

CHAPTER II

THE SECOND-PERSON STANDPOINT

STEPHEN DARWALL

DRAFT: PLEASE DO NOT QUOTE OR DISTRIBUTE WITHOUT THE AUTHOR'S PERMISSION

CHAPTER II

THE MAIN IDEAS II

MORAL OBLIGATION'S PURPORTED NORMATIVITY

The most familiar characterization of moral obligations' purported normative force is in terms of their putatively *categorical* character, their purporting to be what Kant called "categorical imperatives." (Kant 1996c) Philippa Foot points out that this cannot mean just that moral obligations are categorical in form, that they "apply" independently of their relation to the agent's ends. (Foot 1972) Requirements of etiquette are unconditional in this way also, but nobody thinks that their putative normative force is anything like that of moral requirements. Foot concludes that the Kantian claim must rather be that moral obligations "necessarily give reasons for acting to any man." (Foot 1972: 309) Frequently this claim is strengthened to what Scheffler calls the "thesis of overridingness": "it can never be rational knowingly to do what morality forbids."¹ (Scheffler, 1992) In other words, moral obligations always give agents conclusive reasons for acting that outweigh or take priority over any potentially competing considerations; or, at least, they always purport to do so.

This surely captures part of what is felt to be special about moral requirements, but only part. Compare logical requirements. If a conclusion follows from known premises, then we are under a requirement of reason to infer it that apparently applies to all rational beings also, considerations to the contrary notwithstanding. But there is an important difference between moral obligations and requirements that are imposed by logic. Morality involves a distinctive kind of accountability by its very nature. If I fail to act as I am morally required without adequate excuse, then distinctively second-personal responses like blame and guilt are thereby warranted.² But it is only in certain contexts that responses like these seem appropriate to logical blunders, and even here what seems to be in question is a moral error

of some kind (as when I have a special responsibilities for reasoning properly).³ Moreover, although a connection to accountability is, I shall argue, part of the very concept of moral obligation, it is obviously no part whatsoever the idea of a logical requirement.⁴

Mill famously remarked 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.” (Mill 1998: Ch.V) We may think there are strong moral reasons for people to do something and “dislike or despise them for not doing” it; but unless we think “blame” or some other form of “punishment” is warranted, perhaps just the feeling of guilt, we do not think it “a case of moral obligation.” (Mill 1998: Ch. V) Many philosophers today broadly concur with these Millian sentiments.⁵ (Adams 1999: 238; Baier 1966; Brandt 1979: 163-176; Gibbard 1990: 42; Skorupski 1999: 142) Matters of moral obligation are the moral community’s “business,” as Kurt Baier puts it. (Baier 1966: 223) When we are morally obligated, we are not morally free to act otherwise; members of the moral community have the authority to hold us responsible if we do. Moral obligations are thus to others in a more robust way than those of logic are.⁶ Anyone who feels guilty about logical errors would seem to have a “moralized” sense of the logical.

I shall develop this theme and argue that any account of the distinctive normativity of moral obligation that fails to capture this second-personal element is deficient. What’s more, I shall argue that this second-personal aspect also helps to explain why moral obligations purport to be supremely authoritative in the more familiar sense of being categorical and overriding.

The root of this latter explanation is the conceptual connection between moral obligation and responsibility, specifically, obligation’s tie to warranted blame. Along lines

similar to Bernard Williams and others, I shall argue that blame implies that the blamed agent had reason enough not to do what we blame him for doing in something like the way that Moore held that asserting p implies that one believes that p . (Williams 1995: 40-44; Gibbard 1990: 299-300; Skorupski 1999: 42-43; Shafer-Landau 2003: 181-183; Moore 1942: 540-543) Just as it is unintelligible to assert something but deny that one believes it, so also, I claim, does it make no sense to blame someone for doing something and then add that he had, nonetheless, sufficient reason to do it, all things considered. It is common to blame and other second-personal reactive attitudes through which we hold one another responsible that they presuppose, not just that the person shouldn't have done what he did "morally speaking," but that he shouldn't have done it period. A person who in one moment "admits" his guilt, but in the next, asserts that he had, nonetheless, good and sufficient reasons for doing what he did could hardly be said to have accepted responsibility for his action. Or to put the point the other way around, if someone can establish that he had sufficient reason to do what he did, then he will have accounted for himself and shown thereby that blame is unwarranted.

