

AUTHORITY AND SECOND-PERSONAL REASONS FOR ACTING

In recent work, I have been arguing that a distinctive kind of reason for acting, a *second-personal reason*, is an ineliminable aspect of many central moral categories, including rights, moral responsibility, moral obligation, respect for and the dignity of persons, and the very concept of moral agent or person itself.¹ Second-personal reasons are distinguished from reasons of other kinds by their conceptual relations to authoritative claims and demands that must be able to be *addressed to* those to whom they apply (second-personally, as it were). I have argued, more specifically, that there are four interdefinable, irreducibly second-personal notions: the authority to make a claim or demand, a valid (authoritative) claim or demand, responsibility *to* someone (with the relevant authority), and a second-personal reason for acting (that is, for complying with an authoritative claim or demand and so discharging the responsibility). Each one of these notions entails the other three, and no proposition that does not already involve one of these four concepts can entail any that does.

Call the authority I am concerned with *practical authority* to distinguish it from various forms of theoretical authority or expertise, including the kind of authority on practical matters that a trusted advisor might have. A challenge that can be posed to my thesis is that it is possible to establish claims to practical authority by invoking solely non-second-personal reasons. In what follows, I consider Joseph Raz's influential *normal justification thesis* according to which claims to practical authority can be established by showing that an "alleged subject" is likely to comply better with reasons that apply to him independently already if he accepts the directives of an alleged authority as binding and tries to follow them than he would if he were to act on his own assessment of independent reasons. (Raz 1986: 53) I shall argue that the normal justification thesis does not hold for practical authority and, moreover, that the correct diagnosis of why it doesn't shows that both practical authority and reasons that are conceptually related to it are irreducibly second-personal.

Let me try to get the intuitive idea of a second-personal reason across with an example. Compare two different ways in which you might try to give someone a reason to stop causing you pain, say, to remove his foot from on top of yours.

One would simply be to get him to have sympathetic concern for you in your plight and to want you to be free of pain. In desiring this, he would see relief of your pain as a better way for the world to be, a possible outcome or state that, as Moore put it, "ought to exist for its own sake." (Moore 1993: 34) And he would most naturally see his desire, not as the source of the reason, but as a kind of access to an *agent-neutral* (and *state-of-the-world-regarding*) reason for removing his foot that is there anyway.² The reason would not be

¹ For example, in Darwall 2004 and 2006.

² On this point, see Darwall 1983; Bond 1983; Pettit and Smith 1990; Quinn 1991; Hampton 1998; Scanlon 1998: 41-55; and Dancy 2000. Agent-neutral reasons contrast with agent-relative reasons, those whose formulation includes an ineliminable reference to the agent for whom they are reasons (like "that it will keep a promise I made," "that it will avoid harm to others (i.e., people other than me," and so on). Agent-neutral reasons can be stated without such a reference: "that it would prevent some pain from occurring to someone (or some being)." On the distinction between agent-relative (also called "subjective" or "agent-centered") and agent-neutral (also called "objective") reasons, principles, values, etc., see Nagel 1970; Scheffler 1982; Parfit

essentially *for him* as the agent causing another person pain. It would apparently exist, most fundamentally, for anyone who is in a position to effect the state of relief of your pain, and *therefore* for him, since he is well placed to do so.³ Finally, in “giving” him the reason in this way, you might not need to address or relate to him in any way at all. Anything that would get him to see your being in pain as a bad thing, like an unaddressed grimace or whimper, might serve. In no sense, not even epistemic, need he be taking any reason to move his foot on your authority.

Alternatively, you might lay a claim or address a purportedly valid demand. You might say something that asserts or implies your authority *to* claim or demand that he move his foot and that simultaneously expresses this demand. You might demand this as the person whose foot he is stepping on, thereby claiming and exercising what you take to be a right against him. Or you might demand it as a representative of the moral community, whose members understand themselves as holding one another to a (moral) demand not to step on each other’s feet. Or you might do both simultaneously. Whichever, the reason you would address would be agent-relative rather than agent-neutral. It would concern, most fundamentally, your addressee’s relations to others, viewed from a perspective within those relations, in this case, that his keeping his foot on yours causes another person pain, causes inconvenience, and so on, and that this is something we can and do reasonably demand that people not do. The reason would not be addressed to him as someone who is simply in a position to alter a bad state, whether of someone’s being in pain or even of someone’s causing another pain. If he could stop, say, two others from causing an identical gratuitous pain by the shocking spectacle of keeping his foot firmly planted on yours, this second, claim-based reason would not recommend that he do so. The reason would be addressed to him as someone who is himself *causing* gratuitous pain to another person, something we persons normally assume we have the authority to demand that we not do to one another.

What is important for our purposes is that someone can sensibly accept this second reason for moving his foot, one embodied in your claim or demand, only if he also accepts your *authority to demand* this of him (second-personally). That is just what it is to accept something *as a valid claim or demand*. And if he accepts that you can demand that he move his foot, he must also accept that you will have grounds for complaint or some other form of accountability-seeking response if he does not. Unlike the first reason, this latter is second-personal in the sense that although the first is conceptually independent of forms of second-personal address involved in making claims and holding people responsible, the second is not. A *second-personal reason* is thus one whose validity depends upon presupposed authority and accountability relations between persons and, therefore, on the possibility of the

1984; Nagel 1986; Darwall 1986; and McNaughton and Rawling 1991. For a discussion that raises a question about the value of this distinction, see Korsgaard 1996a.

I argue for the claim that sympathetic concern involves its seeming that there are agent-neutral reasons to further someone’s welfare in Darwall 2002: 68-72. I do not deny, of course, that someone who already accepted various agent-relative norms, might not be moved through empathy and sympathy, to feel some special responsibility for relieving the pain. My point is that this would not come through sympathy alone.

