

Second Person Rules: Resolving the Idealization Puzzle for Second-Person Normativity

Kevin Vallier

DRAFT: Do not cite without permission.

Stephen Darwall's recent work, *The Second-Person Standpoint*, has provided a valuable set of insights into the nature of what I shall call our *social-moral* practices, our practices of holding one another responsible for wrongdoing, making moral demands of others and blaming them for infractions. Darwall describes a network of normative concepts that together he takes to ground these practices, practices that are "second-personal" as they involve community members taking up a common standpoint where they hold one another accountable on common terms. Despite a flurry of criticism, I assume here that Darwall's basic insight about the nature of moral obligation is sound: for John to have bona fide moral authority over Reba, she must have some second-personal reason to recognize his authority and thereby permit him to hold her accountable for wrongdoing and blame her accordingly. If Darwall is right, we can move beyond his account of second-personality and examine how his understanding of second-person normativity provides the framework for a continuing research program in moral philosophy.

The purpose of this paper is to help develop such a framework by identifying a puzzle for Darwall's account of second-personal normativity and resolving it in favor of an alternative account. Specifically, I shall argue that Darwall's account of second-personality is torn between actual and hypothetical interpretations, that is, between different levels of *idealization*. In some writings, Darwall stresses that the second-person standpoint is tied up with real-world authorities and reactive attitudes and that obligations require some form of actual recognition by the relevant moral community. But in other writings, Darwall repeatedly insists that the appropriate form of recognition occurs merely among hypothetical members of a kingdom of ends, members we serve as representatives. I maintain that Darwall that both interpretations present serious difficulties, even if Darwall could have it both ways (which he cannot). In response, I argue that second-personality should be applied first and foremost to publicly recognized *rules of conduct*. Rules play an intermediary role between actual practices of moral address and accountability and hypothetical rational endorsement of these practices. I argue that real-world moral address, justification and accountability occurs via moral rules that are themselves the object of hypothetical endorsement.

The introduction of social-moral rules into the second-person standpoint resolves Darwall's puzzle of idealization.

Resolving the puzzle in favor of rules has an additional advantage: it makes better sense of the notion of a second-personal reason. In some articles, Darwall claims that a second-person reason is any reason issued in the relevant second-personal context. But this, I shall argue, leaves the notion of a second-person reason obscure. I suggest that second-personal reasons be understood in terms of the rules to which they are attached; a second person reason turns out to be a reason to make demands or comply with demands in accord with a justified moral rule.

My argument proceeds in six parts. In Section I, I review Darwall's understanding of second-personal normativity. Section II discusses the puzzle for idealization. In Section III, I give an account of social-moral rules that retain the conceptual features that Darwall claims are essential for second-personal normativity. In Section IV, I show how social-moral rules resolve the puzzle of idealization. Section V develops a new account of second-personal reasons. Section VI concludes.

I: Darwall's Account of Second-Personality

The second-person standpoint is intended to ground contractualist theories of right action. Contractualism proposes to explain the rightness of actions and the authority of the moral demands associated with them. Acts are wrong if the principles they violate cannot be reasonably rejected. Or as Thomas Scanlon puts it, "An act is wrong if its performance under the circumstances would be disallowed by any set of principles for the regulation of behavior that no one could reasonably reject as a basis for informed, unforced general agreement."¹ Darwall points out that contractualisms share the idea "that we are accountable for regulating ourselves by principles that are acceptable ... to each as equal moral persons."² We can appreciate this fact best when we take the second-person standpoint. Right and wrong are thereby grounded in "the perspective you and I take up when we make and acknowledge claims on one another's conduct and will."³ This standpoint "commits us to regulating our conduct by principles that are acceptable, or not reasonably rejectable, to each as free and rational agents."⁴

¹ Scanlon 1998, p. 153.

² Darwall 2006, p. 304.

³ *Ibid.*, p. 3. I take it our practice of making and acknowledging claims just is to practice interpersonal morality.

⁴ *Ibid.*, p. 300.

For Darwall, the second-person standpoint can only be explicated and defended by appealing to a “circle of irreducibly second-personal concepts.”⁵ No concept can be fully understood apart from the others and no explanation of the second-person standpoint can begin from outside of the circle. Darwall claims that there are *four* such concepts: (i) practical authority, (ii) accountability to others, (iii) valid claims or demands, and (iv) second-personal reasons.⁶

(i) For Darwall, practical authority obtains when each person recognizes that others can require certain lines of action from her via second-personal reasons. Or as Darwall puts it: “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 accountable to the former for so doing.”⁷ We acknowledge this authority because we see other people as morally related to us, as standing in relations of accountability. A relationship of authority obtains when A obligates B to φ in the appropriate context. But for Darwall, this relationship only obtains when parties recognize that an authority exists. This mutual recognition arises when second-personal reasons are taken to apply to the parties in question.

(ii) Darwall understands the idea of accountability to others in terms of P.F. Strawson’s conception of the connection between responsibility and normativity. Darwall terms “Strawson’s Point” the claim that “desirability is a reason of the wrong kind to warrant the attitudes and actions in which holding someone responsible consists in its own terms.”⁸ According to Darwall, Strawson claims that we can only hold others morally responsible for their wrongful actions when they recognize a reason not to act wrongly. The attitudes of praise and blame associated with holding others to account are “reactive attitudes” that consist in “an expectation of, and demand for” behaviors from others in a particular context.⁹ Since reactive attitudes concern what claims we can hold others to, Darwall holds that they “presuppose the authority to demand and hold one another responsible for compliance with moral obligations.”¹⁰ We cannot obligate others merely by willing that we do so, by desiring that the duty obtain. Instead, our social practices and attitudes commit us to addressing the reasoning of others. Darwall has recently explained the idea of accountability in

⁵ Ibid., p. 11. The key notions ... “comprise an interdefinable circle; each implies all the rest. Moreover ... there is no way to break into this circle from outside it.” Ibid., p. 12.