THE SCOPE OF MORAL OBLIGATION

I should emphasize that nothing in the idea of moral obligation as involving reciprocal accountability rules out its scope or content extending beyond the needs and interests of free and rational individuals considered as such. For all I shall say in this book, what we are morally responsible for might include, for example, the protection of cultural treasures, wilderness, and/or the welfare of other sentient beings, quite independently of the relation any of these have to the interests of free and rational persons. I take no stand here on whether or not this is so. What I say will entail that if we have such moral obligations then these are among the things we free and rational agents have the authority to demand of

one another. Of course, even if we do not, even if, say, harming wilderness or members of other species were not in itself to violate any demand for which we can be held morally accountable, there would still be weighty reasons against such harm. In any case, what I shall be arguing is that moral obligations essentially include demands free and rational individuals have standing to make of one another as such and that we are committed to the standing to make these demands by presuppositions of the second-person standpoint. Whether we are morally accountable for more is obviously an important question, but it is not one I can address here.⁷

Of course, even if moral obligations concerning other species, say, can be accommodated within the kind of framework I am sketching, it may still be objected that there is no place in it for the idea that these obligations are owed to the beings themselves if they lack second-personal competence and so, according to my argument, the authority to demand anything. But the kind of view I am developing may have resources for response even here. First, although I am bound to insist that moral obligation, like the concept of a right, cannot be understood independently of authoritative demands, the thought that moral obligations can be owed to beings who lack second-personal competence might be able to be elaborated in terms of trustees' (for example, the moral community's) authority to demand certain treatment on their behalf (perhaps also to claim certain rights, compensation, and so on, for them). Thus, The Lorax (a free and rational person) declares, "I speak for the Trees." (Dr. Seuss 1971) Second, to the extent that we find the thought that we owe obligations to nonrational beings a natural thing to think, it seems likely that we also impute to them a proto- or quasi-second-personality, for example, that we see an animal's or an infant's cry as a form of complaint. In any case, what I shall seek to show in this book is that the second-personal competence that makes us subject to moral obligation

also gives us an authority to make claims and demands of one another as members of the moral community. Whether the scope or content of moral obligation extends farther, is a question I shall not here consider.⁸

VINDICATING MORAL OBLIGATION:

THE KANTIAN PROJECT

Whether moral obligations purport to be supremely authoritative is one question, but whether they actually are is another.⁹ The project of vindicating morality's putative authority poses one of moral philosophy's great challenges. A mark of the Kantian tradition has been to take the challenge seriously and attempt to meet it. (Nagel 1970; Gewirth 1978; Darwall 1983; Korsgaard 1996c) Kant's own strategy apparently changed significantly from the *Groundwork* to *The Critique of Practical Reason*, however. What held constant throughout was his "reciprocity thesis," Kant's conviction that the moral law is equivalent to what he called "autonomy of the will": "the property of the will by which it is a law to itself independently of any property of the objects of volition." (Kant 1996b: 440; Allison 1986)

Consider, first, how Kant argues in the *Groundwork*. At the end of *Groundwork's* Section II, Kant remarks that it is consistent with his arguments to that point that morality is nothing but a "figment of the mind." He believes he has already proven in Sections I and II that morality's fundamental principle is the Categorical Imperative (CI) and that it holds if, and only if, the will has the property of autonomy (the reciprocity thesis). But so far he has simply analyzed the "generally accepted" concept of morality. (1996c, 445) All he has shown, consequently, is that *if*, and only if, there is such a thing as morality is conceived to be, then "the categorical imperative, and with it, autonomy of the will, is true and absolutely necessary as a principle *a priori*." (1996c: 445) To vindicate the moral law, however, it is insufficient to establish this biconditional, since it is merely an analytic truth that is

consistent with morality's and autonomy's both being mere "figments." He must establish either that autonomy really holds or that the CI does (either will do because of the reciprocity thesis). And both, he says, are synthetic *a priori* claims that require the kind of critique of practical reason he sets about providing in Section III.

