to Scanlon, being good or having value is not itself a reason to have any attitude or take any action with respect to the valuable thing. Rather, value "passes the buck" to the good-making or value-making features of the valuable thing, and it is these features that provide reasons for the relevant (valuing) attitudes and actions toward it.[9] That the thing is valuable or good does not provide a further reason for the relevant attitudes or actions. If a buck-passing account is plausible for evaluative (good-making) properties, however, then why would it not be plausible for deontic (right- or wrong-making) properties also? Why not think that, as with goodness or value, moral wrongness passes the buck to wrong-making features, and it is (only) those features that provide reasons for the relevant attitudes and actions, including not performing the wrongful act?[10]

This has not, however, been Scanlon's view of wrongness. He has taken the position that being wrong is indeed a reason not to do what is wrong.[11] In what follows, I shall argue that Scanlon is right; taking a buck-passing approach to normative concepts in general, and to the concepts of moral obligation, right, and wrong, in particular, does not entail that an act's wrongness is not a reason not to perform it. To the contrary, I shall argue, the most plausible buck-passing theory of the concepts of moral obligation and wrong entails that an act's being wrong is a reason not to perform the act that is additional to the act's wrong-making features (in the normative sense of features that *ground* or are normative reasons for its wrongness).[12] I will argue this in three stages. First, in Section II, I will argue that although a buck-passing theory of a normative concept (or similarly, a "warranted-attitude" or "fitting-attitude" theory of the concept) entails that the fact that something instantiates the concept (say, is good, right, or wrong) is not a further reason for the specific attitudes that are conceptually tied to the normative concept, the theory may not entail that that fact is not a reason for some relevant choice, intention, or *action*, since the conceptually implicated attitude may not entail any relevant

[8] I take this to be a more pointed version of the first line of thought mentioned above.
[9] Scanlon, *What We Owe to Each Other,* 95–100.
[10] That is, why not think that moral wrongness passes the buck to features that are wrong-making in the sense of providing normative grounds of an act's wrongness? This differs from any features that might be wrong-making in the sense of being what an action's being wrong consists in metaphysically.
[11] See, e.g., ibid., 11; and T. M. Scanlon, "Wrongness and Reasons: A Re-examination," *Oxford Studies in Metaethics,* vol. 2 (Oxford: Oxford University Press, 2007), 5–20.
[12] See note 5 above.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

action attitude (like intention or choice). Second, in Section III, I will present a buck-passing theory of the concepts of moral obligation, right, and wrong, according to which the fact that something instantiates one of these concepts is normative, at least in the first instance, not for intention or action, but for attitudes through which we hold people responsible for their actions and blame them (what P. F. Strawson called "reactive attitudes").[13] Third, in Section IV, I will argue that this theory entails that the fact that an act is wrong is indeed a further reason not to perform the action, and that the theory tells us what this further reason consists in. It consists in the fact that refraining from the act is something we legitimately *demand* of ourselves and of one another. Finally, in Section V, I will discuss some light the theory sheds on whether there is (always) overriding or conclusive reason to act as we are morally obligated to act (all things considered) and not do moral wrong.

II. Normative Concepts, Attitudes, and Action

Although it is possible to hold a buck-passing theory of value on other grounds, one natural line of thought that leads to it is a *warranted-attitude theory of normative concepts* in general.[14] It is plausible to suppose that what normative concepts all have in common is—as Allan Gibbard says, using a phrase of Wilfrid Sellars's—that they are "fraught with ought."[15] The basic idea derives from A. C. Ewing and ultimately from Henry Sidgwick.[16] Sidgwick said that all "ethical judgments" include "the fundamental notion represented by the word 'ought'."[17] "Ought" here has a "flavorless" sense, to use Gibbard's term, of thin justification or warrant—distinct, for example, from more thickly qualified senses like "morally ought" or "prudentially ought."[18] According to warranted-attitude theories, every normative concept is tied conceptually to some specific atti-

[13] P. F. Strawson, "Freedom and Resentment," in Strawson, *Studies in the Philosophy of Thought and Action* (London: Oxford University Press, 1968). I defend this analysis in greater detail in *The Second-Person Standpoint: Morality, Respect, and Accountability* (Cambridge, MA: Harvard University Press, 2006).

[14] In my view, the warranted-attitude theory (or, more specifically, the fitting-attitude theory, which I shall discuss presently) is the most plausible approach to analyzing normative concepts. It is not necessary to the argument of this essay, however, that it actually succeeds. We shall simply assume that it does succeed as the most plausible line of thought leading to the view that wrongness is not itself a reason for acting. My point will be that accepting a fitting-attitude theory of wrongness would not commit us to this. To the contrary, on the fitting-attitude theory of wrongness that I shall propose, wrongness is indeed itself a practical reason.

[15] Allan Gibbard, "Knowing What to Do, Seeing What to Do," in *Ethical Intuitionism: Re-evaluations,* ed. Philip Stratton-Lake (Oxford: Oxford University Press, 2003), 212.

[16] A. C. Ewing, "Suggested Non-Naturalistic Analysis of Good," *Mind* 48 (1939): 1–22; Henry Sidgwick, *The Methods of Ethics,* 7th ed. (London: Macmillan, 1967).

[17] Sidgwick, *The Methods of Ethics,* 25.

[18] Allan Gibbard, *Wise Choices, Apt Feelings* (Cambridge, MA: Harvard University Press, 1990), 7. See also Owen McLeod, "Just Plain 'Ought'," *The Journal of Ethics* 5 (2001): 269–91.

tude or set of attitudes; each is the concept of being a warranted object of its distinctive attitude or attitudes.

For example, the concept of the *desirable* is that of being a justified object of desire, of being such as there is normative reason to desire. The concept of the *estimable* is that of what is warrantedly esteemed. The concept of *dignity* is that of what warrants respect. And so on. Every normative concept can be represented in terms of its distinctive attitude or attitudes and the fundamental normative notion of warrant, justification, ought, or (normative) reason.

There is, however, a crucial qualification. As recent discussion of the "wrong kind of reasons" problem makes clear, not just any reason for having some attitude is a reason of the "right kind" to justify or warrant an attitude in the terms that are distinctively relevant to the normative concept that is conceptually tied to it.[19] For example, instrumental or strategic reasons for having some attitude—say, a desire or an intention to drink a toxin if one will be rewarded for having either attitude, as in Gregory Kavka's famous "toxin puzzle"—are reasons of the wrong kind to establish that drinking the toxin would itself be either desirable or choiceworthy.[20] Reasons of the right kind are "fittingness" reasons; they show that the attitude "fits" or is a fitting response to the object.[21]

On what we might better call *fitting-attitude theories* of normative concepts, therefore, every normative concept can be analyzed in terms of "fittingness" (that is, support by warrant or reasons of the right kind) and the attitude or set of attitudes that is distinctive of the concept. Thus, the concept of the *desirable* is that of being a fitting object of desire; the *estimable,* that of being a fitting object of esteem; *dignity,* that of being a fitting object of respect; and so on for every normative concept.