³ Roughly speaking, again, a reason is agent-neutral if it can be formulated without essential reference to the agent (as such); otherwise it is agent-relative. It should also be noted that superficially agent-relative reasons may be grounded more deeply in agent-neutral considerations and values, and/or *vice versa*. For example, rule-utilitarianism holds that rules of right conduct include agent-relative principles, for example, those defining rights of promise and contract, on grounds of overall agent-neutral value.

reason's being addressed person-to-person within these relations. Reasons of this kind simply wouldn't exist but for their role in second-personal address and in mediating our relations to one another. And their second-personal character explains their agent-relativity. As second-personal reasons always derive most fundamentally from agents' *relations* to one another, they are invariably agent-relative at the most fundamental level.⁴

It is important also to see, however, that a norm or reason can be agent-relative without entailing or presupposing second-personal reasons. We can imagine someone who accepts and scrupulously observes a universal norm of foot-avoidance but who also denies, consistently with that, anyone's authority to claim or demand his compliance with this norm, hence denies that he is responsible to anyone for compliance, even to God. Such a person might conceive of the norm as mandatory in the sense of entailing categorical, indeed supremely authoritative, or "silencing," or even "pre-emptive" or "exclusionary" reasons, without accepting that he is accountable to anyone for complying with it.⁵ However, he could not then consistently accept that anyone has a *right* to his foot-avoidance, and in respecting the norm of avoiding people's feet he would not be respecting *them* as persons, since he would not be recognizing any authority anyone might claim as a person to demand anything, in particular, that he avoid their feet. (Feinberg 1980; Darwall 2004) Neither, in my view, could he consistently accept that he is *morally obligated* not to step on others' feet since moral obligation is related to moral responsibility conceptually. It is conceptually impossible for one to be morally obligated to do something but not responsible for doing it, neither to the moral community, nor to God, nor to anyone. So someone who thought he was accountable to no one could not think he was morally obligated not to step on others' feet, whatever priority he might give to a norm requiring him not to do so.

There is, in other words, a significant difference between the idea of an authoritative claim or demand, on the one hand, and that of an authoritative, valid, or in some thin sense "binding" norm or normative reason, or even of a normative requirement, on the other. There can be requirements *on* us that no one has any standing *to require of* us. We are under a requirement of reason not to believe propositions that contradict the logical consequences of known premises, for example. But it is only in certain contexts, say, when you and I are trying to work out what to believe together, that we have any standing to demand that one another reason logically, and even here that authority apparently derives from a moral or quasi-moral aspect, namely, our having undertaken a common goal.⁶ Requirements of

⁴ The formulation of the reason may not always be agent-relative, however. Suppose, for example, that the best way of grounding the Categorical Imperative is, as I believe, in an equal authority to make claims and demands that persons presuppose when they address one another second-personally. It is at least conceivable that what the categorical imperative itself requires is a principle of conduct that can be specified agent-neutrally. R. M. Hare, for example, believes that the categorical imperative can be seen to entail the sort of universal prescriptivism he favors *and* that this entails a form of act-utilitarianism (an agent-neutral theory). See Hare 1993.

⁵ A reason "silences" other reasons if it cancels their weight (and thus does not simply outweigh them. (McDowell 1979) A reason is "exclusionary" or "pre-emptive" if it is not to be added to other reasons, but to replace or "exclude" them. (Raz 1976 and 1986: 46)

⁶ Of course, these further constraints are frequently in the background, as they are, for example, whenever we do philosophy, say, right now. Because of the relationship you and I are currently in, each of us *does* have authority to call one another to account for logical errors, a standing that, without some such context, we lack. But however frequently that or some relevantly similar context obtains, the authority comes, not just from the requirement of reason, but from some other presupposed feature of the context.

logical reasoning are, in this way, fundamentally different from moral requirements. I follow Mill and a number of contemporary writers in holding that it is part of the very idea of moral obligation that moral requirements are what those to whom we are morally responsible have the authority to demand that we do.⁷ Clearly this is no part whatsoever of the concept of a demand of logic or a requirement of reason.⁸

Now, as I've said, I believe that second-personal reasons are related conceptually to, respectively, the authority to make a claim or demand, a valid or authoritative claim or demand, and responsibility or accountability *to*. To appreciate the second-personal aspect of moral responsibility, consider Strawson's famous critique of "pragmatist" or consequentialist approaches to responsibility in "Freedom and Resentment." (Strawson 1968) Strawson argued influentially that social desirability cannot provide a justification of "the right *sort*" for practices of moral responsibility "as we understand them." (1968: 74) When we seek to hold people accountable, what matters is not whether some sanction is desirable, either in a particular case or in general, but whether their actions are culpable and whether we have any authority to demand their acceptance of a sanction. Desirability is a reason of the wrong kind to warrant the attitudes and actions in which holding someone responsible consists *in their own terms*.

Strawson's point is an instance of the *wrong kind of reason problem*. For example, there might be pragmatic reasons to believe (or perhaps just to want to believe) some proposition, but that doesn't make that proposition *credible*. It doesn't justify believing it in terms of reasons and standards that distinctively apply to belief. Similarly, as D'Arms and Jacobson have pointed out, it is a "moralistic fallacy" to conclude from the fact that being amused by a certain joke is morally objectionable that the joke is therefore not itself funny.⁹ (D'Arms and Jacobson 2000)

To be a reason of the right kind, a consideration must justify the relevant attitude in its own terms. It must be a fact about or feature of some object, appropriate consideration

⁷ "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. This seems the real turning point of the distinction between morality and simple expediency. It is a part of the notion of Duty in every one of its forms, that a person may rightfully be compelled to fulfil it. Duty is a thing which may be exacted from a person, as one exacts a debt. (Mill 1998: Ch. V) John Skorupski points out that calling an act "morally wrong . . . amounts to blaming the agent" and maintains that the idea of moral wrong can't be understood independently of that of blameworthiness. (Skorupski 1999: 29, 142) 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." (Gibbard 1990: 42) And we can find versions of this Millian idea in other writers also. (Baier 1966; Brandt 1979; Shafer-Landau 2003)

⁸ I am indebted to Peter Graham for this point.