In *Groundwork* III, Kant attempts to establish autonomy of the will as an independent premise and to prove the CI on its basis. He claims that autonomy is a necessary presupposition of the practical standpoint—we “cannot act otherwise than *under the idea of freedom*”—and then argues (via the reciprocity thesis) that in presupposing autonomy a rational agent is committed to the CI (the moral law). (Kant 1996b: 448) Contemporary Kantians, most notably Christine Korsgaard, often pursue some version of this strategy also. (Korsgaard 1996b, 1996c, 1996d; Hill 1985)

I shall argue that any such attempt must fail. Autonomy of the will, as Kant defines it, is a form of freedom that is distinctively practical and that lacks any analogue in theoretical reasoning. However, the only argument that Kant gives in *Groundwork* III that we must presuppose autonomy in practical reasoning is that we must assume a freedom of reason in forming any judgment, whether practical or theoretical. (Kant 1996c: 448) Kant's idea seems to be that if we see our thinking or judgments in any area as simply directed from without, we will not be able to regard ourselves as reasoning or making rational judgments in that area. No doubt this is true, but I shall argue that it doesn't follow that the reasons on the basis of which we judge are not themselves drawn from features of our judgment's object (and hence, in practical deliberation, from properties of the objects of our desire and volition). When we reason about what to believe, at least about empirical matters, for example, there is a clear sense in which we need to see our thinking as responding to independent objects and states of the world about which we are attempting to form beliefs.¹⁰

Theoretical reasoning evidently involves no form of freedom that is analogous to autonomy of the will (the will's being "a law to itself independently of any properties of the objects of [desire and] volition").

But neither, I argue, need autonomy be assumed in just any intelligible practical reasoning. Consider what practical reasoning would be like if it were structurally analogous to rational belief formation about the empirical world. (I am not saying for a moment that this is a plausible picture of practical reasoning, only that it is an intelligible one, and that since autonomy would not be assumed in such reasoning, it cannot, therefore, be an inescapable presupposition of the practical standpoint.) A naïve theoretical reasoner takes her experience as evidence of states of an independent world (the objects of correct beliefs). An analogously "naïve" practical reasoner might take her desires and other forms of "practical experience" like pleasure and pain as forms of epistemic access also, only this time, not to the world as it is, but to how it would good or bad for the world to be, how it should or should not be.¹¹ To an agent with a desire that p it will seem that the world should be such that p , as if p would be a good thing, hence, as if there is reason for her to bring p about, and similarly for experiences of being pleased or pained that p . Naïve practical reasoning on this basis would therefore see reasons for acting as deriving from features of the objects of desire and volition and so would not presuppose autonomy in Kant's sense.

So I shall argue that Kant's *Groundwork* argument fails, as must, indeed, any argument that attempts to prove that autonomy is an inescapable assumption of any possible practical reasoning.¹² The problem, as I will diagnose it, is that it is only deliberation from the second-person standpoint that requires us to assume autonomy of the will.

By the time Kant wrote *The Critique of Practical Reason*, he had apparently abandoned the *Groundwork's* strategy of trying to establish the moral law from a prior premise of

autonomy as an inescapable assumption of any practical reasoning. There he holds that we have no independent way of establishing autonomy other than through our consciousness of a particular kind of reason, namely, being bound by the moral law. What Kant calls “the fact of reason,” the awareness of our freedom through consciousness of being bound by the moral law, itself assumes morality’s supremacy and cannot therefore be used to establish it. (Kant 1996d: 31)

I think Kant was right to give up the strategy of the *Groundwork* and to argue that consciousness of autonomy of the will involves an awareness of being bound by and able to act on a reason of a distinctive kind. Perhaps we can now glimpse why I shall claim that the requisite kind of reason must be second-personal. First, as I mentioned in the last chapter, I shall be arguing that any second-personal address or acknowledgment commits us to the equal dignity of persons, conceived as equal second-personal authority, and to the idea that addresser and addressee alike are capable of acting on reasons that are grounded in this authority and so are irreducible to the value of any outcome or state (hence to any property of any object of desire). Second-personal address thus commits us to the idea that agents capable of entering into relations of mutual accountability must have within this capacity (second-personal competence) a second-personal authority that is a source of reasons and norms independently of any features of desired (or desirable) outcomes or states along with the ability to act on these norms and reasons. When we acknowledge the summons of another free and rational agent, we confront, in effect, the “fact of reason.” We presuppose the equal dignity of rational beings and our ability to act on a “law” or reason—a second-personal reason grounded in this dignity—that derives not from the value of any state of affairs or outcome that might be the object of a desire, but, ultimately, from what it is to be one free and rational person interacting with others.