Now for fitting-attitude theories to be plausible, there must be some way of distinguishing reasons of the right kind from those of the wrong kind that does not itself depend upon the relevant normative concept. If the only way of distinguishing the right kind from the wrong kind of reasons to desire, say, is that the former bear on an object's desirability, then any attempt to analyze the concept of desirability in terms of the right kind of reasons to desire will be circular.

In my view, various promising suggestions have been made. Wlodek Rabinowicz and Toni Ronnøw–Rasmussen have proposed that reasons of

[19] Justin D'Arms and Daniel Jacobson, "The Moralistic Fallacy: On the 'Appropriateness' of Emotions," *Philosophy and Phenomenological Research* 61 (2000): 65–90; Justin D'Arms and Daniel Jacobson, "Sentiment and Value," *Ethics* 110 (2000): 722–48; Wlodek Rabinowicz and Toni Ronnøw-Rasmussen, "The Strike of the Demon: On Fitting Pro-Attitudes and Value," *Ethics* 114 (2004): 391–423.

[20] The "toxin puzzle" refers to a situation Kavka described in which there is an instrumental reason to form an intention to drink a toxin (e.g., that one would be rewarded if one did), but in which this apparently gives one no reason to act on the intention. See Gregory S. Kavka, "The Toxin Puzzle," *Analysis* 43 (1983): 33–36.

[21] D'Arms and Jacobson, "Sentiment and Value."

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

the right kind appear also in the object or content of the attitude they purport to support. In other words, what the reason is strictly for is not just some specific attitude, but a specific-attitude-toward-something-on-account-of-the-reason.[22] Pamela Hieronymi has argued that every attitude has its distinctive deliberative question and that reasons of the right kind are those that bear on that question.[23] And I have proposed that the right kind of reasons for an attitude are those that are capable of being *the subject's reason* for the attitude, that is, those *on account of which* someone can directly form the attitude.[24] Thus, it is not psychically possible for *an agent's own reason* for intending or desiring intrinsically to drink a toxin to be that she will win a reward if she intends or forms an intrinsic desire to drink it. She can desire to have an intrinsic desire to drink the toxin for this reason, but she cannot desire intrinsically to drink the toxin for this reason. Likewise, if she were offered a reward were she to believe that the U.S. war in Iraq was an unqualified success in every respect, she could *desire,* for that reason, to believe that proposition, but that reason could not be her reason for having the belief itself. It would be psychically impossible to form the belief *for that reason.* One cannot reason: "I will get a reward if I believe *p;* therefore, *p,*" as one would have to do if one were to believe *p* for the reason that one will get a reward.

Perhaps none of these proposals is fully satisfactory, but we need not worry about that. Since the most powerful line of thought leading to buck-passing is a fitting-attitude approach to normative concepts of some kind, we can simply assume that a solution to the "wrong kind of reasons" problem can be found. What we want to know is whether *if* a fitting attitude account of wrongness were true, this would show that wrongness is not itself a reason for acting—that is, whether it would support being a buck-passer about wrongness with respect to reasons for acting.

This brings us to Scanlon's buck-passing theory of value. As I said, the "letter" of the buck-passing theory—that being good or valuable is not itself a reason to hold any relevant attitude toward the valuable thing or to promote it—does not require a warranted- or fitting-attitude theory of normative concepts. Some things Scanlon says, however, suggest such a theory. For example, he says that "to claim that something is valu*able* (or that it is 'of value') is to claim that [one and] others have reason to value it."[25] This suggests a warranted- or fitting-attitude account of the valuable (in terms of warrant for some valuing attitude or other). So under-

[22] For example, it might be a reason not just for a desire to eat an apple, but for a desire to eat an apple on account of its taste. (Rabinowicz and Ronnøw-Rasmussen, "The Strike of the Demon," 414.) This is related to a suggestion of W. D. Falk's in "Fact, Value, and Nonnatural Predication," in W. D. Falk, *Ought, Reasons, and Morality* (Ithaca, NY: Cornell University Press, 1986), 117.

[23] Pamela Hieronymi, "The Wrong Kind of Reason," *The Journal of Philosophy* 102, no. 9 (2005): 437–57.

[24] Darwall, *The Second-Person Standpoint,* 66.

[25] Scanlon, *What We Owe to Each Other,* 95.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

stood, value would be a generic evaluative concept, related conceptually to a set of valuing attitudes. Being desirable, estimable, having dignity, and so on, would then be distinct species of value.[26]

A buck-passing theory about value, at least with respect to the relevant valuing attitudes—that being valuable or good is not itself a reason to have the (valuing) attitude toward that thing—would then follow straightaway. If being valuable *just is* there being reason to have one or another valuing attitude toward something, and if how valuable that thing is depends on the weight and character of the reasons for valuing it, then the fact that something is valuable cannot be a further reason to value it. By creating a further reason, the exact value of something (say, *A*) would thereby make the thing more valuable (*A* + *B*, where *B* is the increased value owing to the additional reason created by the thing's having value *A*).[27] But that would seem to be absurd.[28]

Clearly, the reason why a fitting-attitude theory of value apparently entails buck-passing about value is that the attitudes that are implicated in the concept of value, and the attitudes with respect to which the question arises whether value does or does not provide a reason, *are the very same attitudes.* If the concept of the desirable is the concept of what there is reason (of the right kind) to desire, then the fact that something is desirable cannot provide a (further) reason to desire it. According to a fitting-attitude theory, something is exactly as desirable as the (right kind of) reasons for desiring make it. Its being desirable does not make it more desirable yet.

Does that mean, however, that being valuable cannot itself provide a reason for *action?* Presently, I will argue that there are kinds of value that do indeed provide reasons for action, namely, where the relevant valuing attitude does not itself conceptually implicate choice, action, or some other agential disposition. With respect to the desirable, however, there does seem to be a fairly tight conceptual connection. Some disposition to bring the desired state about seems to be part of the very concept of desire. But if that is so, then a buck-passing theory of the desirable would arguably entail not just buck-passing about reasons for desire, but also about reasons to bring about desire's object.

[26] For a defense of such a view, see Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge, MA: Harvard University Press, 1990).

[27] See Zimmerman, "The Good and the Right," for a parallel argument against holding that wrongness creates an additional reason.