⁹ D'Arms and Jacobson argue that this poses a problem for response-dependent or, as they call them, "neo-sentimentalist" accounts of various evaluative and normative notions, since it shows that, say, the funny can't be understood in terms of amusement's making sense or being warranted by just *any* reasons. There is a distinction between an emotion or attitude's being "the right way to feel" and it's "getting [the relevant value] right." For an excellent discussion of how what they call "fitting-attitude" (or "FA") analyses can deal with the problem of distinguishing reasons of the right from reasons of the wrong kind, see Rabinowicz and Ronn ow-Rasmussen 2004. (See also Olson 2004) I am indebted to Julian Darwall for discussion of this general issue and to Joe Mendola for a question that helped me to see that Strawson's point is an instance of it.

of which could provide someone's reason for a warranted attitude of that kind towards it.¹⁰ It must be something on the basis of which someone could (and appropriately would) come to hold the attitude as a conclusion of a process of considering (deliberating about) *whether* to do so. In considering whether to believe some proposition *p*, for example, it is simply impossible to conclude one's deliberation in a belief that *p* by reflecting on the desirable consequences of believing *p*. That is a reason of the right kind for *desiring* to believe that *p*, but not for believing that *p*.¹¹ The *desirable* concerns norms and reasons that are specific to desire, and the *credible* concerns norms and reasons that are specific to belief.

Similarly, the (*morally*) *responsible* and the *culpable* concern norms for the distinctive attitudes and actions that are involved in holding people responsible and blaming them. The desirability—whether moral, social, personal, or otherwise—of holding someone responsible or blaming her, or reasons why that would be desirable, are simply reasons of the wrong kind to warrant doing so in the sense that is relevant to whether she *is* morally responsible or blameworthy. The former concerns reasons and norms of desire (even if from the moral point of view), and what is thus desirable is simply a different question from whether we are justified in holding someone responsible or blaming her in the relevant sense. The latter concerns reasons and norms that are distinctively relevant to these latter attitudes.

Strawson dubbed the distinctive attitudes involved in holding people responsible “reactive attitudes,” with prominent examples being indignation, resentment, guilt, blame, and so on. And Strawson himself pointed out what more recent commentators, notably Gary Watson and Jay Wallace, have since also noticed, namely, that reactive attitudes implicitly address *demands*. They involve “an *expectation of*, and *demand for*” certain conduct from one another.¹² (1968: 85, emphasis added) To feel a reactive attitude is to feel as though one has a warranted expectation *of* someone. Reactive attitudes, and actions that express them, must therefore presuppose the authority *to* expect and hold one another responsible for compliance with moral obligations (which must then be standards to which we can warrantably hold each other as members of the moral community). Strawson claimed also that holding one another responsible with reactive attitudes presupposes that those we hold responsible have that standing as well. In holding someone accountable in this way, we “view him as a member of the moral community; only as one who has offended against its demands.” (Strawson 1968: 93) In these ways, reactive attitudes can mediate *mutual* accountability.

It follows on Strawson's analysis that culpability and moral responsibility are second-personal phenomena in my sense. They implicitly involve the address of authoritative claims and demands and, therefore, of second-personal reasons. Moral responsibility is responsibility *to*—in Strawson's view, to the moral community. It concerns what the moral community can warrantably expect and demand of each other. When we hold one another

¹⁰ Rabinowicz and Ronnøw-Rasmussen put essentially the same point by saying reasons of the right kind also appear in the content of the attitude for which they are reasons: the attitude is toward something “on account of” these reasons. (Rabinowicz and Ronnøw-Rasmussen 2004: 414). As W. D. Falk pointed out, a favoring that is relevant to value is “by way of true comprehension of what [the object] is like.” (Falk 1986: 117)

¹¹ More accurately, it entails that reasons of the right kind for desiring to believe the proposition exist.

¹² Gary Watson stresses this in Watson 1987: 263,264. Note also, R. Jay Wallace: “there is an essential connection between the reactive attitudes and a distinctive form of evaluation . . . that I refer to as holding a person to an expectation (or demand).” (Wallace 1994: 19) See also Bennett 1980 and Scanlon (1998: 272-290).

responsible, we express these demands and presuppose the authority to do so. So if Strawson is right, as I think he is, then responsibility *to* is conceptually related to the idea of authoritative demands and, therefore, to practical authority and second-personal reasons, that is, to reasons for acting that are thought to consist in or to derive from authoritative demands.

But the conceptual relations also seem to run in the opposite direction. If one person has practical authority with respect to another, then this would seem to mean, not just that the latter has a reason of whatever priority or weight—whether overriding, silencing, exclusionary or pre-emptive—for acting as the former directs, but also that the latter has some responsibility *to* the former for doing so, that the latter is, in some way or other, answerable to the former. This point can be illustrated by altering our earlier thought experiment to imagine someone, call him now B, who accepts a mandatory norm requiring him always to do what someone in a certain position tells him to do. Suppose that A is now in that position. B will then accept that he must do what A tells him to do. We can imagine that B thereby accepts a reason for complying with A's directives of whatever weight or priority. B might think that the reason always overrides, silences, or pre-empts any potentially conflicting reasons. But B could still deny, consistently with all of that, that he is in any way responsible or answerable to A. Were he to do so, he would seem to be questioning A's authority over him. "Just because I always have overriding or pre-emptive reason to do what you say doesn't make me answerable to you or give you authority over me," he might consistently say. A cannot have authority with respect to B, it seems, unless B is in some way answerable to A.

