

# Second Person Rules: An Alternative Approach to Second-Person Normativity

Kevin Vallier<sup>1</sup>

Published online: 23 December 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** Stephen Darwall’s moral theory explains moral obligation by appealing to a “second-person” standpoint where persons use second-person reasons to hold one another accountable for their moral behavior. However, Darwall claims obligations obtain if and only if *hypothetical* persons endorse them, despite tying the second-person standpoint to our real-world moral practices. Focus on hypothetical persons renders critical elements of his account obscure. I solve this problem by distinguishing two ideas quietly working in tandem, (i) the hypothetical endorsement of moral norms and (ii) the hypothetical recognition of these norms. Hypothetical endorsement is a plausible source of normativity; hypothetical recognition is not. A more plausible account of second-person normativity must combine hypothetical endorsement with *actual* recognition. I term these alternative conceptions *justification* and *easy publication*. To combine justification and easy publication in an account of moral obligation, second-person normativity should be applied first to rules. Following other moral philosophers, I introduce the concept of a “social-moral” rule into an account of moral obligation. Social-moral rules acquire normative force when they are justified for and easily published by the relevant moral community. I conclude that a *rule-centric* account of second-personality is superior to Darwall’s *reason-centric* account.

**Keywords** Normative reasons · Second-person reasons · Moral rules · Darwall

Stephen Darwall’s work on the second-person standpoint (*SPS*) provides valuable insight into what I shall call our *social-moral* practices, those social practices associated with moral obligation, including holding one another responsible for wrongdoing, making moral demands of others and blaming them for infractions. These

---

✉ Kevin Vallier  
kevinvallier@gmail.com

<sup>1</sup> Department of Philosophy, Bowling Green State University, Bowling Green, OH, USA

'second-personal' practices are grounded in a network of normative concepts whose force can only be appreciated by persons who take up a shared *SPS* where they hold one another accountable on mutually acceptable, authoritative terms. I assume here that Darwall's basic claim 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 must thereby permit him to hold her accountable for wrongdoing and blame her accordingly. The flurry of criticisms of Darwall's work, however, suggests that we may need another account of the nature of second-person normativity. In this paper, I offer such an alternative.

We can accurately describe Darwall's account of second-person normativity as *reason-centric* because the notion of a second-person reason is the ground of second-person normativity. I believe an alternative *rule-centric* interpretation of second-person normativity is superior. Following moral philosophers Kurt Baier, H.L.A. Hart, P.F. Strawson and Gerald Gaus, I place *social-moral rules* at the heart of interpersonal morality. Social-moral rules derive their normative force from two sources: their endorsement by a hypothetical community and their recognition and publication by an actual community. On the rule-centric approach, hypothetically endorsed rules must be actually recognized rules of conduct. As such, they can coordinate and ground real-world moral practices while still being rooted in the rational commitments of the moral community.

The rule-centric account of second-person normativity has two advantages over Darwall's account. First, it clears up what I regard as the confusion of hypothetical endorsement and hypothetical recognition in Darwall's account of second-person normativity. While the idea of hypothetical endorsement of a norm by a hypothetical community makes sense, it is less clear how to understand the claim that moral norms are in force when they are only hypothetically recognized. Settling for hypothetical recognition undermines the interpersonal element of second-person normativity by detaching second-personality from real-world recognition of the value of other persons as sources of reasons. By rooting moral norms in actual recognition, the rule-centric account of second-person normativity avoids this problem. The rule-centric account also makes better sense of the notion of a second-personal reason. Darwall sometimes claims that a second-person reason is any reason issued within second-personal accountability relations. But this, I shall argue, is too general to be helpful. The rule-centric approach to second-person normativity derives the second-personality of reasons from the rules to which they are attached. A second person reason is understood as a reason to make demands or comply with demands in accord with moral rules justified for and publicly recognized by persons in a moral community.

My argument has eight parts. I first review Darwall's understanding of second-personal normativity and suggest that the hypothetical interpretation of second-person normativity is in tension with Darwall's commitment that standards of right and wrong be publicizeable, and so appropriately serve as grounds for public expectations. I then argue that these problems can be resolved by distinguishing between the recognition and justification of norms. Next I introduce the rule-centric approach to second-person normativity and shows that it retains the attractive features of Darwall's approach but by appealing to more plausible conceptions of

recognition and justification. I then try to show how the rule-centric account avoids Darwall's appeal to hypothetical recognition. The following section addresses the common concern that actual recognition accounts of obligation cannot account for societal blind spots. The next articulates the rule-centric conception of second-personal reasons; the final section concludes.

## Darwall's Account of Second-Personality

The *SPS* is supposed to provide the best foundation for contractualist theories of right action. Contractualism proposes to explain the rightness of (many) 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' (Scanlon 1998, p. 153). Darwall points out that all contractualisms share the idea 'that we are accountable for regulating ourselves by principles that are acceptable... to each as equal moral persons' (Darwall 2006, p. 304). We can appreciate this fact best when we take the *SPS*. 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' (Darwall 2006, p. 3).<sup>1</sup> This standpoint 'commits us to regulating our conduct by principles that are acceptable, or not reasonably rejectable, to each as free and rational agents' (Darwall 2006, p. 300).