[28] For our purposes, we do not strictly need to suppose that a fitting-attitude theory actually does entail buck-passing about value. The argument of this essay is that even if it were to support buck-passing about the relevant valuing attitudes in this way, a fitting-attitude account of wrongness of the sort I will propose would not entail buck-passing about wrongness with respect to reasons for action. Here again, then, we can just assume that a fitting-attitude theory of value of the sort Scanlon gestures to entails buck-passing about value with respect to the relevant valuing attitudes. I am indebted here to discussion with Ruth Chang.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

Whether a fitting-attitude theory of moral wrongness will entail buck-passing about reasons not to perform wrongful acts, therefore, will depend on precisely which attitude or attitudes are conceptually implicated in the concept of moral wrong. We must therefore ask: If the concept of moral wrong is a normative concept, what attitude or attitudes is it normative for?

The proposal whose implications we shall consider is one I have defended at greater length elsewhere, namely, that the concepts of moral obligation and moral wrong are irreducibly second-personal concepts, since they conceptually implicate attitudes through which we implicitly address agents and hold them accountable.[29] What is new in the present essay are the implications I draw from this analysis of moral obligation for the question of whether an action's wrongness is in itself a reason not to perform the action. The proposal I discuss is a version of John Stuart Mill's idea, taken up more recently by P. F. Strawson, Richard Brandt, Allan Gibbard, and others, that the concept of moral wrong is the concept of what would be *blameworthy*—that is, a warranted object of Strawsonian reactive attitudes including blame—if an agent were to perform the act without some adequate excuse.[30] I follow what I take to be the thrust of Strawson's point by analyzing reactive attitudes as invariably involving *demands* that are implicitly *addressed* to their attitudes' objects second-personally, if only in imagination. Mill said that "we do not call anything wrong, unless we mean to imply that a person ought to be punished in some way or other for doing it; if not by law, by the opinion of his fellow creatures; if not by opinion, by the reproaches of his own conscience."[31] As I interpret Mill through Strawsonian lenses, "punishment" in this passage should be taken in the broad sense of holding someone responsible with a reactive attitude like indignation, blame, or guilt.[32] *Holding* someone accountable, in this way, is implicitly a second-personal relation, since it involves the (imagined, at least) address of a demand to its object. The connection between the concepts of moral obligation and wrong (on

[29] See Darwall, *The Second-Person Standpoint*, 91–118. There I define a set of second-personal concepts, which conceptually involve the idea of claims and demands that can be *addressed* (to an addressee, second-personally), and I argue that these include the concepts of moral obligation, wrong, rights, the dignity of persons, and the very concept of a moral person (as a subject of obligations).

[30] Strawson, "Freedom and Resentment." Strawson uses the term "reactive attitude" to refer to a set of attitudes (which include indignation, resentment, guilt, and moral blame) that implicitly hold their objects responsible and thus regard them in a distinctively interpersonal (as I put it, "second-personal") way. For ease of expression, I will sometimes speak of actions, and not agents, as blameworthy; strictly speaking, however, it is the agent who is appropriately blamed for performing a "blameworthy action" (speaking loosely).

[31] John Stuart Mill, *Utilitarianism*, chap. V, sec. 14.

[32] I use "blame" throughout this essay to refer to an attitude rather than to any specific activity or speech act. It is thus possible to blame someone without ever saying anything to him or to anyone. Blame in this sense also differs from a belief or judgment that someone is blameworthy. One might say, for example: "I know that she still blames me for what I did twenty-five years ago, though we haven't talked in many years."

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

the one hand) and the reactive attitude of blame (on the other) is then this: What is morally wrong is what is blameworthy—that is, what is warrantedly blamed, if the action is done without adequate excuse.

If this is so, then the attitudes for which the concepts of moral obligation and moral wrong are conceptually normative are attitudes other than intention, choice, or any agential disposition like desire. The object of a reactive attitude is, in the first instance anyway, a person, and not some action or outcome as this might appear from an agent's deliberative standpoint. Consequently, if the concepts of moral obligation and moral wrong are normative for reactive attitudes, then the reasons for being a buck-passer about evaluative concepts like the desirable will not simply transfer directly to being a buck-passer about moral wrong with respect to reasons *for action,* specifically, for avoiding moral wrong. To be sure, if what is wrong is what is warrantedly blamed if done without adequate excuse, then, on the assumption that someone lacks such an excuse, the fact that he acted wrongly cannot provide a reason for holding him responsible through reactive attitudes and blaming him that is additional to those reasons provided by the act's wrong-making features.[33] On the proposal we shall consider, the latter reasons (the wrong-making features) consist in whatever reasons (of the right kind) there are for the reactive attitudes through which one holds someone responsible for performing an act of that kind. By analogy with our earlier point about the desirable, something does not become more blameworthy than its wrong-making features make it by virtue of being morally wrong; its being wrong just *is* its being blameworthy (if done without excuse) to whatever degree its wrong-making features make it. Anyone who accepts the fitting-attitude account of moral obligation and wrongness I will propose should therefore be a buck-passer with respect to reasons for blaming and holding morally responsible. But that would not entail being a buck-passer with respect to reasons for *action,* specifically, for avoiding moral wrong. The fact that an action is wrong might still itself be, and I shall argue actually is, a reason, indeed a decisive reason, not to perform the act or to intend or choose to do so.

Before I sketch this proposal in more detail, I would like to note briefly some other normative concepts where buck-passing about the relevant reasons for action is not very plausible. The explanation for this is that the attitudes that are conceptually implicated in the normative concept are attitudes other than intention and choice. (I will be arguing that this is also the case with moral wrong.)

Consider, for example, the concepts of the estimable and its contrary, the contemptible. To be estimable is to be worthy of esteem, and if some-

[33] That is, it cannot provide such a reason on the assumption we have been making, namely, that a fitting-attitude theory of a normative concept entails buck-passing with respect to reasons for the attitudes that the concept conceptually involves.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

thing is estimable, that will be on account of its estimable-making features.[34] These features are all, and only, whatever reasons of the right kind there are for esteeming it. For now-familiar reasons, something's being estimable cannot be a further reason for esteeming it. And the same is true, mutatis mutandis, for the contemptible. The features that make something contemptible are all, and only, reasons of the right kind that exist for having contempt for it. But the rationale for being a buck-passer about reasons for esteeming (or having contempt) does not directly transfer to being a buck-passer about reasons for *acting* in some way or other with respect to the estimable or the contemptible. Even if esteeming some trait generally gives rise, say, to some desire to acquire the trait, the desire does not seem to be *conceptually* implicated in the former. We might imagine beings who admire some trait while lacking any desire whatsoever to act in any particular way with respect to it. Whatever relation exists between esteem and the desire to acquire or emulate seems different from the conceptual connection between a desire for some outcome and a disposition to bring that outcome about.