We might put this point by saying that practical authority is not just a relation in the logical sense; it is a standing in a relationship.¹³ If A has authority with respect to B, then certain things follow about how A and B may and must relate to one another. If A has the authority to demand that B stay off A's feet, then not only is it the case that B has a reason of whatever weight or priority for doing so on the occasion of such a directive. B is also answerable to A for doing so.

This point is illustrated further by early modern theological voluntarist moral theories, like those of Pufendorf and Locke. The fundamental premise of such theories is that God has superior authority over his creatures. It is absolutely central to the voluntarist picture, moreover, that, because God has authority over us, not only must we do what God says, but we are also accountable to him for doing so. Unless we were responsible to God for compliance, any sanction God might attach to violating his commands would amount to coercion by threat. Of course, it might be within God's authority so to threaten us, but voluntarists like Pufendorf thought it critical to distinguish between even such justified threats and God's authority to hold us accountable with sanctions. For us genuinely to accept God's authority, Pufendorf thought, we must see ourselves as responsible to God for doing as he directs.¹⁴ (As we shall see, Raz also distinguishes between direction by justified threats and by authoritative directives.)

¹³ I have been helped here by discussion with Jules Coleman.

¹⁴ For discussion see Darwall 2003.

I take it, then, that there is a strong *prima facie* case for thinking that the four second-personal notions I mentioned at the outset—practical authority (to claim or demand), valid (authoritative) claim or demand, responsibility *to*, and second-personal reason for acting—are interdefinable in something like the following fashion.

Practical authority: Someone has practical authority with respect to another if, and only if, the latter has a second-personal reason to comply with the former's valid claims and demands and is responsible to the former for so doing.

Responsibility to: Someone is responsible to another if, and only if the latter has the authority to make some valid claim or demand of the former that the former is thereby given a second-personal reason to comply with.

Valid claim or demand: A valid claim or demand is one that is within the authority of someone having practical authority with respect to another to make of the latter and that the latter thereby has a second-personal reason to comply with and some responsibility to the former for so doing.

Second-personal reason: A second-personal reason is one consisting in or deriving from some valid claim or demand of someone having practical authority with respect to the agent and with which the agent is thereby accountable for complying.

Presently, we shall consider the challenge that Raz's *normal justification thesis* poses to my claim that these concepts are irreducibly second-personal. First, however, I want briefly to summarize the ways in which, as I see it, the moral categories that I mentioned at the beginning all involve these four interdefinable notions. Perhaps the clearest case is that of a moral *right*. Feinberg put the point best when he said that "it is claiming that gives rights their special moral significance." (Feinberg 1980: 155) Without rights, others might still accept norms that require treating us in whatever ways we have a right to be treated, but we could not claim this is as our right. We would have no guaranteed standing in relating to one another, no authority to claim or demand anything of each other.

We have already seen how claims of culpability and *moral responsibility* are second-personal. Holding someone responsible and blaming her involves relating to her in a way that presupposes an authority to make demands of her and that she thereby has a reason to comply with these demands. But if this is right, and if *moral obligation* is conceptually connected to moral responsibility—that is, if violations of (all-things-considered) moral obligations are blameworthy lacking adequate excuse—then it follows that moral obligation is a second-personal concept also. What we are morally obligated to do is what members of the moral community have the authority to hold us to.¹⁵ Of course, wrong-making features of action will include facts that are not themselves second-personal, such as that stepping on someone's foot causes avoidable pain, suffering, and inconvenience. But the fact that an action violates a(n all-things-considered) obligation, and so is morally wrong, is an additional

¹⁵ Depending on what we take the reference of 'moral community' to be, this may go farther than what is guaranteed by the concept of moral obligation (viz., that moral obligations are what those to whom we are morally responsible have the authority to demand that we do). According to theological voluntarists, we are morally responsible, not to other moral subjects, but to God.

reason, a second-personal reason, for not performing it, since it includes the fact that those to whom we are morally responsible have the authority to demand that we not so act.

Membership in the moral community is thus not just being subject to moral demands, but also having the standing to hold one another (and ourselves) responsible for complying with them. I believe this to be a significant aspect of the *dignity of persons*, as was illustrated in our earlier thought experiment of an individual who is committed to forbear acting in various ways that, as it happens, persons claim an authority to demand he not act, but who also recognizes no such authority. Such an individual fails to have respect for persons in some clear sense. It follows that both the dignity of persons and respect for this dignity are irreducibly second-personal. (Darwall 2004) To be a person is, as Rawls put it, to be a self-originating sourc[e] of valid claims.” (Rawls 1980: 546) Or, in Kant’s words in *The Metaphysics of Morals*, a person’s dignity is that “by which he exacts respect [that is, as Kant also says, “demands” it] for himself from all other rational beings.” (Kant 1996: 434-435)

Finally, I take the concept of *moral agent* or *person* to be a second-personal notion also itself. Locke makes the same point when he says that ‘person’ is a “forensic term.” (Locke 1975: 346) To be a person is to be apt for moral accountability, a being to whom second-personal reasons can intelligibly be addressed. But what capacities does this involve? This is the same question that arises in Strawson’s “Freedom and Resentment” of how to draw the line between beings whom we appropriately regard from an “objective” perspective as objects of self-protective coercive measures that cannot count as genuinely holding them responsible, however justified these might be, on the one hand, and those who are warranted objects of reactive attitudes and eligible to participate fully in mutually accountable human relations, on the other. Implicit in Strawson’s analysis is the following guiding thought. We can only intelligibly hold responsible those we regard as capable of holding themselves responsible, that is, people who are able to take a second-person standpoint on themselves, acknowledge the validity of moral demands and the authority of the moral community, and, in effect, address the demands to themselves through self-reactive attitudes like guilt. In a word, to be a person, and so apt for accountability, is to have *second-personal competence*.

The emergent conception of the *moral community* is one of *mutually accountable* moral agents or persons, where the second-personal competence necessary for being intelligibly held morally responsible, and so a moral subject, is also sufficient for the standing or second-personal authority to hold others and oneself morally accountable.