For Darwall, the *SPS* can only be explicated and defended by appealing to a 'circle of irreducibly second-personal concepts' (Darwall 2006, p. 11). No concept can be fully understood apart from the others and no explanation of the *SPS* can begin from outside of the circle. Darwall claims that there are (at least) four such concepts: (i) practical authority, (ii) accountability to others, (iii) valid claims or demands, and (iv) second-personal reasons (Darwall 2006, p. 13; Darwall 2010a, p. 266; Darwall 2010c, pp. 217–218).

- (i) Practical authority obtains when persons can require certain lines of action from an agent 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' (Darwall 2010a, p. 266). 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 can obligate B to  $\phi$  in the appropriate context. But this relationship only obtains when parties recognize the capacity for moral authority in one another. This mutual recognition arises when second-personal reasons are taken to apply to the parties in question.

<sup>1</sup> Making and acknowledging claims is to practice interpersonal morality.

- (ii) Darwall understands moral authority as requiring mutual accountability to others. He cashes out the idea of mutual accountability by drawing on similar work by P.F. Strawson. 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’ (Darwall 2006, p. 15).<sup>2</sup> According to Darwall, Strawson claims that we can hold others morally responsible for their wrongful actions only when they can recognize a reason not to act wrongly (and would be unreasonable not to). 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 (Strawson 1974, p. 15). Since reactive attitudes concern what claims we can hold others to, Darwall argues that they ‘presuppose the authority to demand and hold one another responsible for compliance with moral obligations’ (Darwall 2006, p. 17). We cannot obligate others merely by willing that we do so, by desiring that the obligation obtain. Instead, our social practices and attitudes commit us to addressing the reasoning of others. Darwall has recently explained the idea of accountability in 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’ (Darwall 2010a, p. 266). 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, two preconditions must be met. First, those making claims must take the *SPS* and offer second-personal reasons. Second, claim-makers acquire their second-personal authority by acknowledging—at least implicitly—a norm of reciprocity, that others can make claims on our will and conduct 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...’ (Darwall 2006, p. 21). Further, valid claims and demands occur in terms of legitimate forms of address. A legitimate form of address is one that can obligate. 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 (Locke 2003, p. 225). 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 along these lines: ‘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’

<sup>2</sup> For the original essay, see Strawson (1974, pp. 1–28).

- (Darwall 2010a, p. 266). 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 despite Reba's belief to the contrary. His attitude expresses a failure of reciprocity that deprives him of obligation-making power.
- (iv) The *SPS* generates obligations 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 ineliminable pronominal back-reference' to the person who has the reason in question (Ridge 2003, p. 339). Agent-relative reasons are reasons that are essentially 'reasons for' a particular agent. Second-person reasons are agent-relative because they are essentially *recognizable* reasons. For John to obligate Reba, she must be able to 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 cannot ground an obligation. Second-personal reasons are thus agent-relative reasons that have the power to obligate those to whom they are addressed within already recognized moral relations. 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' (Darwall 2006, p. 8). Without presupposed authority and accountability relations, second-personal reasons cannot generate obligations. Thus, Darwall understands a second-personal reason as 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' (Darwall 2010a, p. 266). Second personal reasons are second-personal because they stem 'from a legitimate demand, and hence, from the authority to issue the demand' (Darwall 2010d, p. 260). Reasons arise from such demands that, in turn, can only obligate via preexisting relations of mutual accountability and reciprocal practical authority. 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*' (Darwall 2010b, p. 151).

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. The position is now clear enough to discern an important problem that troubles this account of normativity.

## The Problem of Publication

One common criticism of Darwall is that his conception of second-person normativity is too 'actualist' because it seems to imply that 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. R. Jay Wallace argues that Darwall's example of a second-personal 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' (Wallace 2007, p. 26). 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' (Wallace 2007, p. 27). Wallace also insists that Darwall's account of obligation makes 'moral obligation hostage to the actual responses of the individuals implicated in interactions with each other, in ways that are problematic' (Wallace 2007, p. 27). 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' (Watson 2007, p. 39). Watson realizes that the difference must appeal 'to certain implicit understandings and hypothetical conditions' but he worries that 'it is never well worked out' (Watson 2007, p. 40). 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 (Smith and Strabbing 2010, p. 239, ft. 2).

However, 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 argues that when we hold Reba accountable for violating her obligations and blame her accordingly, our demands are not issued as individuals but as people who represent the moral community (Darwall 2010b, p. 146). Darwall claims that our moral practices 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.' While real-world actions are authoritative, they are so in virtue of expressing the conclusions of an ideal human community. Darwall's commitment to a hypothetical interpretation leads him at times to completely sever the link between actual persons and hypothetical persons as when he 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* (Darwall 2010b, p. 146). 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' (Darwall 2007, p. 64). 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' (Darwall 2007, p. 65).<sup>3</sup>

If any part of second-person normativity has an actualist component, it only arises in the context of what Darwall now calls 'bipolar obligations' or those specific normative states that 'always involve a relation between two 'poles': an

---

<sup>3</sup> Emphasis mine.

agent who is obligated (the obligor) and an individual, group, etc., to whom she is obligated, tied, or bound (the obligee)' (Darwall 2012, p. 333). Bipolar obligations are created only through the explicit exercise of normative powers, such as promising, contracting, forgiving and other actions that create and alter moral status, that in other words alter the moral landscape. But the more general form of moral obligation, 'moral obligation period' involves norms for which anyone can hold violators responsible. Consider an obligation not to litter. If John litters, then anyone, including his wife, can blame him. In this case, those who blame John for littering experience the impersonal reactive attitude of indignation and when they demand that he pick up the trash, they exercise *representative* authority, authority that anyone can exercise in virtue of the existence of moral obligations as such. But moral obligation[s] period obtain even if no moral power like promising is exercised, nor need they be actually acknowledged or recognized. Instead, they remain in force if and only if a hypothetical community endorses them. In this way, Darwall can provide a powerful reply to Watson, etc., who asks Darwall to supply a distinction 'between 'demands' that remain in force whether or not one issues them and demands that do not' (Watson 2007, p. 39). Moral obligations period and their associated demands remain in force whether or not one issues them, even if the rarer and more specific bipolar obligations do not.<sup>4</sup>