⁶ Ibid., p. 13. Also see Darwall 2010a, p. 266 and Darwall 2010c, pp. 217-8. I initially counted the ideas of equal dignity and membership in a moral community as distinct concepts, increasing the list of concepts to six. Perhaps Darwall’s network of concepts can be individuated in different ways.

⁷ Darwall 2010a p. 266.

⁸ Darwall 2006, p. 15. For the original essay, see Strawson 1974, pp. 1-28.

⁹ Strawson 1974, p. 15.

¹⁰ Darwall 2006, p. 17.

this way: “Someone is accountable 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.”¹¹ Accountability relations themselves are second-personal because they hold in virtue of the common recognition of valid claims and demands as grounded by second-personal reasons.

(iii) To make a valid claim or demand on the conduct and will of others, several preconditions must be met. First, those making claims must take the second-person standpoint and offer second-personal reasons. But claim-makers acquire their second-personal authority by acknowledging—at least implicitly—a claim of reciprocity, that others can make claims on our conduct and will just as we can make the same claims on theirs. As Darwall argues, there exists “a form of reciprocal respect that is built into all second-personal reason-giving....”¹² Further, valid claims and demands occur in terms of legitimate forms of address. A legitimate form of address is one that can obligate. To give an example, consider John Locke’s famous discussion of the logic of persecutors as a case of an inappropriate form of address. Suppose John demands that Reba, a member of another church, not practice her religion openly. The reason John provides Reba is that he affirms the true religion and that Reba does not. When Locke reminds his reader that *every church is orthodox to itself* he is arguing that John’s claim that he affirms the true religion is the wrong kind of address.¹³ In this case, the demand fails to be grounded in a second-personal reason, a reason whose force the addressee has reason to acknowledge. Darwall defines valid claims or demands in this way: “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 accountability to the former for so doing.”¹⁴ In Locke’s case, John lacks authority because he fails to issue a second-personal reason to Reba, and instead insists that she comply because he is right and she is wrong.

(iv) The second-person standpoint generates duties and justifies moral demands via *second-personal reasons*. An essential feature of second-personal reasons is that they are “agent-relative” because they make a “non-trivial and uneliminable pronominal back-reference” to the person who has the reason in question.¹⁵ This means that agent-relative reasons are reasons that are essentially

¹¹ Darwall 2010a p. 266.

¹² Darwall 2006, p. 21.

¹³ Locke 2003, p. 225.

¹⁴ Darwall 2010a p. 266.

¹⁵ Ridge 2003, p. 339.

“reasons for” a particular agent. The agent-relativity of second-personal reasons generates their recognitional component. For John to obligate Reba, she must recognize his demand as having moral force. She must do so through her appreciation of the reasons that apply to her or that are relative to her. Unless these reasons can be so recognized, they fail to obligate. Second-personal reasons are thus agent-relative reasons that have the power obligate those to whom they are addressed. To put it another way, second-person reasons obligate others because their soundness as reasons “depends on presupposed authority and accountability relations between persons and, therefore, on the possibility of the reason’s being addressed person-to-person.”¹⁶ Without presupposed authority and accountability relations, second-personal reasons cannot generate obligations. Thus, Darwall understands a second-personal reason to derive “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.”¹⁷ Second personal reasons are second-personal because they derive “from a legitimate demand, and hence, from the authority to issue the demand.”¹⁸ Or as Darwall puts it elsewhere, “Because the notion of an *addressable* demand is itself part of this further reason, I call it a *second-personal reason*.”¹⁹

In sum, Darwall’s four concepts are meant to explain our distinctively second-personal moral practices, which in turn explain our obligations to one another and the conditions under which we can hold one another responsible for not complying with those obligations. There is more to be said, but for now the position should be clear enough to discern an important puzzle that troubles this account of second-person normativity.

II: The Puzzle of Idealization

A number of prominent moral philosophers have addressed Darwall’s views at length.²⁰ One common criticism is that Darwall’s view seems too “actualist” in that it looks as if generating obligations requires real, discrete acts on the part of members of the moral community. This is to say that persons’ obligations are too tightly tied to their real-world behavior. Consider a representative case from R. Jay Wallace; Wallace argues that Darwall’s example of a second-personal

¹⁶ Ibid., p. 8.

¹⁷ Darwall 2010a p. 266.

¹⁸ Darwall 2010d, p. 260.

¹⁹ Darwall 2010b, p. 151.

²⁰ See, for instance, the symposia in *Ethics* and *Philosophy and Phenomenological Research* cited in this article.

reason issuing from stepping on someone else's toe "suggests that you did not have a second-personal reason to refrain from stepping on the victim's toe until the protest was issued."²¹ Even if Darwall responds that the "demand is not explicitly addressed by the person whose toe you step on, it is present in the disposition of that person" Wallace complains that Darwall has illegitimately expanded the commonsense idea of "what it is to address a demand to a person."²² Wallace also insists that Darwall's account of obligation makes "moral obligation hostile to the actual responses of the individuals implicated in interactions with each other, in ways that are problematic."²³ Gary Watson has expressed a similar worry: "Darwall needs a distinction between 'demands' that remain in force whether or not one issues them and demands that do not."²⁴ Watson realizes that the difference must appeal "to certain implicit understandings and hypothetical conditions" but he worries that "it is never well worked out."²⁵ Michael Smith and Jada Strabbing express the same concern. Interpreting Darwall in this actualist manner leads them to wonder whether "a third party who makes ... evidence available thereby make[s] available a second-personal reason to get off your foot that existed all along" or whether the third party creates the reason.²⁶

In response, Darwall repeatedly insists that his account of obligation does not rest on the behaviors of an actual second-person community, but rather on a hypothetical community. For instance, Darwall writes that when we hold others accountable for violating her obligations and blame her accordingly, our demands are not issued as individuals but "*as representatives of the moral community* (not, of course, any actual collectivity ... but a regulative ideal like Kant's Realm of Ends)."²⁷ Darwall claims that our behaviors are authoritative merely because "we give voice to demands that could come from any of us and that we presume to come from all of us."²⁸ This passage suggests that real-world actions are still authoritative, but merely in virtue of representing the conclusions of an ideal human community. But in other cases Darwall severs the link between actual persons and hypothetical persons almost entirely. For instance, Darwall claims that "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*."²⁹ In *SPS*,

²¹ Wallace 2007, p. 26.