I claim that the four interdefinable notions of practical authority, valid claim or demand, responsibility *to*, and second-personal reason are all irreducibly second-personal. If this is so, then, to the extent that the moral categories I have just mentioned involve these notions, it will follow that they all have an irreducibly second-personal aspect also. In my view, this is a fact of great importance for moral theory. It means that no set of propositions that do not already involve the interdefinable second-personal notions I have identified can possibly entail any proposition of moral right, responsibility, obligation, and so on. To paraphrase a slogan of Bernard Williams’s: Second-personal authority out, second-personal authority in.¹⁶ (Williams 1985: 181) Of course, it may still be true that, for example, we are

¹⁶ Williams’s slogan was “Obligation out, obligation in.” One way of putting the thesis that moral obligation must be understood in terms of second-personal reasons is to say that the formulation in the text is not just a

morally obligated to maximize overall utility, but this cannot possibly follow simply from the fact that the world would be better if we did. There must be some showing that this is something we reasonably demand of one another.

But what makes it the case that someone has practical authority of any kind (as I contend, for example, that beings with second-personal competence have authority as equal members of the moral community)? And what makes it the case that any claim or demand from someone to another is valid, one that that person has the authority to make on the other? If it is possible to justify practical authority fully in terms of reasons that are not themselves second-personal, then perhaps the notions I have identified are not *irreducibly* second-personal. Maybe their apparently second-personal character is a superficial aspect that can be explained away or otherwise accounted for at some more fundamental level.

This is the challenge that is posed by Raz's *normal justification thesis* which Raz formulates as follows:

the normal way to establish that a person has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him (other than the alleged authoritative directives) if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly. (1986: 53)

Now the reasons that already apply to an "alleged subject" independently of the second-personal reasons putatively generated by some *specific* alleged authority may themselves be partly second-personal, since they may be conceptually related to claims and demands of some other authority. This will turn out to be important if reasons of moral obligation are themselves second-personal in the way I have claimed. Of course, if the normal justification thesis is correct, then it should be possible to establish every authority, including whatever authority we have as equal members of the moral community, by reference to independently existing reasons. So it ought to be true that whatever standing we have to hold one another answerable for moral demands, like any other practical authority, can be accounted for by virtue of its being the case that we would do better, were we to accept this authority, at complying with reasons that apply to us independently of this authority. And if we could account for all practical authority by the normal justification thesis, then it would seem to follow that practical authority and the other notions that are conceptually related to it, although second-personal on their face, are not irreducibly second-personal.

Meeting the standards of the normal justification thesis is not, however, sufficient to establish practical authority. There are cases where one person might very well do better to follow someone else's directives where it seems clear that the latter has no claim whatsoever on the former's will and actions and consequently no practical authority with respect to him. And cases where an "alleged subject" would do better in complying with independent reasons where genuine authority *does* seem to be involved all also seem to involve some assumed background accountability relation that gives the authority's

paraphrase of Williams's slogan, but that the point it formulates is the genus of which that formulated by Williams's slogan is a species.

directives standing as second-personal reasons. In these cases, it is the latter that establishes the directives' authority, not the former.

It is worth pointing out that much of what Raz says about practical authority in *The Morality of Freedom* aside from the *normal justification thesis* seems well attuned to practical authority's second-personal character. First, Raz notes that authority of this kind is different, not only from power and coercive threats, but also from *justified* power and coercion. One does not exercise authority over "people afflicted with dangerous diseases," Raz notes, "if [one] knock[s] them out and locks them up to protect the public" even if one "is justified in doing so." (Raz 1986: 25) Such a justified use of coercion would miss, Raz says, an essential element of any genuine exercise of authority, namely, "an appeal for compliance" and "an invocation of the duty to obey." (25-26) This seems exactly right. Such an appeal is an instance of what I am calling second-personal address. It is an appeal *to* the alleged subjects to recognize the alleged authority and therefore to comply with directives that are authorized by it. Moreover, someone can credibly make such an appeal only if he can expect his alleged subject to accept that the subject has some duty or obligation to follow his directives. Without such a duty or obligation in place, which an alleged authority cannot of course create by his own directives, no genuine authority exists.

Second, Raz makes an important distinction between the idea of a kind of authority that can give one new reasons for *belief*, including beliefs about what there is reason to do, on the one hand, and the idea of genuine "practical authority" (as he also calls it), which can give one reasons for *acting* that one did not previously have, on the other. Raz rightly rejects a "recognitional conception" of practical authority according to which an authority gives agents no new reasons for acting, just new reasons to believe what reasons for acting they already have. Applying the "recognitional conception" to authority *überhaupt* gives rise to what Raz calls the "no difference thesis": "the view that authority does not change people's reasons for acting." (Raz 1980: 30) Raz points out that one way of seeing that the recognitional conception and no difference thesis cannot be correct is to reflect on the role that practical authority can play in solving coordination problems. There are cases where it is important for everyone to coordinate on one of two options that, so far as antecedently applicable reasons go, are equally choiceworthy. It's fine if everyone drives on the right, or on the left, but everyone had better coordinate on one side or the other. "A wise man," Raz notes, "can tell me which options belong" to a set of optimal choices, "but he cannot tell me which option to choose before it is known what others will do." (Raz 1980: 30) This is, of course, correct. Raz rightly distinguishes between *practical* authority and the kind of authority that a trusted advisor can have in directing us to independently choiceworthy options. There is, as Hobbes famously noted, a distinction between "counsel" and "law" or "command." (Hobbes 1983: XIV.1) Genuine practical authority can give those who are subject to it new reasons for acting by addressing valid claims or demands to those subjects. As I would put it, the relevant reasons are second-personal reasons; their validity depends on the alleged authority's actually having the authority he presupposes when he addresses the demand (or "appeal") for compliance.