Second, consider briefly how the “fact of reason” functions within Kant’s own argument in the second *Critique*. Kant asks us to imagine someone whose prince demands that he give false witness against an honorable man on pain of death. (Kant 1996d: 30) Kant notes that no one can say for sure what he would do in such a situation, but Kant nonetheless insists that anyone in the situation would have to admit that he could refuse the prince’s request, since he would judge that he should do so. Because we all already implicitly acknowledge the authority of the moral law, we are forced to conclude that we have the capacity to act on it. We cannot coherently think that it is impossible to do something and still think, from the practical standpoint, that it is something we should do. If we can’t do something, then the practical question is what else to do in light of that fact.

Obviously, if acknowledging the authority of moral obligation is what requires us to assume we can act on a law that binds us independently of our desires (including for survival), then we cannot assume autonomy in an argument for the authority of moral obligation. But as I’ve said, Kant is no longer interested in that strategy. Nonetheless, Kant clearly takes this example as evidence, not just that anyone must admit that he could comply with this moral obligation (in accepting it), but that he can comply with the Categorical Imperative in general, since Kant presents a formulation of the CI (“So act that the maxim of your will could always hold at the same time as a principle in a giving of universal law”) just after the passage we have been discussing.¹³ (1996d: 30) And that, after all, is what would be necessary for him to have established that we must accept autonomy of the will—that we are bound by a *formal* principle of the will. But why should that be? All we know so far is that we can act on a principle of right we accept, irrespectively of our desires for any outcome. It is consistent with this that, as deontological intuitionists like Richard Price and W. D. Ross believed, this is explainable by an independent fact of right rather than by

anything like autonomy of the will. (Price 1974; Ross 1930) Why suppose there must be some formal principle of the will like the CI that stands behind our moral obligations?

Suppose we were to consider, however, not just that anyone in the situation Kant describes ought to reject the prince's demand, but also that he would be morally responsible for doing so. In the second-personal address of holding him responsible we would have to assume that he had within him a source of motivation to do as he was morally obligated. We cannot intelligibly hold someone to a demand as a moral agent without supposing that he could hold himself to that same demand by acting on the relevant second-personal reasons. And if we think that any rational person in that situation would be thus answerable, we are committed to thinking that what makes a rational person subject to moral obligation must itself include a source of motivation to do as he is morally obligated. The need for this presupposition comes, not from supposing there to be normative reasons for someone to do something, regardless of their priority or weight; but rather from the distinctively second-personal reasons that are involved in moral obligation. It comes from what Gary Watson calls "constraints of moral address." (Watson 1987: 263,264) It follows, I shall argue, that the second-personal competence that makes us subject to moral obligation must include a source of the (second-personal) reasons in which moral obligation consists along with the capacity to act on these reasons. In presupposing this, I shall claim, we effectively presuppose autonomy of the will. The capacity of will that make us apt to be held responsible, second-personal competence, is a "law to itself," since it is the basis of second-personal authority.

A SECOND-PERSONAL INTERPRETATION OF THE CI

What then does second-personal competence consist in? Well, it must consist in something like the capacity to make demands of oneself from a second-person standpoint: in

being able to choose to do something only if it is consistent with demands one (or anyone) would make of anyone (hence that one would make of oneself) from a standpoint we can share as mutually accountable persons. But that is just a second-personal version of the CI. In fact, I shall argue, the most natural way of interpreting the CI and kindred moral principles like the Golden Rule, is second-personally, in terms of demands that one (anyone) would sensibly make of all from the shared standpoint of a member of the moral community. What matters for moral obligation is not what one would like or prefer all people do, but what one would expect of others, what demands one could endorse anyone's being able to hold others to as members of a community of mutually accountable equals. What forces an assumption of autonomy of the will, therefore, is the second-personal aspect of moral obligation, that is, that what is morally obligatory is what we are responsible to one another for doing.