²² *Ibid.*, p. 27.

²³ *Ibid.*

²⁴ Watson *ibid.*, p. 39.

²⁵ *Ibid.*, p. 40.

²⁶ Smith and Strabbing 2010, p. 239, ft.2.

²⁷ Darwall *ibid.*, p. 223. Emphasis author's.

²⁸ *Ibid.*

²⁹ Darwall 2010b p. 146, ft. 36. Emphasis author's.

Darwall claims that “We might think of moral demands as being ‘in force’ if members of ‘the moral community’ are prone to make them. But the moral community as I understand it is not any actual community composed of actual human beings.”³⁰ Similarly, and more radically, “It takes neither an explicit actual demand *nor a demand that is implicit in actual human beings* prone to make it, either individually or collective in order for a claim or demand to be in force.”³¹

Now if Darwall emphasizes that his account involves hypothetical agreement and reason-giving, why would so many different authors mistake his point? Did Darwall simply forget to emphasize the hypothetical elements in his account? Or are his critics picking up on some deeper issue? From what I can tell, various features of Darwall’s account suggest an actualist reading, but I will focus on four: (i) Darwall’s insistence that second-personal reasons depend on real-world second-personal authorities, (ii) his claim that second-personal reasons depend on real-world second-personal competence, (iii) the connection Darwall draws between obligation and the reactive attitudes and (iv) his insistence on the publicity of second-person norms.

Let’s begin with (i) and (ii). Darwall claims that “Second-personal reasons exist only if the relevant authorities also exist to address and acknowledge claims and demands.”³² Darwall therefore appears to make the very possibility of obligation depend on the existence of actual persons who have the authority to address real world others. Darwall also regularly argues that practical authority is something that actual persons have over one another and that real obligations depend on the possibility of real-world second-personal address. Again: “What makes the reason second personal ... is its conceptual connection to an authority to make claims and demands of one another (second personally).”³³ Further, “What makes a reason second-personal ... is that it derives from a legitimate demand and, hence, from the authority to issue the demand.”³⁴ Both passages seem to imply that second-personal reasons can only *exist* if certain real-world conditions are met, specifically that there are real-world authorities and legitimate demands. In fact, Darwall’s actualist “face” seems most prominent in his most well-known statements of his position; after all, the second-person standpoint is one that “you and I” take up together in order to address one another as persons with equal dignity.

³⁰ Darwall 2007, p. 64.

³¹ *Ibid.*, p. 65. Emphasis mine.

³² Darwall 2010a, p. 263.

³³ Darwall 2007, p. 61.

³⁴ Darwall 2010d, p. 260.

Moving to (iii), Darwall's emphasis on the reactive attitudes requires that reactive attitudes somehow license or make sense of second-person reason-giving and the formation of obligations. The reactive attitudes themselves presuppose that the relevant second-person authorities are in place and that the relevant second-personal practices obtain. But it is hard to see how the reactive attitudes presuppose anything *hypothetical*. We feel reactive attitudes in real-world cases of moral address (including in our role as third-parties) because we are in real-world situations. How can what a hypothetical member of the moral community recognizes as appropriate address license our real-world reactive attitudes? Reactive attitudes are, after all, reactions to observed or imagined events. Further, reactive attitudes are supposed to permit us to hold one another accountable for our actual behaviors. Reactive attitudes consist in "an expectation of, and demand for" behaviors from others in real-world contexts.³⁵ And yet Darwall ties the reactive attitudes to a hypothetical moral community when he requests that we "understand the moral community as being prone to the reactive attitudes in a contractualist way."³⁶ Reactive attitudes are essential for second-personal activities, and they have their home in our actual emotional systems.

Finally, let us focus on (iv): Darwall's writings lend themselves to an actualist reading when he argues that moral principles should be public. As Darwall notes, "Standards of right and wrong are, in their nature, bases for public expectations, for what we justifiably expect of one another. So they must be able to be publicly articulated and accessible."³⁷ Darwall reveals his actualist inclinations most clearly when he claims that,

It would be unreasonable to hold people accountable to standards whose application required some special sensibility ordinary moral agents could not be assumed to have or that could not be formulated in ways that ordinary agents can understand. In this way, standards of moral right are a moral 'law', modeled on law as we generally understand it.³⁸

Principles of right "essentially involv[e] publicly articulable principles" since they are connected to "mutual accountability."³⁹ But if moral laws and principles are those endorsed by a merely hypothetical community, that is, if real-world moral demands can be in force solely on the basis of

³⁵ Strawson 1974, p. 15.

³⁶ Darwall 2007, p. 65.

³⁷ Darwall 2006, pp. 313-4.

³⁸ *Ibid.*, p. 314.

³⁹ *Ibid.*, p. 313.

the rational commitments of a hypothetical moral community, then how can those principles be public? The principles might be publicly recognized, but they will not thereby be publicly recognized *as justified* because the justification will be obscure to ordinary persons. For principles to provide real public expectations and to solve real problems, then they should be strongly tied to an actual community.

Now that said, a number of reasonable considerations move Darwall in a hypothetical direction. For instance, he wants to avoid the worries raised by Wallace, Watson, Smith and Strabbing. They are concerned that obligations seem in force even when no real-world acts of address or endorsement have occurred. Duties seem to exist without any actual person creating them. Further, it seems that individuals can be obliged to act in certain ways even if they are not presently aware of the second-personal reasons that apply to them, perhaps due to deep irrationalities. The push towards idealization in ethics is always motivated by worries that defective minds and communities will screw up or sully the generation and maintenance of duties.