Setting aside bipolar obligations, then, we can safely declare that Darwall is committed to a hypothetical account of second-person normativity. A hypothetical account of moral obligation holds that a moral obligation  $\phi$  is in force if and only if a hypothetical community would endorse it in the relevant second-personal sense. That is, hypothetical persons would accept or endorse the norm on which the obligation is based.

The main problem for Darwall's account of hypothetical endorsement is that it is in tension with the claim that moral obligations are publicizeable; I call this the *problem of publication*. Darwall notes that, 'Standards of right and wrong are, in their nature, bases for public expectations, for what we justifiably expect of one another. So they [the standards and bases] must be able to be publicly articulated and accessible' (Darwall 2006, pp. 313–314). Principles of right 'essentially involv[e] publicly articulable principles' since they are connected to 'mutual accountability' (Darwall 2006, p. 313). But if moral laws and principles are those endorsed by a merely hypothetical community, then the connection between moral justification and actual publicity is weakened. A hypothetical community can endorse norms for reasons that are entirely obscure to real-world individuals. This suggests that real-world moral demands can be in force solely on the basis of the rational commitments of a highly idealized group. Yet this would entail what Darwall wishes to deny, namely that we cannot justifiably hold people accountable to standards that require 'some special sensibility' that ordinary moral agents arguably lack—the ability to access the reasoning of a hypothetical community (Darwall 2006, p. 314). Darwall could argue that the proper characterization of a hypothetical community will demonstrate that their standards can be understood or articulated by actual individuals under the proper circumstances. For instance, given

---

<sup>4</sup> I shorten 'moral obligations period' to moral obligations.

that the hypothetical community will be aware of the limitations of their actual counterparts, they may restrict their endorsement requirements to accommodate their simpler brethren. But to do so, they will in effect build in the conditions that I will develop below.

Darwall could reply by distinguishing between the reasons persons have to endorse a moral principle and the publicizeability of the principle itself. Perhaps the rationales appealed to by a hypothetical community can be so complicated that, while they can *in principle* be explained to persons, doing so would require an enormous amount of education and effort. Nonetheless, the principles could be published with the understanding that there are in fact good, accessible rationales for those principles. As an analogy, we hold people to policies that can only be justified based on the judgments of experts, whose reasoning is sound but hard to access; we can permissibly hold persons to those policies because we have reason to think that there are successful rationales for the policy. Similarly, we can expect persons to recognize mathematical truths without having the capacity to evaluate the logical foundations of those truths, many of which even the most celebrated logicians cannot spell out. Persons do not require ‘some special sensibility’ in order to understand the justification of the relevant principle or policy; instead, they would only need to endure a lengthy and time-consuming educational process.

The problem with this reply is that Darwall needs the relevant justificatory reasons to be more accessible to ordinary persons than complex mathematical truths. This is because part of the point of the *SPS* is to provide an account of the sorts of reasons that are *generally available* for use in issuing moral demands. On Darwall’s view, we cannot simply insist that others comply with justified moral demands without offering or being prepared to offer the *reasons* that justify the moral demands. In Darwall’s core cases, such as the second-personal reason we have not to step on another person’s toe, we owe it to *the other person* not to cause him pain without good reason (as opposed to the third-personal reason that we should avoid causing others pain). Reasons, on Darwall’s view, are the currency of moral practice, and so cannot be obscure or hard to make sense of. Second-person reasons need to be *public* reasons in that diverse agents can offer the reasons to others as comprehensible rationales and allow others to use the same reasons to them on reciprocal terms. This commitment to giving an account of the reasons that undergird interpersonal obligation limits Darwall to the reasons that can be generally deployed in ordinary moral discourse, such as reasons not to harm others, to keep our promises and to help those in need. Darwall’s view becomes much less plausible if we permit others to hold us accountable based on reasons that we may find deeply alien and even hostile to our own conception of our interests, even if a hypothetical community endorses those reasons.

## Recognition, Justification and Publication

On second-person accounts of normativity, the concept of *recognition* is central to making sense of moral obligation. The hypothetical account of second-personal normativity supplies a natural interpretation of the idea of recognition, specifically



an idealized recognition. Second-person normativity also relies on a notion of moral *justification*, where principles must be justifiable to hypothetical persons in second-personal terms. The hypothetical interpretation of second-personal normativity attempts to capture both functions. Given the *problem of publication* with the hypothetical interpretation, we should look for an alternative. I believe our search must begin by distinguishing the ideas of recognition and justification and developing alternative accounts of both. I shall call the new recognition condition *easy publication*; I shall simply refer to the moral justification condition as *justification*.