Thus, whatever rationale may exist for being a buck-passer about reasons for esteem and contempt, it could still be the case, and arguably is the case, that the fact that something is estimable is itself a reason to *act* in various ways regarding it—say, to promote the estimable trait or to realize it in one's own life. Thus, when Aristotle advances the view that virtuous activity (chosen for its own sake as noble or estimable) is the final good for us human beings and is therefore what we have most reason to seek, he is not simply mouthing an empty truth, such as that what we have reason to seek is what we have reason to seek (or an "almost empty" truth, such as that what we have reason to desire is what we have reason to bring about). He is making a substantive normative claim.

Or consider the concept of welfare or well-being. Most philosophers believe that this too is a normative concept. But if it is, the question arises, for what attitude or attitudes is it normative? It seems clear that, unlike the concept of the choiceworthy, the concept of well-being is not simply that of what a person herself has reason to choose, intend, or do, since whether or not we do, in fact, have reason to do anything other than promote our own welfare, it seems clear that no incoherence or conceptual confusion is involved in thinking we do (for example, in thinking that *others'* well-being, or moral considerations, or a whole host of other considerations give us reasons to act that are additional to those that concern our own well-being). But if that is so, then the concept of welfare cannot possibly be identical with the concept of what the person herself

[34] Again, such features are "estimable-making" in the sense of being normative *grounds* of something's being estimable, not in the sense of being facts in which the thing's being estimable consists metaphysically.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

has reason to do. Nor can welfare simply be the concept of the desirable, for similar reasons. We often think that things that have nothing to do with us or our own well-being are desirable—such as the survival of the planet long after we ourselves are dead, the flourishing of our children in the future, or whatever. And whether or not any of these thoughts are true, there is no doubt that they are coherent.

I have argued elsewhere that the most plausible account of well-being as a normative concept holds that it is what is desirable for someone *for that person's sake;* that is, well-being is what there is reason to desire for someone out of care or sympathetic concern *for her.*[35] A *rational care theory of welfare* understands the normativity of welfare to be for desires for someone (or for some being) for her (or its) sake, that is, from the perspective of care or concern for her (or it). In other words, this concept of welfare is the one we need and make use of when, in caring for someone, we face the normative question of what to want (what we should want) for her. In yet other words, a person's welfare is what there is reason to desire for someone, conditional on there being reason to care for her.

If some account in this general direction is correct, then the warranted attitude that is conceptually involved in welfare or well-being is actually quite complex and does not itself conceptually involve intention or choice. Granted, if one cares for someone and is thereby disposed to have desires (and thereby to act) for her sake, then any reasons why something would be good for her—welfare-making features—will automatically become reasons to bring that thing about. But so far as the concept of welfare goes, this does not entail that, if something, X, is a welfare-making feature for Y (i.e., X would contribute intrinsically to Y's welfare), then there is reason for anyone, whether Y or anyone else, to bring X about. It is conceptually coherent to hold that something would contribute intrinsically to someone's welfare (by realizing some welfare-making feature), but nonetheless hold that there is no reason whatsoever to bring that thing about. Y herself might coherently judge this. Y might not care for herself or think she is worth caring for, and might think therefore that considerations of her own welfare provide no reasons for acting, either for her or for anyone else. What is conceptually incoherent is caring for Y, or thinking one should care for her (for reasons of the right kind), and not taking welfare-making features as reasons to desire that these features be realized in Y's life for Y's sake.

I take it as evidence that something in this general neighborhood is correct that there is no attraction whatsoever to holding a buck-passing theory of welfare with respect to reasons for action. No philosopher I am aware of seems tempted to deny that the fact that something will promote an agent's welfare or well-being is itself a reason for her to realize or

[35] Stephen Darwall, *Welfare and Rational Care* (Princeton, NJ: Princeton University Press, 2002).

promote it. To the contrary, it is hard to think of a more widely held view in the history of ethics. If a fitting-attitude approach to welfare is correct, something's being for someone's welfare must consist in there being reasons for some attitude or other toward that thing. And this would naturally lead to a buck-passing view *for that attitude or attitudes.* But a rational care theory of welfare illustrates the possibility of a fitting-attitude theory of welfare that does not entail buck-passing for welfare with respect to reasons for action.

III. Moral Obligation, Moral Responsibility, and the Second-Person Standpoint

Even if a fitting-attitude theory of a normative concept entails buck-passing with respect to reasons for the attitudes that the concept analytically implicates, therefore, the theory may nonetheless not support buck-passing with respect to reasons for action. What, then, about the concept of moral wrong? As with other normative concepts, whether a fitting-attitude theory of wrongness supports buck-passing about reasons for acting will depend on the attitude or attitudes that are distinctive of the concept.

In *The Second-Person Standpoint,* I argue that the concepts of moral obligation and moral wrong are irreducibly second-personal in the sense that they conceptually involve *addressable* (and so second-personal) authoritative demands. When we violate moral obligations, we violate legitimate expectations and demands that we have and can make, in principle anyway, of one another and ourselves as representative persons or members of the moral community.[36] When we think that an action is wrong, as Richard Nixon might or might not have thought about subverting the Watergate investigation by paying hush money to Hunt, we are committed to thinking that not performing that action is something we therefore warrantedly demand of ourselves and one another. My claim here is that, so understood, the fact that an act is wrong gives us a reason not so to act that is additional to reasons consisting of the act's wrong-making features—and that this reason is a *second-personal reason* owing to its conceptual connection to (second-personally) addressable demands.

The thesis that the ideas of moral obligation and moral wrong conceptually involve the notion of (second-personally) addressable legitimate demands is rooted in two further ideas. The first is the one I mentioned before, namely, Mill's and Strawson's idea that moral obligation and wrong are conceptually related to warranted reactive attitudes. What is morally wrong is what one is warrantedly held accountable and blamed for (with

[36] By the "moral community," I mean no actual community, but a regulative ideal like Kant's "kingdom of ends." We could as well say that the authority is one we have *as representative persons.* I am indebted to David Velleman and Samuel Scheffler for discussion on this point.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

reactive attitudes) if one so acts without adequate excuse. The second idea is that holding someone responsible with a reactive attitude has an implicitly second-personal character (which is why Strawson contrasts its role in interpersonal relationships with taking an "objective" or third-personal view toward someone).[37] Holding someone responsible involves the putatively authoritative address of a demand, if only in imagination.

To prepare the way for considering this thesis and its implications, we need first to focus more precisely on the concepts of moral obligation and moral wrong in order to notice how they differ from other concepts with which they might be conflated. I take it as obvious, first, that "wrong" in the sense we are interested in refers to a narrower notion than the broader sense that is involved in just any kind of mistake, failure, or incorrectness. If, in doing a math problem, I make a mistake and get the wrong answer, I do no wrong in the sense with which we are concerned; I violate no moral obligation.