To this point, we have not probed as deeply as we might into the sense in which Darwall's account is hypothetical and the sense in which it is actual. First, for Darwall, the hypothetical community is the true moral community. When we engage in second-personal moral practices we act as representatives of this realm of ends. For Darwall, moral demands are in force if members of the hypothetical community are disposed to make them, and so we real-world individuals can make authoritative demands of others only derivatively. When questioning the coherence of non-second-personal duties to God, Darwall claims that "It is not our actual acceptance that our obligation to him is conditional on but rather that we can reasonably be expected to accept it, that we would be unreasonably if we did not."⁴⁰ The hypothetical community is somehow the source of the force of moral demands. Real-world community members make legitimate demands only when their demands are instances of the demands that hypothetical persons would make of one another in the relevant contexts. In some way then, the actions and reasons of hypothetical community are the right and wrong-makers for real world action. Sometimes it looks as if the hypothetical community makes certain actions wrong in virtue of what they would *accept* and in other cases they make actions wrong or obligatory in virtue of what they would *demand*.⁴¹ But either way, the hypothetical community reigns supreme.

⁴⁰ Darwall 2007, p. 67.

⁴¹ I'm sure these two activities are related, but I'm not sure how.

If so, one might ask, what's the *point* is of a real-world community? The actions of real-world persons seem neither necessary nor sufficient to make a moral difference in the world. Second-personal real-world activities are not sufficient for moral obligations, as moral obligations depend on what the hypothetical community would demand, accept, etc. But neither are they necessary, as Darwall insists that moral demands can be in force so long as members of the hypothetical community are prone to make them. Real-world actors need not even be *disposed* to make such demands for them to be binding. Our real-world actions and our real-world feelings do not change the moral landscape. They only determine our obligations by setting the relevant context.

Darwall might reply that real-world activities are not meant to make a normative difference in the world. Instead, the aim of his account of second-person normativity is to use a hypothetical community to *explain* why our real-world second-personal practices have normative force. The endorsements and demands of a hypothetical community imbue real-world endorsements and demands with normative force insofar as the latter map onto the former. The challenge for an "explanation-only" account of second-personal practices is that it seems to entirely undermine the spirit of Darwall's theoretical enterprise. Our real-world social practices of holding responsible, mutual accountability and interpersonal address are supposed to be a *source* of normativity, not merely a *recipient*.

We are therefore left with a puzzle. On the one hand, Darwall wants our real-world moral behaviors to make some kind of normative difference. The content and existence of obligations should somehow depend on what we individuals actually do or at least what we can presently recognize. Further, second-personal norms are supposed to guide action and so they must be publicly recognizable—members of the actual community should be able to recognize these rules as their own. On the other hand, Darwall wants to avoid the trappings of actual communities and so makes the ultimate source of moral obligations and second-personal normativity depend on what an emphatically hypothetical community would accept or demand. If Darwall is too actualist, his account of obligation depends too much on the real-world activities of individuals, yet if he is too hypothetical, his account of obligation seems no longer rooted in our real-world practices. The connection with the reactive attitudes, actual address and publicity seems lost. This is the puzzle of idealization.

III: Social-Moral Rules

I shall argue that the puzzle of idealization can be resolved by rooting second-person normativity in rules rather than standpoints and reasons. Specifically, second-personal activities acquire their normative force insofar as they are instances of classes of behavior governed by social-moral rules. I shall then argue that social-moral rules acquire their normative force from being justified to the reason of each person, which can be *represented* by endorsement by a hypothetical community. Social-moral rules mediate the interaction of the actual and the hypothetical. They have a foot in both camps: they are actual because they are actually practiced and recognized as such, but they are hypothetical because they are endorsed by the idealized reasoning of members of the moral community.

Before using social-moral rules to resolve the puzzle of idealization, we must explain what they are. I will rely on an account of moral rules that Gerald Gaus calls the “Baier-Strawson” conception of social morality.⁴² Gaus follows Kurt Baier and Strawson in developing an account of social-moral rules that mediate second-personal address. Let’s review Baier and Strawson’s conception of moral rules and Gaus’s extension and unification of their views. We can then see how social-moral rules have “second-personality.”

Gaus defines “social morality” as “the set of social-moral rules that require or prohibit action, and so grounds moral imperatives that we direct to each other to engage in, or refrain from, certain lines of conduct.”⁴³ Social-moral rules “structure social interaction” and “allow us to live together in cooperative, mutually beneficial, social relations.”⁴⁴ They extend our aims, while “constrain[ing] our choices about how to pursue them.”⁴⁵ Social-moral rules are also imperative, for social morality provides “the basis for issuing demands on others that they must perform certain actions.”

Gaus defines a moral rule as a description of a class of moral actions or demands. For example, a moral rule that requires keeping one’s promises describes a class of actions—promise-keeping—and requires that each promise is kept. Gaus argues that moral address must occur primarily by means of “middle-level” social rules in order to solve coordination problems and produce stable expectations. Rules provide “guidance which is general enough to apply to unforeseen future circumstances while specific enough that we have common understandings of

⁴² Gaus 2011, p. 4.

⁴³ *Ibid.*, p. 2.

⁴⁴ *Ibid.*, p. 4.

⁴⁵ *Ibid.*, p. 6.

what the rule requires.” Specific cases of address are not general enough to coordinate individuals and general principles are too vague to do the same. Rules, for Gaus, “regulate the specific through the general” and play the role of specifying instantiation relations for particular acts of moral address and moral behavior.⁴⁶ Thus, moral address is usually made by means of moral rules—moral rules *mediate* second-person address. If John blames Reba for not keeping her promise, he blames her for not following the promise-keeping rule.