Let us begin by developing an account of justification. On the hypothetical view of second-person normativity, a moral norm is justified when members of a hypothetical, rational community would endorse it. But it is hard to see how a norm can be justified for John based merely on the fact that a hypothetical John would endorse it. The problem of identifying with one's idealized counterpart is familiar in normative ethics, but I believe we can set the problem aside if we understand hypothetical endorsement as a *heuristic for rational commitment*. What norms hypothetical John would endorse are fixed by facts about which norms John must endorse on pain of irrationality. In this way, hypothetical endorsement does not make it true that John is obligated to comply with a norm. Instead, the truth-maker for the obligation is the fact that his values and beliefs commit him to accepting the obligation. When John is rationally committed to X, or when X is *justified* for John, then it is true that a hypothetically rational and informed John would endorse X. On this view, then, hypothetical endorsement is a buck-passing heuristic for rational commitment.

I understand rational commitment as a form of *epistemic justification*, specifically access internalism. In epistemology, access internalism holds that whenever one is justified to believe P one can become aware by reflection of all of one's justifiers that P (Chisholm 1989, p. 17). Justifiers can be understood in terms of the idea of a belief-value set, the set of John's beliefs, desires, goals, and plans, i.e., everything he thinks and wants (Williams 1981, p. 102). He can also reason by inference from some of these elements to generate new endorsements. This is the sense in which John can psychologically access his commitments. On an internalist view of epistemic justification, we can treat a system of beliefs and reasons 'as open to new information and arguments and, from this external perspective, make judgments about what would then be justified' in that system (Gaus 1996, p. 31). The key feature of this form of access internalism is that it takes John's 'current belief system as the point of departure' (Gaus 1996, p. 32). To epistemically justify reason X to John is to show that it can be properly inferred from his present belief-value set conjoined with adequate information, reasoning ability and some degree of belief coherence. In sum, then, we can replace hypothetical recognition and endorsement with the more precise notion of access internalist justification, or justification for short.

We must now offer an alternative to the idea of purely hypothetical public recognition. For Darwall, norms have second-personal force only when they are recognized by all members of a hypothetical moral community. But to maintain the publicity constraint on second-personal obligations, we need a standard of *easy*

*publication*. Easily published norms need not be explicitly endorsed and articulated by the community, but they should be relatively easy to psychologically access and *make* a matter of public knowledge.

To understand ease of publication, first consider an analogy with norms constituting monetary practices. Monetary norms cannot serve as the basis for economic cooperation unless it is relatively easy to tell what counts as money. Most Americans actually recognize that dollars serve as money, but were a foreign visitor confused about which currency Americans use, she could easily figure out the answer. Or contrast the relative publishability of tomorrow's weather forecast and the cure for polio. While I may not presently recognize the probabilities meteorologists assign to various weather conditions for the next day, I can easily access their predictions via television, computer or smartphone. Information about tomorrow's weather is thereby publishable to ordinary moral agents. The cure for polio is a matter of public record—anyone of normal intelligence can acquire the relevant training to make sense of the cure and know where to find it. But it is not easily publishable given the complex knowledge required to understand how it works. The norms on which moral obligations are based must be closer to information about tomorrow's weather for a conception of second-person normativity to be plausible.

Easy publication does not require an especially strong form publicity, where 'the reasons justifying morality are accessible to all,' but rather that members of a moral community can make one another aware of the norms on which their practices of holding responsible are based (Gaus 2011, p. 296). Easily published rules need merely be teachable to ordinary moral agents; their rationales need not be. It would be too much to require that moral agents know the reasons justifying moral norms for all persons, as reasons may vary considerably between persons. A shared public morality cannot depend on exercising the extraordinary cognitive capacities required to satisfy a strong publicity condition. Otherwise our practices of holding responsible are rooted in esoteric cognitive capacities and cease to be applicable by normal members of the moral community. Instead, ease of publication is determined by the degree of reasoning required to identify the norm as *in force*, or as I shall say below, part of the mores of the moral community. The standard of reasoning required to identify a norm as in force is considerably simpler than the standard of reasoning required in order for us to determine whether the norm is justified for all in the access internalist sense. Returning to our monetary analogy, it is much easier to determine *that* dollars are the American medium of exchange than why each person has reason to endorse dollars as the medium of exchange. That is, we can much more easily determine whether a norm is part of the moral community's mores or social practices than we can determine whether the norm *ought* to be part of those mores or practices.

Any adequate account of second-personal obligations should appeal to both easy publication and justification. Unless a norm can be easily published, it cannot serve as a common basis of expectations on which our practices of holding responsible depend. Real-world moral authority requires that moral address proceed in terms of norms that are part of the recognized social mores or practices of the moral community. And unless a norm can be justified in the access internalist sense,

members of the moral community will lack sufficient reason to accept and comply with the norms that seek to bind them. Without justification as a requirement for obligation, we could be subject to obligations based on easily correctable errors in reasoning and informational omissions. I submit, then, that any plausible second-personal explanation of moral obligation must affirm two individually necessary and jointly sufficient conditions for having an obligation: the norm on which the obligation depends must be (i) justified and (ii) easily published.