But isn't acting on demands that others can make of one heteronomy rather than autonomy, being governed by them rather than by oneself? Or, to put it another way, how could autonomy consist in a law that comes from nothing outside of the will itself if it is realized in second-personal interaction? The response to this objection is that when one decides to reject the prince's demand because this is what the moral community authoritatively demands, the second-personal perspective of a member of the moral community is as much one's own as it is anyone else's. One demands the conduct of oneself from a point of view in which one shares as a free and rational person.

A FOUNDATION FOR CONTRACTUALISM

An account of the second-person standpoint along the lines that I am attempting to develop in this book provides, I believe, the most natural way of motivating the moral theory known as "contractualism." Contractualists sees morality as most fundamentally concerned

with “how *persons are to relate to one another*,” to quote one insightful commentator.¹⁴ (Kumar 1999: 284) Or as Scanlon puts it, “the contractualist ideal . . . is meant to characterize” a specific “relation” with others: “mutual recognition.” (Scanlon 1998: 162) Moral principles have a distinctive *role* on a contractualist view: they mediate mutual respect.¹⁵ They don’t just tell us what actions we should or must perform; they mediate a fundamental interpersonal *relation*: mutual respect between mutually accountable persons.¹⁶ These are all, I shall be arguing, irreducibly second-personal matters.

For Scanlon, the problem of accounting for morality’s authority is that of explaining “the *priority* of right and wrong over other values” or accounting for their special “importance.” (1998: 148, 149) In the paper in which he originally formulated the position, Scanlon maintained that the “motivational basis” of contractualist morality is a desire to act in a way that can be justified to others. (Scanlon 1982) In *What We Owe to Each Other*, however, Scanlon departs from his prior view and claims that contractualism’s basis is the “value” and “appeal” of standing to others in the relation of mutual recognition, both in itself and as an ineliminable aspect of other valuable relations such as friendship. (1998: 158-168)

This is obviously an attractive idea, but like Mill’s attempt to explain the bindingness of morality in the appeal of “unity with our fellow creatures,” to which Scanlon in fact compares it, it is hard to see how it can adequately account for the nonoptional character of moral obligation. (Mill, 2002: Ch. III) This is another instance of Strawson’s Point; such an argument seems to provide a reason of the wrong kind. By contrast, the approach I shall sketch promises to explain how the standing to make claims and demands on one another as free and rational persons is something to which we are jointly committed whenever we take up the second-person stance more generally rather than simply being an ineliminable aspect

of interpersonal relations we find appealing.

A central issue in the debate between consequentialist theories of right and deontological theories such as contractualism has been whether a satisfying philosophical rationale can be provided for agent-relative restrictions (“deontological constraints”). (Scheffler 1982; Nagel 1986) Consequentialists since Moore have argued that if, for example, betraying or causing harm to others is to be avoided, then that must be because these are bad things.¹⁷ But if they are, then they are no less bad no matter whom they are done by. So if, for example, an agent could do something that would amount to her betraying another person but would also prevent an exactly similar betrayal involving two further people, say because another would-be-betrayer is sufficiently shocked at seeing the first agent’s betrayal that she gives up her treacherous plans, these two outcomes should cancel each other out from the moral point of view. The fact that one betrayal would be by the agent herself should make no intrinsic difference to what there is reason for her to do. There will end up being one betrayal in the world whatever she does.

It is well known that, in addition to being part of moral common sense, agent-relative constraints can be derived within contractualism or, for that matter, within indirect consequentialist approaches such as rule utilitarianism.¹⁸ The problem has remained of how to give any of these derivations a deeper philosophical rationale that itself confirms rather than undermines the case for agent-relative constraints. The problem with indirect consequentialism, for example, is that even if we think there are (agent-neutral) consequentialist reasons to call actions that violate common sense agent-relative moral constraints “wrong,” or to want people (ourselves included) to think them wrong, it is not obvious how a consequentialist can herself sensibly think (simultaneously with accepting the deeper agent-neutralist reasons) that agents really should avoid performing such actions,

when doing so would advance overall agent-neutral value.