Moral rules are required for social coordination because they help “individuals with diverse ends to live together.” They do so, again, by maintaining stable expectations. Strawson emphasizes that social order can only exist when “certain expectations of behavior on the part of its members should be pretty regularly fulfilled.”⁴⁷ Further, what Strawson calls “the minimal interpretation of morality” understands moral rules as “a kind of public convenience” as a condition “of everything that matters.”⁴⁸ Like Darwall’s idea of moral norms, social-moral rules provide a public basis for social expectations and coordination.

Social-moral rules must also be *socially recognized*. Baier maintains that social-moral rules must be “part of the mores of the group” in that they are generally acted upon and regarded as binding.⁴⁹ For Strawson, these rules are “socially sanctioned demand[s] made on an individual” that are advanced based merely upon “his membership of the society in question, or in virtue of a particular position which he occupies within it or a particular relation in which he stands to other members of it.”⁵⁰ Gaus, like Darwall, understands social morality as an extant social practice. His account of moral rules is partly based on H. L. A. Hart’s conception of social rules. Rules involve (i) a social convergence in judgment, where those in a particular community “must look upon [and endorse] the behavior in question as a general standard to be followed by the group as a whole” and (ii) observed collective behavior where individuals generally follow the rule and hold one another accountable to it.⁵¹ Following Hart, Gaus emphasizes that rules have an “internal aspect” where the rule can only hold when at least some community members recognize it as binding.⁵² Darwall too argues that second-personality requires public recognition, despite not applying second-personality to rules.

⁴⁶ Ibid. pp. 122-123.

⁴⁷ Strawson 1974

⁴⁸ Ibid.

⁴⁹ Baier 1954, p. 108.

⁵⁰ Strawson 1974, p. 36.

⁵¹ Hart 1961, pp. 54-6. Gaus 2011, p. 165.

⁵² Darwall is well known for defended an “internal ought” that requires rational recognition. See Darwall 1995.

Social-moral rules must also be enforced with social sanction. Baier argues that moral rules must be “supported by the characteristically moral pressure.”⁵³ A socially sanctioned demand, for Strawson, “is doubtless a demand made with the permission and approval of a society; and backed, in some form and degree, with its power.”⁵⁴ Similarly for Gaus, social-moral rules ground moral imperatives that license blame or sanction when they are violated, as it gives persons authority to direct the action of their fellows.⁵⁵ Again, like Darwall’s account of second-person norms, social-moral rules license holding others accountable and blaming them for infractions.

But this social sanction also exacts costs from those who impose it, for social-moral rules are essentially reversible or universal. Moral rules are essentially *reversible* in that they apply to all persons: when John demands that Reba not violate a rule, he must also regard himself as subject to the rule.⁵⁶ Strawson insists that social morality cannot exist without “reciprocal acknowledgment of rights and duties.”⁵⁷ A universal feature of morality is “the necessary acceptance of reciprocity of claim.”⁵⁸ Baier also claims that moral rules are “universally teachable and therefore universalizable.”⁵⁹ Darwall obviously agrees that moral norms imply relations of reciprocity.

Finally, for Baier, Gaus and Strawson, moral rules must generally override other reasons for action. Baier notes that “moral reasons are superior to all others.”⁶⁰ Social-moral rules cannot perform their coordinating function unless they are so regarded: “The very *raison d’être* of a morality is to yield reasons which override the reasons of self-interest in those cases where everyone’s following self-interest would be harmful to everyone.”⁶¹

Social-moral rules therefore have five key features; they are:

- (i) Required for social coordination.
- (ii) Socially recognized.
- (iii) Enforced with social sanction.
- (iv) Reversible or universalizable.

⁵³ Baier 1954, p. 108.

⁵⁴ Strawson 1974, p. 38.

⁵⁵ Gaus 2011, pp. 8-9. Also see Gaus’s discussion of blame and punishment, pp. 193-201.

⁵⁶ Darwall acknowledges this point, but does not apply it to moral rules. See Darwall 2006, p. 23.

⁵⁷ Strawson 1974, p. 40.

⁵⁸ Ibid.

⁵⁹ Baier 1954, p. 108. Also see pp. 111-12.

⁶⁰ Baier 1958, p. 148. Note that Baier does not always restrict his conception of self-interest to non-altruistic conceptions of interests and value. In later work, Baier understands moral rules as trumping all “self-anchored” reasons. Baier 1995, p. 72.

⁶¹ Baier 1958, pp. 149-150.

(v) Able to generally override other reasons for action.

The Baier-Strawson-Gaus (BSG) conception of social-moral rules is strikingly similar to Darwall's second-person standpoint. The second-person standpoint includes the exchange of recognizable reasons. Not complying with these reasons licenses social sanction and indignation. The standpoint requires recognizing reciprocity and overrides many first-personal considerations. The main difference between BSG social-moral rules and Darwall's second-person standpoint is that social-moral rules explicitly mediate moral address.

IV: How Rules Resolve the Puzzle of Idealization

Recall the puzzle of idealization: an attractive account of second-person normativity should include both actualist and hypothetical elements, that is, real-world practices should make a bona fide difference in the moral landscape but moral norms and demands must be in force in the absence of explicit acts of moral address. An account of second-person normativity should include a publicly accessible set of moral norms that simultaneously acquire their justification for *good* reasons, not just the reasons acknowledged by ordinary, often irrational and misinformed individuals. Darwall wants to have it both ways, but it is not clear how he can.

Social-moral rules resolve the puzzle of idealization by illuminating two roles that actual and hypothetical demands and endorsements might play. Social-moral rules can be both *instantiated* and *justified*. That is, a particular action in accord with a social-moral rule is permissible because it instantiates the rules—it is an instance of following the rule. Further, a rule can be justified to all persons in virtue of their *rational* commitments, not their *actual* commitments. To put the solution to the idealization puzzle as a slogan: second-personal rules are *actually* practiced and *hypothetically* endorsed.