These two conditions mandate a further constraint on a plausible second-personal explanation of moral obligations, specifically a constraint on the individuation of norms that ground moral obligations. The norms that ground obligations can be individuated in terms of principles, rules or acts. We can set an act-individuation standard aside, as it is ruled out by the contractualist aspects of second-person normativity as based on maxims or act-types. On contractualist accounts, judging whether we are obligated to act in some way always depends on appealing to some maxim that describes which type of act is permitted, required or forbidden. However, obligations can neither depend on sophisticated philosophical judgments of the sort required to apply abstract principles to real-world situations. We need the norms on which moral obligations depend to be somewhere in the middle, perhaps based on rules. Rules have fallen out of favor in contemporary moral philosophy, but a number of prominent moral philosophers have placed rules at the heart of moral practice, including Kurt Baier, H.L.A. Hart and P.F. Strawson. Gerald Gaus has recently built on their efforts by arguing that only ‘middle-level’ social rules can solve coordination problems and produce stable expectations. For only rules can provide ‘guidance which is general enough to apply to unforeseen future circumstances while specific enough that we have common understandings of what the rule requires.’ In contrast with principles, rules ‘regulate the specific through the general’ by specifying instantiation relations for particular acts of moral address and moral behavior (Gaus 2011, pp. 122–123). Briefly, a plausible account of second-personality must appeal to *justified and easily published rules*. It is fortuitous that such an account is already on offer.

## Social-Moral Rules and the Rule-Centric Approach to Second-Personality

I base the rule-centric approach on the idea of a social-moral rule. This account of moral rules is in turn based on what Gaus calls the ‘Baier-Strawson’ conception of social morality, following Baier and Strawson (Gaus 2011, p. 4). A social morality is ‘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’ (Gaus 2011, p. 2). These social-moral rules have six key features: (1) they are *moral* rules (2) required for social coordination, (3) socially recognized, (4) enforced with social sanction, (5) reciprocal or universalizable, and (6) that are taken to override other reasons for action.

- (1) Let us define a moral rule as a rule that specifies features of actions and attach deontic operators (must, must not, must allow) to them. 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. Social-moral rules are thereby imperatival, for they provide 'the basis for issuing demands on others that they must perform certain actions' (Gaus 2011, p. 6). Thus, what makes the rule moral is first that it has categorical, rather than hypothetical force. Moral rules are also distinct from other types of rules because they license blame and resentment when violated, a feature we shall explore further below.
- (2) Social-moral rules also 'structure social interaction' and 'allow us to live together in cooperative, mutually beneficial, social relations' (Gaus 2011, p. 4). They do so by 'constrain[ing] our choices about how to pursue [our ideals]' (Gaus 2011, p. 6). Accordingly, they help 'individuals with diverse ends to live together' by maintaining stable expectations. Strawson emphasizes that social order can only exist when 'certain expectations of behaviour on the part of its members should be pretty regularly fulfilled' (Strawson 1974, p. 34).
- (3) 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 (Baier 1954, p. 108). 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' (Strawson 1974, p. 36). The recognition of social-moral rules resembles Hart's conception of social rules, which require a social convergence in judgment where community members 'must look upon [and endorse] the behavior in question as a general standard to be followed by the group as a whole,' along with observed collective behavior where individuals generally follow the rule and hold one another accountable to it (Hart 1961, pp. 54–56).<sup>5</sup> Following Hart, social-moral rules have an internal aspect where the rule can only hold when at least some community members recognize it as binding.
- (4) Social-moral rules must also be enforced with social sanction. Baier argues that moral rules must be 'supported by ... characteristically moral pressure' (Baier 1954, p. 108). 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' (Strawson 1974, p. 38). Similarly for Gaus, social-moral rules ground moral imperatives that license blame or sanction when they are violated, as they give persons authority to direct the action of their fellows (Gaus 2011, pp. 8–9, 193–201).
- (5) Social-moral rules must also be reciprocal in the sense that they apply equally to those who demand compliance with them: when John demands that Reba

<sup>5</sup> For Hart, only public officials accept rules of recognition.

not violate a rule, he must also regard himself as subject to the rule (Darwall 2006, p. 23).<sup>6</sup> Strawson insists that social morality cannot exist without ‘reciprocal acknowledgment of rights and duties’ (Strawson 1974, p. 40). A universal feature of morality is ‘the necessary acceptance of reciprocity of claim.’ Social-moral rules are universalizable in this sense. Baier claims that moral rules are ‘universally teachable and therefore universalizable’ in virtue of the fact that they license claims against all members in the moral community (Baier 1954, pp. 108, 111–112).

- (6) Finally, for Baier and Strawson, moral rules must generally override other reasons for action. Baier notes that ‘moral reasons are superior to all others’ (Baier 1958, p. 148). 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’ (Baier 1958, pp. 149–150).

The Baier-Strawson conception of social-moral rules is strikingly similar to Darwall’s *SPS*. The second-point standpoint consists in the exchange of recognizable moral reasons. Second-personally justified norms must be public in order to coordinate behavior and maintain stable expectations. Accordingly, such norms must be in some broad sense socially recognized in that they can be easily published as the norms of the moral community in question. Failure to comply with second-personal reasons licenses social sanction and the negative reactive attitudes. And the *SPS* requires recognizing relations of reciprocity. Finally, and obviously, the *SPS* justifies norms whose moral force overrides the normative force of other reasons for action.

The main difference between Darwall’s view and the Baier-Strawson conception of social-moral rules is that on the latter view social-moral rules mediate moral address. Darwall applies norms to persons via a direct exchange of reasons, whereas the Baier-Strawson view applies second-personality to persons by means of justified rules.<sup>7</sup> On the Baier-Strawson view, if John blames Reba for stepping on his toe and causing him pain, he blames her for failing to comply with social-moral rules that prohibit her from stepping on his toe and causing him pain. Social-moral rules mediate moral address. Thus, while the social-moral rule conception of second-personal normativity differs in important respects from Darwall’s view, it retains critical common features. As a result, it would be mistaken to claim that second-personality is rendered non-basic or trivial in the social-moral rule account of moral obligation, as I am only arguing that second-personality is realized in a complex social order of recognized rules (Darwall 2006, pp. 310–313).<sup>8</sup> So we can now

<sup>6</sup> Darwall acknowledges this point, but does not apply it to moral rules.