Nonetheless, one might still think that what grounds practical authority and so makes it the case that someone can give reasons of this distinctive, second-personal kind are

further reasons that are not themselves second-personal. This brings us back to the *normal justification thesis*. The basic idea, again, is that someone earns practical authority with respect to another person if the latter would do better in complying with already applicable reasons were she to accept the former's directives "as authoritatively binding and tr[y] to act on them" than she would if she did not and tried to act on her own assessment of reasons. There is, however, a problem we face right at the outset in interpreting the normal justification thesis. What is it to accept someone's "directives as authoritatively binding"?

If we understand accepting a directive as authoritatively binding to be accepting it as a valid demand in the second-personal sense I have been pointing to, then the alleged subject's own normative thought must then be interpreted in second-personal terms. In particular, we must suppose that she not only takes herself to have a reason of whatever weight or priority (exclusionary or pre-emptive, say), but that she also sees herself as *responsible* for compliance, that is, as answerable to the alleged authority. We must suppose that she, at least, takes it that the fact that she is thus responsible is in itself a reason for her to comply of this distinctively second-personal kind. But if this is so, then a version of Strawson's "wrong kind of reason" problem will arise from her perspective. Just as the fact that there might be desirable effects of treating people as responsible and their actions as culpable is a reason "of the wrong sort" for practices of holding people responsible "as we understand them," so also will the fact that an alleged subject would do better in complying with independently applicable reasons seem to her to be a reason of the wrong kind to think that she really is responsible for complying with some (alleged) authority's directives and therefore that his putative practical authority is genuine.

If this is right, moreover, the beneficial effects of the subject's acceptance of an alleged authority may themselves be hostage to the subject's not also accepting the normal justification thesis herself, since the putative authority she accepts may then be susceptible, from her perspective at least, to "the dissolving force of [the] analysis" that the normal justification thesis purports to provide.¹⁷ Were she to think that the only legitimacy the "alleged authority" can claim is that she would do better if she were to accept it, then this may tend to undermine her acceptance of it, and so the beneficial effects.

It seems more accurate to Raz's idea, however, to suppose that by accepting an alleged authority's "directives as authoritatively binding," Raz means nothing essentially second-personal, but simply that the alleged subject takes the authority's directive as pre-emptive reasons, that is, reasons that are "not to be added to all other relevant reasons when assessing what to do," but that "exclude and take the place of [at least] some of them." (Raz 1986: 46) Thus A acquires practical authority with respect to B if B would do better in

¹⁷ The reference is to Mill in a different context: "But moral associations which are wholly of artificial creation, when intellectual culture goes on, yield by degrees to the dissolving force of analysis: and if the feeling of duty, when associated with utility, would appear equally arbitrary; if there were no leading department of our nature, no powerful class of sentiments, with which that association would harmonize, which would make us feel it congenial, and incline us not only to foster it in others (for which we have abundant interested motives), but also to cherish it in ourselves—if there were not, in short, a natural basis of sentiment for utilitarian morality, it might well happen that this association also, even after it had been implanted by education, might be analyzed away." (1968, Ch. III, §9)

actually complying with independently applicable reasons if B were to treat A's directives as pre-emptive reasons in this sense.

It is the pre-emptive character of the reasons that practical authority purports to create, according to Raz's analysis, that enables Raz to distinguish between genuinely practical authority and the kind of authority a trusted advisor might have, that is, to make Hobbes's distinction between command and counsel. The only claim an advisor makes as such is on an advisee's beliefs about independently existing reasons and about what actions these reasons support, not on her will. "This is what I think you have good reason to do," an advisor might say, and add, "but I'm not for a moment telling or even asking you to do it," without canceling his advice or anything the advice implied. The improvements to the alleged subject's reason compliance that matter for the normal justification thesis are those that come, not from her being convinced by the alleged authority that there is sufficient independently existing reason to do as he suggests, but from her taking his instructions as *directives* that create *pre-emptive* reasons.

Now I take it that when Raz says that it is sufficient to establish that A has practical authority with respect to B if it can be shown that B would do better in complying with independently applicable reasons were B to accept A's directives as creating pre-emptive reasons, he means this to entail also that under such conditions A's directives actually do create such pre-emptive reasons. It is important to see, however, that these two theses can be distinguished. In fact, all three of the following theses are different:

- I. If B would do better in complying with independently existing reasons were B to treat A's directives as pre-emptive reasons, then B has sufficient reason so to treat A's directives.
- II. If B would do better in complying with independently existing reasons were B to treat A's directives as pre-emptive reasons, then A's directives actually are such pre-emptive reasons for B.
- III. If B would do better in complying with independently existing reasons were B to treat A's directives as pre-emptive reasons, then A has authority with respect to B. (normal justification thesis)

It seems obvious that one could accept I without accepting II and accept both I and II without accepting III. Indeed, one could accept I and II without accepting the normal justification thesis even if one thought that A's having authority with respect to B entails that A's directives are pre-emptive reasons.

For our purposes, however, we can simply put aside any reservations we might have about I and II and simply stipulate that both are true. Our question will still remain, namely, whether the satisfaction of the conditions of the normal justification thesis establishes, not just that the directives of an alleged authority should be taken as pre-emptive reasons or that they are pre-emptive reasons in fact, but that that makes it true that the alleged authority is genuine.¹⁸

¹⁸ Of course, Raz might hold that A's having authority over B is no more and no less than that A's directives to B give B pre-emptive reasons. I consider this possibility below.