Let's illustrate with an example, which I shall call *Driver John*. Imagine John driving home from work. On the way, he must know and obey a large number of rules, some legal, some conventional. Legal rules include obeying the speed limit, not running red lights, and using his turn signal, whereas conventional rules include not passing cars from the right and the imperfect duty of allowing others into ongoing traffic. John knows when he risks getting a ticket and when he has unjustifiably cut someone off. He knows not to tailgate others and not to use his brights in the face of on-comers. I submit that these traffic rules have second-personality. They are generally regarded

as justified to all and are taken to ground accountability relations. If John cuts Reba off in traffic, she may honk at him to signal her resentment that he broke the rules or simply to enforce the rules so that John will be less likely to break them. Most penalization relies on the common recognition that drivers are part of a cooperative association that allows all to get where they are going. Most of us (including, we assume, John) do not like being honked at because we know that others are angry with us for violating a rule. We sometimes feel embarrassed, angry or ashamed at the slightest indication that a fellow driver is unhappy with our conduct. We recognize our equality on the road as well, since we generally believe that no one deserves to be cut-off or tailgated because on the road, breaking the rules indicates our willingness to exercise domination over equals.

Driver John has the actualist and hypothetical elements I have outlined. The hypothetical element is rooted in the common acceptance of the rules of the road. It may well be that John fails to accept traffic rules when he is late for work, or when he complains that a speed limit is set too low. But a hypothetical John surely endorses traffic laws because he is rationally committed to recognizing that the rules of the road make everyone better off, even if they require that he sacrifice his short term interests, say, in getting to work on time. I must emphasize that the hypothetical element must be understood as a *heuristic* for rational commitment, not as a truth-maker in itself. Frequently contractualists like Darwall appeal to hypothetical endorsement because they take it that the facts about what would be hypothetically endorsed are themselves right-makers for particular norms. But this leads to the obvious question of why a member of a hypothetical community would endorse such a norm. I therefore assume that hypothetical endorsement is a stand-in for *rational commitment*. John is rationally committed to norm X when he has conclusive reason to endorse X. In defining rational commitment in this way, I appeal to an epistemological notion of justification.

Specifically, I appeal to the epistemic notion of *open justification*. The concept of open justification can be understood in terms of the idea of a belief-value set. John's belief-value set is the set of all of his beliefs, desires, goals, and plans, i.e., everything he thinks and wants.⁶² He can also reason by inference from some of these elements to generate new commitments, be they to values or beliefs. I take it that John is epistemically justified in affirming these new commitments when he has some kind of psychological *access* to these commitments. In epistemology, this is a version of access internalism, which holds that whenever one is justified to believe P one can become aware by

⁶² The idea of a belief-value set derives from Bernard Williams's conception of a subjective motivational set. See Williams 1981, p. 102.

reflection of all of one's justifiers that P.⁶³ The idea of open justification is a version of access internalism about moral reasons. On Gaus's view, open justification "treat[s] [one's system of beliefs and reasons] as open to new information and arguments and, from this external perspective, mak[ing] judgments about what would then be justified in [one's system of beliefs and reasons]."⁶⁴ The key feature of open justification is that it "takes Alf's current belief system as the point of departure."⁶⁵ For Gaus, beliefs and reasons can acquire justification inferentially or foundationally and foundational beliefs can be generated by feelings or intuitions which give them *prima facie*, though weak, credibility.⁶⁶ All beliefs relevant to hypothetical endorsement can be accessed either as foundational beliefs or by means of inferential chains from those foundational beliefs. Open justification can be ascribed to beliefs based on information an agent may not presently possess but could possess under the proper circumstances. Individuals are therefore rationally committed to beliefs and reasons that are not currently a part of their belief-value sets. To openly justify reason X to John is to show that it can be properly inferred from John's present belief-value set conjoined with adequate information, reasoning ability and some degree of belief coherence.

We can now say that John hypothetically endorses traffic rules because he is openly justified in affirming traffic rules. The combination of John's present belief-value set combined with sound rules of inference and information requires that he endorse the rules of the road. The rules are valid and binding for John, therefore, *not* because he actually accepts them but because he is openly justified in accepting them (and not justified in rejecting them). *Driver John* is rationally committed to the rules because he can become aware by reflection of his justifiers for the rules, which arguably include the realization that the rules make him better off and reflect relations of reciprocity.

The open justification of rules preserves the various benefits that Darwall derives from his account of a hypothetical moral community. First, it explains how rules can be binding even when no one acts on them. So long as each member of the actual moral community is openly justified in affirming the rules, the rules are in force. Thus, anyone can appeal to the rules and thereby make justified demands on others because all persons are rationally committed to the rules even in lieu of their actual behaviors. Further, they are rationally committed to the rules in lieu of their actual *endorsements*. Rational commitment and actual commitment come apart, and therefore the process of

⁶³ For more on access internalism, see Chisholm 1989, p. 17.

⁶⁴ *Ibid.*, p. 31.

⁶⁵ *Ibid.*, p. 32.

⁶⁶ *Ibid.*, pp. 64-65, 83-84, 108-109.

second-personal justification is not held hostage to real-world irrationalities. In this way, a second-personalist can set aside the worries raised by Wallace, Watson, Smith and Strabbing.