Neither is the idea of moral wrong in the current sense the same as that of acting against the balance of reasons, or even, indeed, against the balance of moral reasons. This is shown by the fact that supererogation, an act's being morally recommended but not morally obligatory or required, is a conceptually coherent possibility. That does not mean that theories that deny supererogation—like act-consequentialism, which holds that it is always wrong to do anything other than the morally best, optimific act—are mistaken. It just means that whether they are mistaken or not—whether there is such a thing as supererogation—is a substantive normative question that is not settled by the concept of moral obligation itself. And if that is so, moral obligation must be a different concept than that of being most morally choiceworthy.[38]

[37] See Strawson, "Freedom and Resentment," 81: "What I want to contrast is the attitude (or range of attitudes) of involvement or participation in a human relationship, on the one hand, and what might be called the objective attitude (or range of attitudes) to another human being, on the other. . . . To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided, though this gerundive is not peculiar to cases of objectivity of attitude. The objective attitude may be emotionally toned in many ways, but not in all ways: it may include repulsion or fear, it may include pity or even love, though not all kinds of love. But it cannot include the range of reactive feelings and attitudes which belong to involvement or participation with others in inter-personal human relationships; it cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other. If your attitude towards someone is wholly objective, then though you may fight him, you cannot quarrel with him, and though you may talk to him, even negotiate with him, you cannot reason with him. You can at most pretend to quarrel, or to reason, with him."

[38] But might the concepts of moral obligation and moral wrong be analyzed in terms of morally *conclusive* reasons—a moral obligation being something there is conclusive moral reason to do, and being wrong being something there is conclusive moral reason not to do? But what is it for reasons to be morally conclusive? On the one hand, if we give this an epistemic sense, such that reasons are conclusive if they conclusively establish that a certain

This should already make us suspicious of taking a buck-passing view of moral wrongness in relation to reasons for acting. If the concept of moral obligation were simply that of an action that is best supported by moral reasons, or best supported by reasons period, then being wrong could not be an additional moral reason, or reason period, not to perform the act. Being morally wrong would already involve being supported by all the reasons, or all the moral reasons, there are for acting. But that is not the way we normally think of it. Anyone who takes seriously the debate about whether views like act-consequentialism "demand too much," for example, must already take for granted that the ideas of moral obligation and moral wrong involve a notion of requirement or demand that differs conceptually from that of the balance of moral reasons.

But how are we to understand this notion? What is it for morality to demand something? Plainly, the idea of demand or requirement extends more widely than the moral. We speak, for example, of the requirements of reason: say, the demand not to have contradictory beliefs or plans. Unlike demands of reason or logic, however, talk of moral demands is conceptually linked to *accountability.*[39] What we are morally obligated to do is, as a conceptual matter, what we are morally answerable for doing.

There is thus an important difference between moral obligations and, for example, requirements that are imposed by logic. Moral obligation is tied to a distinctive kind of responsibility—accountability—*conceptually.* If I fail to act as I am morally required (without adequate excuse), it simply follows straightaway that reactive attitudes like blame and guilt are thereby warranted.[40] In contrast, responses like these seem appropriate to logical blunders only in certain contexts, and even here, what seems to be in question is a moral error of some kind (as when I have a special responsibility for reasoning properly).[41] Moreover, although a connection to accountability is intrinsic to the concept of moral obligation, it is obvi-

act is what morality most recommends, then this really adds nothing to the possibility just canvassed. On the other hand, if reasons are morally conclusive when they warrant a moral *requirement* or demand, then this possibility amounts to the one I go on to consider presently. In order for there to be some other possibility, we would have to have a notion of morally conclusive reason that we understand independently of the idea of moral requirement or demand, which we could then use to understand this latter idea. I doubt that this is so. I am indebted to Samuel Scheffler for discussion on these points.

[39] For this reason, I argue in *The Second-Person Standpoint* that the normativity of moral obligation is not adequately captured by the Kantian idea that moral demands are demands of reason. Of course, it could still be the case that what is morally demanded is also demanded by reason. I argue that the most promising line of argument supporting this, indeed, proceeds from a second-personal account of moral obligation's normativity.

[40] Indeed, the very idea of an "excuse" is not internal to the rules of logic; it must be understood in relation to a broader context that includes other norms.

[41] I do not mean, of course, that logical errors are not subject to criticism, or that we do not sometimes use words like "blame," as when a teacher says that he does not blame his student for a given error on a first try.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

ously no part whatsoever of the idea of a logical requirement or demand of reason.

Mill seems on safe ground, then, in saying that our concept of wrongdoing is essentially related to accountability. Even if it is natural to think that a person falls short of full virtue if she does only what can be required of her in the sense of what she can be warrantedly *held to* through reactive attitudes like blame and guilt, she nonetheless does no wrong. We do not impute wrongdoing unless we take ourselves to be in the range of the culpable, that is, unless the action is such that the agent is not just morally criticizable in some way or other, but aptly *blamed* or the object of some other form of accountability-seeking reactive attitude if she lacks an adequate excuse.

This aspect of the concept of moral wrong has been stressed by a number of contemporary writers. John Skorupski says that calling an act "morally wrong ... amounts to blaming the agent" and that the idea of moral wrong cannot be understood independently of that of blameworthiness.[42] Allan Gibbard quite explicitly follows Mill's lead in proposing that "what a person does is *morally wrong* if and only if it is rational for him to feel guilty for having done it, and for others to be angry at him for having done it."[43] And we can find versions of this Millian idea in other writers also.[44] It is consistent with these views, of course, that there remains a distinction between the wrongness of acts and the blameworthiness of agents. Someone is not to blame for wrongdoing if he has an excuse.

This gives us half of what underlies my thesis that the ideas of moral obligation and moral wrong conceptually involve (second-personally) addressable legitimate demands. The other half concerns the second-personal character of the distinctive attitudes—Strawsonian reactive attitudes—whose warrant is conceptually implicated in the idea of moral wrong. Reactive attitudes invariably involve a kind of "demand," as Strawson himself says, that is imaginatively addressed to its object and that must, therefore, presuppose the authority to address that object.[45]

Consider indignation, for example. Indignation is an instance of what Strawson calls an *impersonal* reactive attitude, since it is felt from an observer's perspective rather than from that of anyone involved in the situation to which one is responding. To feel indignation toward someone

[42] John Skorupski, *Ethical Explorations* (Oxford: Oxford University Press, 1999), 29, 142. This claim is, however, too strong, since one may not blame someone for acting wrongly if the person has some adequate excuse for doing so.

[43] Gibbard, *Wise Choices, Apt Feelings*, 42 (emphasis in the original).