<sup>7</sup> Importantly, Darwall intends his second-person standpoint to provide a foundation for Kantian contractualism, so he could be open to my approach. However, his aim of grounding contractualist theories like Scanlon’s suggest that he does not intend to ground second-personal obligations in social-moral rules. I thank an anonymous referee for this point.

<sup>8</sup> Justifications for social-moral rules can be agent-neutral if acceptable to all.

distinguish between two conceptions of second-person normativity, a *reason-centric* account and a *rule-centric* account.

Before explaining why the rule-centric account of second-person normativity is superior to the reason-centric view, we must dispel a worry about the former. One attractive feature of Darwall's reason-centric account of second-person normativity is that it employs direct moral address in explaining how persons provide their fellows with reasons to treat them with respect. But I have argued that second-personal address is mediated by moral rules. It appears, then, that parties to the *SPS* respect moral rules rather than persons. Justified rules prescribe second-person reasons, not persons themselves. It may appear that something essential to second-person normativity has been lost—direct moral address that generates reasons for action. This criticism can be answered by considering an analogy with the rule of law. In liberal constitutional democracies, citizens typically pride themselves on being ruled by laws rather than persons. Through the application of impartial constitutional rules, citizens retain their status as free and equal rather than becoming subjects of arbitrary power and domination. In one way, then, citizens could be said to respect the law, first and foremost, and not their fellow citizens. But in this case, the reply is obvious. We respect our fellow citizens by complying with the rule of laws that we believe are justified to all and altering those that are not. The rule of law is not an obstacle to treating others with respect but rather a necessary means by which we do so. Social-moral rules can be understood in the same fashion. Moral rules are not an obstacle to treating others with respect but rather a necessary means by which we do so. Persons are still the source of second-person normativity even if that normativity is mediated by rules.

## How Rules Resolve the Problem of Publication

The rule-centric approach to second-personal normativity is less vulnerable to worries raised by hypothetical interpretations of the reason-centric approach because rules can plausibly connect justification and easy publication. This is one of the rule-centric approach's chief merits.

Let's illustrate with an example, *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 one's 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 high beams 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 public 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 do not like being honked at because we know that others resent 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.

*Driver John* has the recognitional and justificatory elements I have outlined. The justificatory component of *Driver John* is rooted in the justification of the rules of the road. John may fail to follow traffic rules when he is late for work and he may complain that a speed limit is set too low. But John is nonetheless rationally committed to holding that the rules of the road make everyone better off. Traffic laws can be binding even when no one actually accepts or endorses them because the rules are justified for them despite their actual behaviors.<sup>9</sup> Consequently, second-personal justification is not held completely hostage to real-world irrationalities or misinformation. The recognitional element in *Driver John* is the easy publication of the rules of the road. Nearly all drivers recognize the rules, having learned them in a class or from experience. Others will learn the rules from legal penalties from the police or moral sanction from their fellow drivers. Much like tomorrow's weather forecast, traffic laws are easily published.

### The Rule-Centric View and Societal Blind Spots

Perhaps the most pressing objection to the rule-centric view is that the easy publication requirement cannot account for sincere and informed moral blind spots (Gaus 2011, pp. 428–434).<sup>10</sup> Consider the case of a slave owner's obligations in a society where slave-ownership is deeply socially entrenched. The slave owner may wholeheartedly believe that his slaves have little moral worth and are thereby the proper subjects of enslavement. In this case, we assume, slavery-prohibiting norms are not easily publishable to him. Yet it may seem that this epistemically impervious slave owner is still obligated to release his slaves because slavery is wrong regardless of whether a slave prohibiting social-moral rule is easily published.

I offer eight related replies. (a) The rule-centric view contains restrictions on which rules can count as genuinely *moral* rules. Moral rules must, for instance, be reversible, applying reciprocally to all members of the moral community. So insofar as the slave-owner sees his slaves as subject to moral obligations, and wishes to hold them accountable, he is committed to treating them according to reversible moral rules. Reversibility will likely imply that the social-moral rule that permits him to forcibly keep his slaves is not a *bona fide* moral rule and so cannot generate public moral permissions or duties. To avoid the trap of reversibility, a slave-owner must be prepared to entirely deny that slaves are moral agents, which means that he cannot rationally blame or resent them for disobeying orders or trying to escape. He must regard them *purely* as animals; he must be *maximally* inegalitarian. But this

<sup>9</sup> Formal legal authorities typically pass traffic laws, so in that sense someone has to endorse them. However, law and convention frequently arise without anyone intentionally ratifying them.