Now let's turn to the actualist element. Real-world second-personal demands acquire their authority by instantiating second-personal rules to which the moral community is rationally committed. *Driver John* makes a valid second-personal demand when he honks at Reba for running a red light because the traffic laws requiring all drivers to stop at red lights is openly justified to Reba. The actual second-personal demands, therefore, hold Reba accountable in the second-personal manner that Darwall describes. Second-personal demands are not required to create a norm, but they are required to hold others accountable for violating the rule and impose penalties upon them. Recall that Darwall claims that second-personal norms exist only if second-personal authorities exist. We can now explain why this is so. Second-personal authorities are all and only those members of the moral community to whom the rules are openly justified. But without these authorities, there would be no such second-personal norms because the norms are only *bona fide* insofar as they are openly justified to the relevant, actual community. Second-person norms thereby depend on second-personal authorities. Also recall that Darwall claims that second-personal reasons derive their force from legitimate demands; second-personal rules explain this dependence relation as well. Legitimate demands derive from openly justified rules. If second-personal reasons are understood as reasons to comply with or enforce second-personal rules, then second-personal reasons cannot exist outside of the possibility of legitimate demands since both second-person reasons and legitimate demands are licensed by openly justified rules.⁶⁷ Attaching second-personality to rules also help explain how hypothetical endorsement and real reactive attitudes are connected. Rules of the road are sufficiently publicly recognized and their rationale sufficiently rationally accessible that we feel entitled to automatically resent or remain indignant due to rule violations. *Driver John* is the proper subject of reactive attitudes by his community when he violates traffic rules. These rules are openly justified (we presume) and publicly recognized and for these reasons, and so they license the reactive attitudes directly. Notice further that the publicity condition is satisfied as well. Again, social-moral rules are essentially part of the mores of a community. In other words, they are actually recognized rules. People can recall and often articulate the rules when led to reflect upon them. They are in this sense *public* because they are the object of real recognition by the community, if not through their words then their actions.

⁶⁷ Thus second-person reasons don't really *depend* on legitimate demands, but correlate with them.

To complete the explanation of how second-personal rules retain the attractive features of an actualist conception of second-person normativity, we must explain how real-world second-person activities make a real moral difference in the world. So far, we have seen that real-world second-person activities derive their validity and authority by instantiating openly justified rules. But this only shows that these activities have derivative authority, not how they help to create and sustain that authority. To see how second-personal activities have *creative* power, consider an analogy with linguistic practice. We all are committed to a set of grammar rules. We by and large know what the rules despite frequently violating them. We cannot always articulate the rules, but we nonetheless accept them as valid. Corrections to others' language use have authority because the rules themselves are definitive of correct usage. In this way, grammar rules and second-person rules have parallel normative force. But notice that grammar rules change depending on actual practice. Within a short time-period, deviations from grammar rules count as errors, but within a long time period, deviations from grammar rules count as *changes in the rules*. In the short-run, proper language use is determined by the rules of language *usage* but in the long-run, the rules of language usage are determined by language *use*. The source of these changes is, somewhat paradoxically, the actual use of language. Our actual linguistic practices sometimes maintain and sometimes alter the grammar rules that bind us all. Second-personal moral activities make a similar normative difference in the world. Their effect is not immediate, but gradual and cumulative. Second-personal moral practices sustain the norms of second-personal normativity in a particular community. This means, perhaps controversially, that what is second-personally justified can change depending on the community and its historical and cultural context. But I submit that this is a merit of the theory, as it is plausible that second-personal norms do shift and develop over time.

Real-world moral behavior can make a normative difference in a second way as well. Second-personal norms often ascribe authority to *speech acts*. Promise-keeping is perhaps the most well-known example. Promises are binding when they are validly made. On the second-person rule account of second-personality, then, actual promises acquire their normative force from being proper instances of a promise-keeping rule. These actual promises only have authority insofar as the promise-keeping rule is openly justified to the community, and so the validity of real-world promise-making depends on what norms are openly justified to the community. However, promises still make a normative difference in the world by creating obligations. Real-world second-personal activity therefore acquires its authority from what community members rationally endorse, while also making a public, normative difference in the moral landscape.

V: A New Conception of Second-Personal Reasons

A benefit of the rule-based account of second-personal normativity is that it allows us to precisely define a conception of second-personal reasons. Darwall defines a second-personal reason as any reason issued in a second-personal context. But this is too vague. Many reasons can be issued in a context where persons share a moral community and second-personal moral practices. The fact that a reason is issued within a practice therefore seems insufficient for the reason to be truly second-personal. Thus, the second-personal aspect of a reason should not be built into the context in which the reason is given but by *the reason's relation to a social-moral rule*. To illustrate the distinction, let's return to *Driver John*. Assume that John is committed to a set of ordinary traffic laws and he cuts Reba off, violating a "no cutting" rule recognized and justified to all. When Reba honks her horn at John for cutting her off, I think she issues John a second-person reason. But the reason is rather simple; it is "Don't violate the rules!" when the rules are those rules relevant to the particular context. Thus the second-personal reason derives its second-personality from being associated and tied to the social-moral rule in question. I call this a *truss* conception of second-person reasons. The truss conception of second-person reasons holds that reasons are second-personal when they take part in justifying, maintaining or reforming a social-moral rule. Social-moral rules are justified and modified in accord with a vast array of moral reasons issuing from diverse sets of reasons and values endorsed by each individual. These reasons are the "trusses" that make up the social-moral rules that bridge each person's values and create a second-personal world of moral relations.

Before defining the truss conception, I should distinguish "moral reasons" from instrumental reasons and second-person reasons. I take a moral reason for A to X to be one that *morally requires* A to X. There is a distinctively moral "ought" or "requirement" relation such that one *does wrong* when she fails to comply with it. When we fail to act on pragmatic or instrumental reasons, we do not do *wrong* even if we act poorly. Moral reasons are ones such that, all else equal, we act rightly when we act on them and wrongly when we do not. Given this definition, not all moral reasons are second-personal in Darwall's sense. I also appeal to the notion of open justification. I can now define a second-person reason as follows:

Second-Person Reason: X is a second-person reason for A iff (i) X is a moral reason for A and (ii) X is a reason to demand, comply with demands, etc. in accord with some social-moral rule M openly justifiably believed by A to be endorsed by members of A's community.