If a foundation for contractualism can be found in commitments within the second-person standpoint, this can provide, I believe, the kind of rationale for agent-relative restrictions that is needed. For it would vindicate reasons that are in their nature relational, that concern, not how it would be good for the world to be anyway, or even what kinds of acts are called for by their intrinsic nature, but how we are to relate to one another owing to claims and demands that we cannot avoid assuming we have the authority to address to each other as one free and rational person among others. It would provide the right kind of reason.

¹ David Brink calls this the “supremacy thesis.” (Brink 1997: 255)

² 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.

³ I don’t mean, of course, that logical errors aren’t subject to criticism, or that we don’t sometimes use words like “blame,” as when a teacher says that he doesn’t blame his student for a given error on a first try.

⁴ Again, I am indebted to Peter Graham for this point.

⁵ See also Ewing’s distinction between two senses of ‘ought’, one concerning the weight of normative reasons, the other conceptually related to blame. (Ewing 1939: 3) I am indebted to Howard Nye for reminding me of this passage.

⁶ See Thompson 2004.

⁷ I am indebted to Allen Buchanan for pressing me to clarify the points in this paragraph.

⁸ I am indebted to Jim Staihar and Howard Nye for discussion of the points in this paragraph.

⁹ Alternatively, we could say that moral obligations are supremely authoritative by definition, but then the question just becomes whether what we take to be moral obligations are really. So far as I can see, nothing hangs on this semantic choice.

¹⁰ This point holds whether or not transcendental idealism is true. Even a transcendental idealist will need to draw a distinction between object- or state-responsive thinking (like that involved in ordinary belief formation about states of the world (the objects of empirical belief), and thinking that isn't. It is significant in this regard that Fichte contrasts practical freedom from any we can have in "representational activity." (Fichte 2000: 19), as does Kant in the *Critique of Practical Reason* (Kant 1996d: 20).

¹¹ This is not far from the picture Dennis Stampe provides in "The Authority of Desire." (Stampe 1988)

¹² As I shall argue in Chapter IX, I differ here with writers like Korsgaard and Wood who attempt to provide a sound reconstruction of the Formula of the End in Itself, autonomy, and related concepts from presuppositions of practical standpoint alone. (Korsgaard 1996b, 1996c, 1996d; Wood 1999) As I will point out in more detail later, although Korsgaard shows a keen appreciation of the role of reciprocity and mutual accountability for a Kantian framework (Korsgaard 1996a: 301 and 1996e), she also believes that autonomy and the Categorical Imperative in its various formulations are transcendental presuppositions of the practical standpoint. I believe this to be mistaken. In my view, there is no way of deriving these ideas outside of the second-person standpoint and within an overall a comprehensive theory of practical reason that is enriched by it.

¹³ Which Kant there calls the "Fundamental Law of Pure Practical Reason." (1996d: 30)

¹⁴ Compare Korsgaard's remark that, on a Kantian view, "the subject matter of morality is . . . how we should relate to one another." (Korsgaard 1995a: 275) See also Schapiro 2001 and 2003a.

¹⁵ As Rahul Kumar says, they provide a "basis for a shared understanding of the kind of consideration . . . persons may legitimately expect of one another, as a matter of mutual respect from one another as persons." (1999: 284)

¹⁶ Compare Elizabeth Anderson's relational theory of equality in Anderson 1999.

¹⁷ For some of the history of this debate and a different proposal about how to justify agent-relative constraints, see Darwall 1986a.

¹⁸ See, e.g., Kumar 1999. For rule utilitarianism, see Brandt 1979; Parfit 1984; Johnson 1991; Hooker 2000. For a (seemingly) dissenting view, see Brand-Ballard 2004. I say "seemingly" because Brand-Ballard's arguments put pressure on the claim that the agent-relative deontological constraints of the strength of common sense can be derived within contractualism, although not, perhaps on the claim that agent-relative constraints of some sort might be.