<sup>10</sup> W.D. Ross's famous critique of T.H. Green's 'rights recognition' thesis is perhaps the most famous criticism of attempts to tie actual recognition to obligation. See Ross (1930, p. 51).

attitude is not common among historical slave-owners, as they will typically insist that the slave has moral agency, at least in order to keep them subservient.<sup>11</sup>

But let us suppose that we restrict our focus to slave-owners who regards his slaves purely as animals, that is, maximally inegalitarian slave-owners. Surely it is a problem for the rule-centric view if the maximally inegalitarian slave-owner lacks an obligation to release his slaves. In response, I offer several points. (b) The rule-centric account of second-person obligation is just that, a theory of obligation. It is not a general theory of immorality or a theory of badness, harm or even duty. The slave owner, therefore, can be subjected to many forms of moral evaluation. We can say that he acts badly, that he has poor character due to oppressing his slaves and that what he does is all-things-considered wrong, and even that were he to start regarding his slaves as something more than animals, he would have an obligation to release them. (c) Remember that the slave owner's slaves will likely have second-person reason to reject slave-permitting rules. If so, they have no obligation to comply with the commands of their master. So even if a maximally inegalitarian slave owner has no obligation to release his slaves, the slaves will have no obligation to comply with his demands. (d) Real-world slave owners' beliefs are not well grounded. This seems especially true for slave-holders we blame most, such as those in the antebellum American South, say, in contrast to slave holders several millennia ago. In the American South, slave-holding norms were already controversial, subject to criticism and not part of the mores of the group and so the purported justification of slave-permitting beliefs was likely illusory. We blame these slaveholders for failing to release their slaves because we believe they knew better. We resent them because we believe they had reason to reject slave-permitting norms and so had no good reason to keep their slaves in chains. Of course, contrary cases are conceivable, but if my account of obligation allows us to say that real-world maximally inegalitarian slave owners are obligated to release their slaves, the force of the slave owner counterexample is diminished.

(e) In cases where the maximally inegalitarian slave owner's belief in the moral inferiority of her slaves is unfounded, alternative social-moral rules, such as those providing for a minimal respect for autonomy or a presumption against harm, may be sufficiently publishable and justified to render the slave owner obligated to release her slaves. That is, were her slavery-permitting beliefs defeated, other social-moral rules will provide her with the slave-releasing obligation. So there are considerable rational resources available to ground a slaveholder's obligation to release her slaves even in hypothetical cases. In this way, even many hypothetical slave owners will have obligations to release their slaves. My account of second-personal obligation and publication only fails to show that *all* hypothetical, maximally inegalitarian slave owners have obligations to release their slaves. But, (f) if the reader insists that a theory of obligation must nonetheless show that *all* mentally component, adult slave holders, in all possible worlds sufficiently similar to our own, have an obligation to release their slaves, no matter their culture, background, time period or social status, then I must simply disagree. It seems to me

---

<sup>11</sup> I thank an anonymous referee for helping me to see that the reversibility requirement can threaten even the epistemically impervious slave-owner's authority.



that we cannot sustain our reactive attitudes against epistemically impervious, maximally inegalitarian slaveholders and this indicates that in practice we would not believe them to be under a moral obligation to release their slaves. Supposing there have ever been any invincibly ignorant and maximally inegalitarian slaveholders, I cannot see that a theory of obligation stands or fall based on whether we can say that all possible slave-owners have an obligation to release their slaves. Our anti-slavery intuitions are robust across many environments, but I do not think our intuitions must have these broad modal implications.

(g) To make this point more vivid, imagine a slave-owner who strongly believes his slaves are robots. He does not believe that his slaves are agents at all, but rather effectively disguised machines. Does such a slave-owner have an obligation to release his slaves, given that he believes, resolutely, that they are, in fact, robots? It seems to me perfectly natural to say that he has no obligation to release them because he thinks they are robots. The discomfited reader may respond by arguing that the more natural interpretation of our slave-owner is that he has an obligation to release his slaves but that his epistemic defects *excuse* him from violations. In that way, one could acknowledge that we cannot sustain our reactive attitudes against epistemically impervious slaveholders but that they nonetheless have an obligation to release their slaves. But this reply seems most natural in cases where someone is already rationally committed to the obligation *generally*, but meets exculpatory conditions for a particular violation in particular circumstances. For instance, John might promise to take Reba to the park, but have to break the promise to take his mother to the hospital. In that case, John is excused from violating his obligation, but his obligation to keep his promises is one he is already rationally committed to. It is less natural to characterize someone as excused from violating her obligation when she is completely epistemically impervious to all sound justifiers for the obligation, such as in cases where a slave owner believes his slaves are robots.

Given the foregoing, I think it fair to conclude that while the rule-centric approach withstands the objection from moral blind spots.

## A New Conception of Second-Personal Reasons

A critical benefit of the rule-centric 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 within second-personal relations of accountability and authority. But this is too vague. Many reasons can be issued in a context where persons share a moral community and second-personal moral practices. For instance, I might appeal to a third-personal reason that pain is agent-neutrally bad when you step on my foot even though we bear second-person relations to one another. Darwall somewhat surprisingly allows that reasons with third-personal content can be second-person reasons if they are issued within a second-personal context. For instance, second-personal contractualism might allow selection of rule-consequentialist standards for evaluating moral norms (Darwall 2006, pp. 312–313). In this case, the reasons issued by rule-consequentialist standards would count as second personal because the moral community endorses

those standards. Yet the idea of a second-personal reason now seems extraordinarily broad. A second-person reason is just any reason that contractors accept as reasons. The distinguishing features of second-person reasons are lost.