[44] See, e.g., Kurt Baier, "Moral Obligation," *American Philosophical Quarterly* 3 (1966): 210–26; Richard Brandt, *A Theory of the Good and the Right* (Oxford: Oxford University Press, 1979); and Russ Shafer-Landau, *Moral Realism: A Defence* (New York: Oxford University Press, 2003).

[45] Strawson, "Freedom and Resentment," 92–93.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

is to feel that he is to blame for some conduct and is therefore appropriately held accountable for it, if only by being subject to reactive attitudes from himself (guilt) and others (blame or indignation). It is essential to Strawson's argument in his classic essay "Freedom and Resentment" that indignation differs from its seeming to one that some sanction would be desirable or even that the sanction would make for a more valuable or fitting whole (that it would realize something like poetic justice). The feeling of indignation invariably includes a sense of legitimate or authoritative demand that may be absent from the thought that some sanction would be either desirable or fitting. We feel indignation only when we feel we can reasonably *expect* or demand of others that they act in certain ways. Indeed, Strawson says, "the making of the demand *is* the proneness to such attitudes."[46]

Similar points hold for *personal* reactive attitudes, which are felt, unlike indignation or moral blame, not as if from the perspective of a disinterested third party, but as if from the victim's point of view or that of someone identified with her.[47] We resent what we take to be violations against ourselves or those with whom we identify. If you resent someone's treading on your foot, for example (or, even more, his rejecting your request or demand that he stop doing so), you feel as if he has violated your valid claim or demand and as if some claim-exacting or responsibility-seeking response by you, or on your behalf, is justified.

Strawson contrasts "the attitude (or range of attitudes) of involvement or participation in a human relationship" (taking a second-person stance, in my terms), on the one hand, with "the objective attitudes (or range of attitudes) to another human being, on the other."[48] We take an "objective" attitude toward those we see as unfit for "ordinary adult human relationships," such as very young children and those with "deep-rooted psychological abnormality," and we regard them as appropriately subject to "treatment" or "management," rather than to reactive attitudes and forms of interpersonal address that involve them.[49] What is distinctive about reactive attitudes, then—whether they are *impersonal* attitudes like indignation or blame, or *personal* attitudes like resentment—is that they are felt as if in second-personal relation with another: they involve an imagined *address* of a claim or demand *to* their object that holds their object answerable for compliance.[50]

[46] Ibid. (emphasis in the original).

[47] In addition to negative personal reactive attitudes, Strawson also mentions gratitude, which he regards as a (positive) personal reactive attitude—an attitude which, in this case, is felt as if from the perspective of a beneficiary.

[48] Ibid., 79.

[49] Ibid., 81.

[50] Thus, an attitude can be second-personal in the requisite sense and be third-party or "impersonal" in Strawson's sense. And the phenomenon of guilt shows that it is possible to take a second-personal attitude toward oneself. What makes an attitude second-personal is its having an implicit *addressee.*

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

IV. Moral Wrongness: A Second-Personal Reason to Act

I shall take it, therefore, that the fact that a certain action (say, stepping unbidden on someone's foot or subverting a criminal investigation) would be wrong amounts to the fact that it violates a legitimate demand we make of ourselves and one another as representative persons. My claim, then, is that this latter fact, the fact that an act violates such a demand, itself provides a reason not to perform the act that is additional to the act's wrong-making features. Because the notion of an *addressable* demand is itself part of this further reason, I call it a *second-personal reason.* Since it is impossible to make a demand of someone without addressing her in some way, the idea of second-personal address is built into the very ideas of moral obligation and moral wrong. Consequently, if the fact that an action would be wrong is a reason not to perform the act, it is a second-personal reason.

Wrong-making features, such as that an action would cause harm or subvert a criminal investigation, themselves entail nothing about legitimate demands and thus are not second-personal reasons.[51] My claim, however, is that in believing that such features *are indeed wrong making*, we are committed to thinking that these features nonetheless ground a legitimate demand not so to act and that this fact—the fact that the act would violate a legitimate moral demand and so be wrong—gives us a further reason not to perform the act.

One way to appreciate this is to consider the relation between the concepts of moral *rights* and moral wrong.[52] So far as I know, no one denies that the fact that an action would violate someone's rights is a reason not to perform the action. Since violating a right is itself a wrong-making feature, however, that may not seem to count against a buck-passing theory of moral wrong with respect to reasons for action. But there is an important difference between moral rights and most other wrong-making features. That violating someone's right is wrong (at least other things being equal) is not just a substantive normative truth; it is a conceptual truth.[53] As we shall see presently, something does not count as a violation of someone's right, unless it is something we legitimately demand that people not do to that person. So the fact that an act would violate someone's right is not itself independent of the fact that it would be wrong, at least other things being equal.

Suppose that I intentionally tromp on your foot to gain an advantage in getting to the table with designer underwear on sale when I could easily

[51] Again, by wrong-making features, I mean features that provide *grounds* or normative reasons for something's being wrong, not the features, if there are any, in which something's being wrong might consist metaphysically.

[52] For convenience, I shall shorten "moral right" to "right."

[53] I take it that it would be incoherent to assert that one has a right to something but that someone would do no wrong, not even other things being equal, if she were to deprive one of it.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

have gone around and risked doing without a pair of Calvin Kleins. Here I clearly violate your right. But in what does this fact consist? What is it to have a right? A right of this sort is a "claim right" in Wesley Hohfeld's sense.[54] I follow Joel Feinberg in holding that it is "claiming that gives [such] rights their distinctive moral significance."[55] Claiming is, in its nature, second-personal; we make claims on others by addressing them second-personally. "I'm sorry, but that's my foot you're stepping on," you might say to me as I'm on my way to the Calvin Kleins, implying that I am where I have no right to be and where you have every right to demand I vacate, apologize, and so on.

Such a demand would not be a *mere* demand, however: it would not simply be an imposition of your will on mine. It would be addressed as a legitimate demand, one you have the authority to make of me—an authority it would be unreasonable of me not to recognize. It seems obvious, moreover, that the authoritative demand in which your right consists gives me a reason not to step on your foot that is additional to the fact that I would be causing you pain, suffering, inconvenience, and so on. Before the concept of human rights became established, people recognized the latter reasons of pain, suffering, and so on, but not yet the former reason consisting in a right not to be gratuitously caused these things. However, when we accept that we have rights we can legitimately claim of one another (that we all are, in John Rawls's phrase, "self-originating source[s] of valid claims"), we recognize an additional reason not to cause pain and suffering: namely, that this is something we legitimately claim from one another as our right.[56]

As Hohfeld points out, the concept of a claim right is conceptually related to that of a moral duty or *obligation to the rightholder.*[57] Your right against me analytically entails that I have a moral duty or obligation *to you.* An obligation *to* someone is a "directed" or "bipolar" obligation.[58] If I violate such a duty to you, I *wrong you.* In so doing, I violate an authority *you* have, as an individual person, to demand that others not tread on *you.* Our moral rights are what we have the authority to demand of one another *as individuals.* You have a right against me only if you have (1) the individual authority to demand that I not act in some way toward *you* and (2) a standing others do not have to hold me answerable, as is shown

[54] Wesley Newcomb Hohfeld, *Fundamental Legal Conceptions,* ed. Walter Wheeler Cook (New Haven, CT: Yale University Press, 1923).