There seem to be clear cases, however, some that Raz himself considers, where it is hard to see how the normal justification thesis could possibly hold. Raz asks us to imagine that “John is an expert on Chinese Cooking” and that if one wants nothing but “to prepare the best Chinese meal [one] can [one] should just follow John’s instructions.” (64) Let us stipulate also that one has no reason to do anything other than prepare the best Chinese meal and therefore, following from our acceptance of II above, that John’s instructions provide pre-emptive reasons. The question remains, does John thereby acquire any practical authority over one? It is hard to see how he does. Of course, one would be foolish not to follow his instructions, but if one didn’t it is difficult to see why John would have any standing to complain or otherwise hold one to account. Raz says that those with practical authority “have the right to replace people’s own judgment on the merits of the case.” But what *right* could John have in such a case? Or recall Raz’s remark that, unlike merely justified coercion, genuine practical authority involves “an appeal for compliance” and “an invocation of the duty to obey.” (Raz 1980: 25-26) How, in such a case, could John warrantably expect that one would have any obligation to follow his instructions, however foolish one might be not to do so? It follows from our stipulations that John’s instructions are not mere advice; they provide pre-emptive reasons. But it is hard to see how that gives John any right to our compliance with his directives or us any obligation to comply with them.

Raz’s position about cases of this kind is somewhat equivocal. He grants that we would not ordinarily think that John would have any authority “over” one even if one’s Chinese cooking would “prosper” by following his instructions, but points out that one can accept the normal justification thesis and still resist that John has such authority because one might think one would enjoy self-directed Chinese cooking more or have other goals. No doubt, but we can simply stipulate that one has no reason to do anything other than prepare the best Chinese meal, and it still doesn’t seem to be the case that John has genuine practical authority as the normal justification thesis requires. (64-65)

However, what if the reasons with which one would better comply themselves concern moral obligations? In such a case, an alleged authority would more plausibly appeal to an obligation to do as he directs. If the only way we can adequately comply with our moral obligations is to treat an alleged authority’s directives as pre-emptive reasons, then there seems to be a sense in which it is plausible to suppose that we would be under an obligation so to treat them. Even so, it wouldn’t follow from this that the alleged authority himself thereby acquires any authority (beyond any he might have had already) to hold others to moral demands. After all, treating some appropriately programmed computer’s “directives” as pre-emptive reasons *might* be no less effective, but neither the computer nor the programmer would thereby acquire any special practical authority by virtue of that.

Even if we could extend the scope of authority to an effective moral director by applying the normal justification thesis in such a case, this would still give us no reason to conclude that the second-personal appearances of practical authority can ultimately be explained away entirely by the normal justification thesis. The reason is that moral obligations provide second-personal reasons themselves. Moral obligations are what is warrantably demanded of us, what we can warrantably be held to. In the kind of case we are imagining, it is, at best, as if the alleged authority functions as a kind of representative who acquires any special authority he comes to have by virtue of the general authority to hold

morally responsible. So even if the normal justification thesis tracks genuine practical authority in a case of this kind, it does so thanks to there being an assumed second-personal authority to hold people to moral demands in the background.

But suppose we attempt to explain the authority to hold people to moral demands itself by the normal justification thesis. If we try to do so based on improvements that would come in our compliance with reasons of other kinds then our moral philosophy will “rest on a mistake,” in Prichard’s famous dictum. (Prichard 2002) Reasons of other sorts are reasons of the wrong kind for distinctively moral authority. But what if we say that moral authority derives from improvements that accepting it would make in our compliance with moral obligations themselves? How then are we to specify our moral obligations? If Mill and I are right, there is no way adequately to characterize what it is to be morally obligatory that does not already involve the authority to hold morally responsible. So we can’t explain moral authority itself by the normal justification thesis by claiming that we will do better in complying with moral obligations if we accept that authority, since nothing will count as a moral obligation unless there is already such a thing as the authority to hold morally responsible. I need not assume, of course, that it is impossible to designate moral obligations extensionally in non-second-personal terms, for example, with some list like: acts of kind A in circumstances of kind C₁, of kind B in circumstances of kind C₂, and so on. The point will nonetheless remain that the mere fact that we will do better in performing acts on the list is impotent by itself to explain the authority to hold morally responsible, since an act’s being morally obligatory, and so on the list, is not itself independent of the authority to hold morally responsible and make demands of one another.

Suppose, however, that Raz were to eschew talk of obligation in this sense “all the way down,” and simply take the position that all he really means by a duty of obedience is that there are pre-emptive reasons for following an alleged authority’s directives. Similarly, he might hold that the latter is all that it is for someone to have practical authority in the sense in which he has in mind. If we interpret the normal justification thesis as applying to practical authority defined in this way, it seems much more plausible. But it should be clear that practical authority so defined is not a thesis that entails anything about any *right* to obedience or about any *obligation* to obey, at least as we ordinarily understand rights and obligations. So understood, the normal justification thesis is simply a thesis about pre-emptive reasons. But if that is so, then it is not even relevant to the question of whether second-personal authority, demands, reasons, and responsibility, in the sense that I have delimited, can ultimately be accounted for in terms of considerations that are not themselves second-personal.

My conclusion is that the normal justification thesis poses no serious challenge to my claim that the four notions I have identified constitute a circle of interdefinable, irreducibly second-personal concepts. If we interpret the thesis as concerning practical authority, understood in second-personal terms, then it is either implausible or, in cases where it seems not to be, it borrows plausibility from the implicitly second-personal character of the reasons on which the alleged authority’s directives are based (as when she directs us to comply with our moral obligations). And if we interpret the thesis otherwise, it just becomes a thesis about pre-emptive reasons for acting and so has no direct relevance to the possibility of grounding or otherwise explaining second-personal authority in non-second-personal terms.

Of course, even if the concepts of practical authority, valid demand, second-personal reason, and responsibility *to*, are irreducibly second-personal in the ways I am claiming, that doesn't entail that any of these concepts are instantiated, that is, that there are any second-personal reasons, any second-personal authority, and so on. So we might regard a position like Raz's as a kind of skepticism about practical authority of the kind I am pointing to and as offering a "reforming definition" in the style of Stevenson or Brandt in terms of pre-emptive reasons. I cannot, of course, establish here that the unreformed second-personal ideas I have described are anything other than what Kant calls "figment[s] of the mind."¹⁹²⁰ However, if what I have said here is correct, it will follow that if these ideas are mere figments, then so are the ideas of moral right, responsibility, and obligation and, indeed, the very concept of moral agent or person itself.