The class of second-person reasons is a subset of the class of moral reasons, thus substantiating condition (i). Second-person reasons are that subset of moral reasons that concern the issuing of demands or responding to them in accord with social-moral rules. They are *bona fide* if they are based upon some moral rule M, that is, they are reasons to comply with M, to issue demands to comply with M, etc. The person who has the second-person reason must also believe that M is openly justified to all others for their own moral reasons. If she did not so believe, she would have no reason to issue demands based upon M or comply with demands justified by M.⁶⁸ X thereby has second-personality in two senses: (a) it relates to demands made upon others and (b) it is a reason to act in accord with a social-moral rule.

There are many types of second-personal reasons. I've mentioned two types already: reasons to demand and reasons to comply with demands. I call these *first-order* second-personal reasons, as they derive directly from M. However, there are also more complex types of reasons, such as second-personal reasons to modify or reject other second-personally justified norms. These second-personal reasons are *second-order* second-personal reasons because they concern the practices of moral reform, of altering, creating or removing a social-moral rule. One advantage of understanding second-person reasons in relation to moral rules is that we can neatly distinguish between types of second-person reasons while showing that they retain a common rule-regarding core.

VI: Conclusion

The point of this piece is to move beyond the standard treatment of second-person normativity provided by Darwall. Again, in my view, Darwall has correctly identified a unique moral domain, but his articulation of the domain need not be the only one on offer. I developed a successor rule-based conception of second-person normativity by articulating and pressing Darwall's account of second-personality with a puzzle, what I call the Puzzle of Idealization. The Puzzle of Idealization holds that any account of second-personal normativity requires both actualist and hypothetical. We saw that Darwall seems torn between the actual and hypothetical because he wants his account of second-person normativity to retain all the relevant benefits available to both the actual (publicity, accountability and real-world reactive attitudes) and the hypothetical (rational endorsement in the absence of acknowledge acknowledgment). Second-person *rules* can resolve this problem by

⁶⁸ We cannot rule out cases where Reba has some independent reason to make demands that just so happen to be in accord with rule M. Further, the beliefs mentioned need not be occurrent when making the moral demand.

mediating between the actual and the hypothetical. These social-moral rules, as articulated by Baier, Gaus and Strawson are *actually* practiced and *hypothetically* endorsed. In this way, second-person rules authorize actions that make a real moral difference in the world, retain publicity and justify the reactive attitudes, but that also avoid basing morality on person's irrationalities and on our obligations depending entirely on real-world second-person address.

My focus on second-person rules has a number of other advantages. One advantage discussed is that we can distinguish types of second-person reasons, but three more bare mentioning. The focus on rules allows us to distinguish between three avenues of future research into second-personal normativity. First, we can focus specifically on the nature of the real-world practices whose authority derives from extant but justified moral rules. How, for instance, do second-personal social practices bind the wills of others by making reference to rules? Is the primary mechanism that speech act? Or is the primary method the fact that second-personal moral practices refine and maintain justified social-moral rules? Second, research can focus on the nature of social-moral rules. In what way are they publicly recognized? How are they sustained through real-world practice? How do they authorize the reactive attitudes? Finally, we can ask how exactly to understand hypothetical endorsement. In this piece, I have defended a standard of rational commitment understood in terms of open justification. But there may well be alternative accounts, some of which may be better than the account I have articulated. Thus, research can divide along three related paths: the actual, the rule and the hypothetical, all now conceptually separated but part of a renewed research project within deontological and contractualist ethics.

References

- Baier, Kurt. "The Point of View of Morality." *Australasian Journal of Philosophy*, vol. 32 (1954): 104-135.
- . *The Moral Point of View: A Rational Basis of Ethics*. Ithaca: Cornell University Press, 1958.
- . *The Rational and the Moral Order*. Chicago: Open Court, 1995.
- Darwall, Stephen. *The British Moralists and the Internal "Ought", 1640-1740*. New York: Cambridge University Press, 1995.
- . *The Second-Person Standpoint: Morality, Respect, and Accountability*. New York: Harvard University Press, 2006.

- . "Reply to Korsgaard, Wallace, and Watson." *Ethics*, vol. 118 (2007): 52-69.
- . "Authority and Reasons: Exclusionary and Second-Personal." *Ethics*, vol. 120 (2010a): 257-278.
- . "'But It Would Be Wrong'." *Social Philosophy and Policy*, vol. 27 (2010b): 135-157.
- . "Precis: The Second-Person Standpoint." *Philosophy and Phenomenological Research*, vol. 81 (2010c): 216-228.
- . "Reply to Schapiro, Smith/Strabbing, and Yaffe." *Philosophy and Phenomenological Research*, vol. 81 (2010d): 253-264.
- Gaus, Gerald. *The Order of Public Reason*. New York: Cambridge University Press, 2011.
- Hart, H. L. A. *The Concept of Law*. Oxford: Clarendon Press, 1961.
- Locke, John. *Two Treatises of Government and A Letter Concerning Toleration*, edited by Ian Shapiro. New Haven: Yale University Press, 2003.
- Ridge, Michael. "Contractualism and the New and Improved Redundancy Objection." *Analysis*, vol. 63 (2003): 337-342.
- Scanlon, Thomas. *What We Owe To Each Other*. Cambridge: Harvard University Press, 1998.
- Smith, Michael and Jada Twedt Strabbing. "Moral Obligation, Accountability, and Second-Personal Reasons." *Philosophy and Phenomenological Research*, vol. 81 (2010): 237-245.
- Strawson, P. F. *Freedom and Resentment and Other Essays*. New York: Routledge, 1974.
- Wallace, R. Jay. "Reasons, Relations, and Commands: Reflections on Darwall." *Ethics*, vol. 118 (2007): 24-36.
- Watson, Gary. "Morality as Equal Accountability: Comments on Stephen Darwall's The Second-Person Standpoint." *Ethics*, vol. 118 (2007): 37-51.