In response, I suggest that the second-personal aspect of a reason should not be identified with the social relations where the reason is presented but with the reason's relation to a social-moral rule. To illustrate the distinction, 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, we can interpret her action as issuing John a second-person demand that in turn generates a reason for action for John.<sup>12</sup> But the reason is rather simple; it is 'Do not violate the no cutting rule!' when the rule is relevant to the context. Thus the second-personal reason derives its second-personality from being associated and tied to the social-moral rule in question. In this way the rule-centric conception of second-person reasons holds that reasons are second-personal when they partake 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 make up the social-moral rules that bridge each person's values and create a second-personal world of moral relations. I can now define a second-person reason:

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

I take a moral reason for A to act on R to be one that morally requires A to act on R. There is a distinctively moral 'ought' or 'requirement' relation such that A commits some moral infraction when she fails to comply with R. Given this definition, not all moral reasons are second-personal in Darwall's sense. But all second-person reasons are moral reasons, satisfying condition (i). More specifically, 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 justified 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 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.<sup>13</sup> Condition (ii) is satisfied. In conclusion, R has second-personality in two senses: it relates to demands made upon others and 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 already: reasons to demand and reasons to comply with demands. But we also have second-personal reasons to modify or reject other second-personal rules. These second-

<sup>12</sup> *Driver John* may involve bipolar obligations but my account of second-person reasons does not concern them.

<sup>13</sup> These beliefs need not be occurrent.

personal reasons concern the practices of moral reform, of altering, creating or removing a social-moral rule. Despite their more abstract formulation, practices of reform and alteration are still grounded in relations of accountability. If I am out to change a norm, I must do so in second-personally justified ways.

## Conclusion

A brief summary of our conclusions: Darwall's reason-centric account of second-person normativity does not adequately distinguish recognition and endorsement, and so gives hypothetical interpretations of both. Doing so, however, arguably deprives second-personal normativity of much of its real-world force. For it is hard to see how norm only hypothetically recognized can be publicized and used to generate public expectations. The rule-centric account of second-personal normativity avoids this confusion by appealing to rationally justified and easily published social-moral rules. Second-person reasons, then, derive their normative force from the justified rules to which they are attached.

Looking ahead, I believe the rule-centric account of second-person normativity helps to distinguish domains of inquiry into the nature of second-person normativity. First, research can focus on the nature of the real-world practices whose authority derives from actually recognized rules. Philosophers might ask how second-personal social practices bind the wills of others via the formation and implementation of rules. Second, research can focus on the nature of social-moral rules themselves. In what way are they publicly recognized? How are they sustained through real-world practice? How do they authorize the reactive attitudes? Finally, moral philosophers might wonder how exactly to understand hypothetical endorsement. Moral philosophers can now focus on three discrete areas of research, all part of a renewed research project within deontological and contractualist ethics.

**Acknowledgments** This piece has been a long time coming, and many people have helped along the way. I am especially grateful to Jerry Gaus, Stephen Darwall, Connie Rosati, Mark LeBar, Jada Twedt Strabbing, Kyle Swan, Matthew Smith, Alex Worsnip, a host of anonymous referees, and a number of participants in Darwall's seminar on moral recognition held at Yale in the fall of 2011.

## References

- Baier, Kurt. 1954. The point of view of morality. *Australasian Journal of Philosophy* 32: 104–135.
- Baier, Kurt. 1958. *The moral point of view: A rational basis of ethics*. Ithaca: Cornell University Press.
- Chisholm, Roderick. 1989. *Theory of knowledge*. Upper Saddle River: Prentice Hall College Division.
- Darwall, Stephen. 2006. *The second-person standpoint: morality, respect, and accountability*. New York: Harvard University Press.
- Darwall, Stephen. 2007. Reply to Korsgaard, Wallace, and Watson. *Ethics* 118: 52–69.
- Darwall, Stephen. 2010a. Authority and reasons: Exclusionary and second-personal. *Ethics* 120: 257–278.
- Darwall, Stephen. 2010b. But it would be wrong. *Social Philosophy and Policy* 27: 135–157.
- Darwall, Stephen. 2010c. Precis: The second-person standpoint. *Philosophy and Phenomenological Research* 81: 216–228.
- Darwall, Stephen. 2010d. Reply to Schapiro, Smith/Strabbing, and Yaffe. *Philosophy and Phenomenological Research* 81: 253–264.

- Darwall, Stephen. 2012. Bipolar obligation. In *Oxford studies in metaethics*, vol. 7, ed. Russ Shafer-Landau, 333–356. Oxford: Oxford University Press.
- Gaus, Gerald. 1996. *Justificatory liberalism*. New York: Oxford University Press.
- Gaus, Gerald. 2011. *The order of public reason*. New York: Cambridge University Press.
- Hart, H.L.A. 1961. *The concept of law*. Oxford: Clarendon Press.
- Locke, John. 2003. In *Two treatises of government and a letter concerning toleration*, ed. Ian Shapiro. New Haven: Yale University Press.
- Ridge, Michael. 2003. Contractualism and the new and improved redundancy objection. *Analysis* 63: 337–342.
- Ross, W.D. 1930. *The right and the good*. Oxford: Clarendon Press.
- Scanlon, Thomas. 1998. *What we owe to each other*. Cambridge: Harvard University Press.
- Smith, Michael, and Jada Twedt Strabbing. 2010. Moral obligation, accountability, and second-personal reasons. *Philosophy and Phenomenological Research* 81: 237–245.
- Strawson, P.F. 1974. *Freedom and resentment and other essays*. New York: Routledge.
- Wallace, R. Jay. 2007. Reasons, relations, and commands: Reflections on Darwall. *Ethics* 118: 24–36.
- Watson, Gary. 2007. Morality as equal accountability: Comments on Stephen Darwall's the second-person standpoint. *Ethics* 118: 37–51.
- Williams, Bernard. 1981. *Moral luck*. Cambridge: Cambridge University Press.