[55] Joel Feinberg, "The Nature and Value of Rights," in Feinberg, *Rights, Justice, and the Bounds of Liberty* (Princeton, NJ: Princeton University Press, 1980), 151.

[56] John Rawls, "Kantian Constructivism in Moral Theory," *The Journal of Philosophy* 77 (1980): 546.

[57] Hohfeld, *Fundamental Legal Conceptions,* 65–75.

[58] For directed obligation, see Margaret Gilbert, *A Theory of Political Obligation* (Oxford: Oxford University Press, 2006), 40; and for bipolar obligation, see Michael Thompson, "What Is It to Wrong Someone? A Puzzle about Justice," in *Reason and Value: Themes from the Moral Philosophy of Joseph Raz,* ed. R. Jay Wallace et al. (New York: Oxford University Press, 2006).

by the fact that were I to violate the right, you would have a distinctive authority to forgive me at your discretion. It is a reflection of this individual authority that you can consent to my acting in ways that would otherwise violate your right, that it is up to you whether to seek compensation (as is recognized in the law of torts), and so on.

When I step on your foot unbidden, I fail to respect, and thus violate, this authority you have as an individual person, and thus I violate and fail to respect *you*. In so doing, moreover, I fail to appreciate a reason not to step on your foot that is additional to the fact that I am causing avoidable harm: namely, that it violates your right (that is, on the current analysis, your legitimate demand of me as an individual person).

As I have said, it is relatively uncontroversial that the fact that an action would violate someone's right, and thus wrong that person, is a reason not to do it. What is in question is whether the fact that something would be wrong *period* is a reason that supplements the wrong-making features of a wrongful action, including, perhaps, that it would violate someone's rights and thus wrong that person. On the second-personal analysis we are considering, however, wrongs and moral obligations *period* are no less conceptually connected to legitimate demands—and hence to the idea of a second-personal reason—than are wrongings and bipolar moral obligations *to* someone.

Moral obligations simpliciter are what we have the authority to demand of one another, not as individuals, but as representative persons—as representatives, as we might say, of the moral community.[59] So wrongdoing also always involves a failure to respect legitimate demands and the authority to make them that is additional to the individual authorities that are involved in violations of rights and wrongings. Wrongdoing violates our authority as representative persons to demand that people not act in that way. This means that it also involves a failure of respect (of this *representative authority*) and thus a failure to appreciate a further reason not to perform the act that is additional to the act's wrong-making features, including, indeed, even the fact that it might violate someone's right (her *individual authority* to demand that one not so treat *her*) and so wrong her.

You have an individual authority to make claims and demands of me regarding my treatment of you that others do not have. You have the authority to allow or forbid actions involving you and your body that others don't, and you also have a distinctive standing to object, to demand apology, to seek compensation, to forgive, and so on. If you want not to "make a scene" by objecting if I step on your foot, then others have a strong reason to respect your wishes. And it is up to you and not others whether, for example, to forgive my brutish treatment. It is a reflection of

[59] Again, "the moral community" refers here not to any actual social collectivity but to a regulative ideal like Kant's "kingdom of ends."

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

all this, that violations of rights (wrongs *to* someone) are, as Strawson points out, the object of warranted *personal* reactive attitudes like resentment.[60] A stranger cannot warrantedly, or even intelligibly, feel resentment when I step on your foot, although you (and others personally related to you) can.

Nonetheless, your individual authority is not the only relevant standing in the case, not even the only relevant standing you have. Although it is up to you in the first instance whether or not to object or complain about my action, it is not distinctively up to you whether or not to feel indignation or blame. The latter is an *impersonal* reactive attitude and thus is felt not as if from the victim's point of view, but as if from a disinterested or impartial perspective that abstracts from personal involvement.[61] Impersonal attitudes are the ones to which the concepts of moral obligation and moral wrong *period* are conceptually related. Something is wrong if it is blameworthy, lacking adequate excuse. But whether to blame someone, whether that person is *to blame,* is not distinctively a victim's question. It is a question we face as representative persons, not as someone personally involved.[62]

This difference between our individual authority to claim our rights, on the one hand, and our representative standing to hold one another and ourselves accountable through blame, on the other, is reflected in the difference between the civil and the criminal law. Whether or not to seek compensation in bringing a tort action is up to the victim. The state legitimately does not proceed on its own to rectify wrongings; rather, the law of torts gives individuals standing to bring cases on their own behalf. But it is not, or at least not always, up to the victim whether to prosecute a crime. That is a question for the people and their representatives. In this way, legal guilt is like moral guilt. Whether to blame someone, whether she is to blame, is a question we face as representative persons. This is reflected also in John Locke's distinction between the authority everyone has in the state of nature to punish transgressions of the law of nature, on the one hand, and the authority that individuals have, "to seek reparations" for violations of their individual rights, on the other.[63]

Because it would violate your right, I have a reason not to step unbidden on your foot that is additional to the fact that doing so will cause

[60] Strawson, "Freedom and Resentment," 72.

[61] Ibid., 84–85.

[62] Thus, if it is a conceptual truth that violations of rights (wrongings) are also, other things being equal, wrongs period—that violations of "bipolar" obligations *to* someone are also, other things being equal, violations of moral obligations period—then it follows that a personal reactive attitude, such as resentment, can be warranted only if an impersonal reactive attitude would be (at least, other things being equal). I believe this conceptual thesis is true, but notice that the claim that the fact that an action is wrong is a reason not to perform it (consisting in the fact that the act violates a legitimate demand we make of one another as representative persons) does not strictly depend on this conceptual thesis. I am indebted here to discussion with Verity Harte and Jules Coleman.

[63] John Locke, *Second Treatise of Government,* sections 7–11.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

pain, suffering, and other things that are bad for you. Stepping on your foot does not simply harm you, it also disrespects your rights, and therefore disrespects you, as someone with the individual authority to demand that I not do that to you. And this gives me an additional reason not to step on your foot.