Stephen Darwall
Practical Reason Conference
Bowling Green State University
April 7-8, 2006

WORKS CITED

- Baier, Kurt (1966). "Moral Obligation," *American Philosophical Quarterly*, 3: 210-226.
- Bennett, Jonathan (1980). "Accountability," *Philosophical Subjects*, ed. Zak Van Stratten. Oxford: Clarendon Press.
- Bond, E. J. (1983). *Reason and Value*. Cambridge: Cambridge University Press.
- Brandt, Richard (1979). *A Theory of the Good and the Right*. Oxford: Oxford University Press.
- Dancy, Jonathan (2000). *Practical Reality*. Oxford: Oxford University Press.
- D'Arms, Justin and Daniel Jacobson (2000). "The Moralistic Fallacy: On the 'Appropriateness' of Emotions," *Philosophy and Phenomenological Research*, 61: 65-90.
- Darwall, Stephen (1983). *Impartial Reason*. Ithaca, NY: Cornell University Press.
- Darwall, Stephen (1986). "Agent-Centered Restrictions From the Inside Out," *Philosophical Studies*, 50: 291-319.

¹⁹ At the end of Chapter II *Groundwork* points out that his arguments that autonomy of the will, the unqualified goodness of the good will, and the Categorical Imperative are mutually entailing has derived from an analysis of our moral concepts, and that nothing yet follows from this about whether any of these are actually realized or valid. That, he says, requires a "critique" of practical reason, to which he turns in *Groundwork* III and, of course, *The Critique of Practical Reason*.

²⁰ I attempt to vindicate these ideas in Darwall 2006.

- Darwall, Stephen (2002). *Welfare and Rational Care*. Princeton, NJ: Princeton University Press.
- Darwall, Stephen (2003). "Autonomy in Modern Natural Law," in *New Essays on the History of Autonomy*, eds. Larry Krasnoff and Natalie Brender. Cambridge: Cambridge University Press.
- Darwall, Stephen (2004). "Respect and the Second-Person Standpoint," *Proceedings and Addresses of the American Philosophical Association*, 78: 43-59.
- Darwall, Stephen (2006). *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge: MA: Harvard University Press.
- Falk, W.D. (1986). "Fact, Value, and Nonnatural Predication," in *Ought, Reasons, and Morality*. Ithaca, NY: Cornell University Press.
- Feinberg, Joel (1980). "The Nature and Value of Rights," in *Rights, Justice, and the Bounds of Liberty*. Princeton, NJ: Princeton University Press.
- Gibbard, Allan (1990). *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Hampton, Jean (1998). *The Authority of Reason*. Cambridge: Cambridge University Press.
- Kant, Immanuel (1996f). *The Metaphysics of Morals*, in *Practical Philosophy*. References are to page numbers of the Preussische Akademie edition.
- Korsgaard, Christine (1996a). "The Reasons We Can Share: An Attack on the Distinction Between Agent-Relative and Agent-Neutral Values," in *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Locke, John (1975). *An Essay Concerning Human Understanding*, ed. Peter H. Nidditch. Oxford: Oxford University Press.
- McDowell, John (1979). "Virtue and Reason," *Monist* 62: 331-350.
- McNaughton, David and Piers Rawlings (1991). "Agent-Relativity and the Doing-Happening Distinction," *Philosophical Studies*, 63: 167-185.
- Mill, John Stuart (1998). *Utilitarianism*, ed. Roger Crisp. Oxford: Oxford University Press.
- Moore, G. E. (1993). *Principia Ethica*, rev. ed. with the preface to the (projected) second ed and other papers, ed. with an intro. Thomas Baldwin. Cambridge: Cambridge University Press.
- Nagel, Thomas (1970). *The Possibility of Altruism*. Oxford: Clarendon Press.
- Nagel, Thomas (1986). *The View From Nowhere*. New York: Oxford University Press.

- Olson, Jonas (2004). "Buck-Passing and the Wrong Kind of Reasons," *The Philosophical Quarterly*, 54: 295-300.
- Parfit, Derek (1984). *Reasons and Persons*. Oxford: Clarendon Press.
- Pettit, Philip and Michael Smith (1990). "Backgrounding Desire," *The Philosophical Review*, 99: 565-592.
- Quinn, Warren (1991). "Putting Rationality in Its Place," in *Morality and Action*. Cambridge: Cambridge University Press.
- Rabinowicz, Wlodek and Toni Rønnow-Rasmussen (2004). "The Strike of the Demon: On Fitting Pro-Attitudes and Value," *Ethics*, 114: 391-423.
- Rawls, John (1980). "Kantian Constructivism in Moral Theory," *The Journal of Philosophy*, 77: 515-572.
- Raz, Joseph (1975). *Practical Reason and Norms*. London: Hutchinson.
- Raz, Joseph (1986). *The Morality of Freedom*. Oxford: Clarendon Press.
- Scanlon, T. M. (1998). *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Scheffler, Samuel (1982). *The Rejection of Consequentialism*. Oxford: Clarendon Press.
- Shafer-Landau (2003). *Moral Realism: A Defense*. New York: Oxford University Press.
- Skorupski, John (1999). *Ethical Explorations*. Oxford: Oxford University Press.
- Strawson, P. F. (1968). "Freedom and Resentment," in *Studies in the Philosophy of Thought and Action*. London: Oxford University Press.
- Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.
- Watson, Gary (1987). "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme," in *Responsibility, Character, and the Emotions: New Essays in Moral Psychology*, ed. F. D. Schoeman. Cambridge: Cambridge University Press.
- Williams, Bernard (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.