But if the fact that it would violate your right gives me reason not to step on your foot, so also does the fact that stepping on your foot would be wrong *period.* A wrongful act violates another authority in addition, not one you have as an individual, but one we all have as representative persons. In being blameworthy (unless excused), the wrongfulness of my act consists in its being a violation of some obligation that representative persons legitimately demand of me—that anyone can, at least in principle, demand (including me of myself) through reactive attitudes like blame (and the self-reactive attitude of guilt). Thus, all wrongs involve a kind of disrespect that is additional to disrespect of any individual's authority. Any wrong violates the legitimate demands of those to whom we are morally responsible, namely, one another as representative persons; thus, any wrong involves a failure adequately to respect this authority. Because this is so, an act's wrongness provides an additional reason not to perform it: namely, that the act would violate a legitimate demand and thus fail to respect our authority as representative persons. Since this reason is conceptually tied to an authority to address demands, an act's wrongness is therefore a second-personal reason not to do it.

Another way of seeing that wrongness creates an additional reason is to see the way in which being wrong interacts with the idea of reasonable demand. Since avoiding your foot puts no unreasonable burden on me, it is something we can demand of one another both as individuals and as representative persons. But what if stepping on your foot is the only way I can avoid significant injury to myself? In such a case, it seems arguable that we (and you) cannot reasonably demand that I not step on your foot, and so it is neither wrong nor a violation of your right for me to do so.[64] There seems to be a continuum of cases, holding fixed the harm to the victim, from ones where the burden to the agent is very great, and so his action not wrong, to ones where the burden is very small, and so his action clearly wrong. Let us suppose there are two different cases, A and B, between which there exists some just noticeable difference in the burdens to the agent, such that in case A the burden is insufficient to make the act of stepping on someone's foot not wrong (so in case A, the agent's stepping on the other's foot would be wrong), whereas in case B, with a somewhat greater cost to the agent, the same act would not be wrong. If

[64] Or, alternatively, though it violates your right and *would* be wrong (period) if I lacked this further justification, it is not wrong because the justification exists. What violates someone's rights would be wrong to do lacking some justification, just as what is wrong would be blameworthy if done without adequate excuse. It follows that what violates someone's rights would be blameworthy if it were done without *either* a justification or an excuse.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

we consider just the wrong-making features of the act—the relative cost to agent and victim—these will, of course, create different and differently weighted reasons. That an agent could avoid stepping on someone else's foot at lesser cost in case A than he could in case B no doubt creates greater reason for the agent to absorb that cost in case A than in case B. But whatever reason the wrong-making features themselves create in these two cases seems not fully to capture the reasons that result from the further fact these features ground, namely, that we can legitimately demand that an agent absorb the cost involved in case A, but not in case B—in other words, that not absorbing the cost in case A amounts to wronging someone and thereby doing moral wrong (assuming there is no further justification).

V. A Conclusion Concerning Conclusive Reasons

Finally, it is an often-noted feature of the concepts of moral obligation and moral wrong that they purport to entail that an agent has not just *some* reason to act, but conclusive reason (or at least, as Scanlon says, reason that is "normally conclusive").[65] Obviously, this is no part whatsoever of the concepts involved in obligation-making or wrong-making features themselves: for example, that stepping on your foot in the case we initially imagined would impose a considerable cost on you at little cost to me. That does not rule out a buck-passing theory of wrongness with respect to reasons for acting by itself, however, since it might be that the concept of wrongness just is that of there being conclusive moral reasons (or reasons period) for doing something, with the reasons being provided entirely by wrong-making features. But we have already effectively excluded these possibilities. If action supported by conclusive moral reasons (or reasons period) just means action recommended by the weightiest moral reasons (or reasons period), we have already seen that this is a different concept from that of morally obligatory or demanded conduct. So far as our concepts go, an act may be thus supported or recommended without being morally required or obligatory. It follows, therefore, that if being morally obligatory purports to entail the existence of conclusive or normally conclusive reasons, then that must be because the facts of moral obligation and wrongness purport to provide some reasons that can guarantee overriding weight *themselves.*

In conclusion, I would like to indicate how the second-personal analysis of moral obligation and wrongness I have been discussing can explain the conclusive-reason-providing purport of moral obligation and moral wrong. The root of the explanation is moral obligation's conceptual tie to accountability and warranted blame. As Bernard Williams and others have noted, when we blame someone we imply that there was conclusive

[65] Scanlon, "Wrongness and Reasons," 6.

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press

reason for her not to have done what we blame her for having done.[66] It makes no sense to blame someone for doing something and then add that she had, nonetheless, sufficient reason to do it, all things considered. Blame implies not just that the person shouldn't have done what she did "morally speaking," but that she shouldn't have done it period. A person who in one moment "admits" her guilt, but in the next asserts that she had, nonetheless, good and sufficient reasons for doing what she did, can hardly be said to have accepted responsibility for her action. Or to put the point the other way around, if someone can establish that she had sufficient reason to do what she did, then she will have accounted for herself, justified her conduct, and shown thereby that blame is unwarranted. The point is not that the fact that one blames someone entails the existence of conclusive reason. Rather, in blaming someone one implies that such reason exists.

When we think that an act is morally wrong, and therefore blameworthy if unexcused, we think that, lacking excuse, blame is the fitting attitude for anyone to take toward someone performing the act. When an agent takes such an attitude toward herself, and addresses a demand to herself through a self-reactive attitude like guilt (whether retrospectively or prospectively), she feels as though there is conclusive reason not to perform the act.[67] But if the facts that *make* an act morally wrong, in the sense of grounding or providing normative reasons for its wrongness, do not purport to be conclusive reasons themselves, and if the fact that an act is wrong is not *just* the fact that such conclusive reasons exist, then what we must be thinking when we take such an attitude toward one another and ourselves is that the fact that an act is morally wrong provides a reason itself not to perform the act, and that this guarantees conclusive reason not to do so. I have tried to say in this essay what this additional reason consists in: that the act is something we legitimately demand that people not do.[68] If this is right, trying to do without this thought will leave us with one thought too few.

Philosophy, Yale University

[66] See Bernard Williams, "Internal Reasons and the Obscurity of Blame," in Williams, *Making Sense of Humanity* (Cambridge: Cambridge University Press, 1995), 40–44; Gibbard, *Wise Choices, Apt Feelings,* 299–300; Skorupski, *Ethical Explorations,* 42–43; and Shafer-Landau, *Moral Realism: A Defence,* 181–83.

[67] For valuable insights about the relevance of prospective guilt, I am indebted to the work of Howard Nye.

[68] It should be clear that I am not saying that this is a wrong-making feature, in the sense of a ground or of normative reason for an act's wrongness. I am saying that this is what being wrong consists in, and that this fact provides a reason not to perform the wrongful act that is additional to the act's wrong-making features (grounds of or normative reasons for its being wrong).

https://doi.org/10.1017/S0265052509990185 Published online by Cambridge University Press