

THE SOURCES AND STANCES OF MORAL NORMATIVITY

CARLA BAGNOLI

*Department of Education and Human Sciences
University of Modena and Reggio Emilia
carla.bagnoli@gmail.com*

ABSTRACT

This paper argues against the view that the issue of moral normativity is best accounted by undertaking the first-person deliberative perspective. Section 1 characterizes Korsgaard's self-constitutivist view of moral normativity in contrast to skepticism and contractualism. Section 2 highlights the role of the value of humanity in the self-constitutivist view. Section 3 formulates an issue about the sources of moral obligations to others, which points to a tension within Korsgaard's theory of moral obligation. Sections 4-5 show that the dominance of the first-person deliberative stance in accounting for moral normativity is related to the deployment of the strategy of reflective endorsement, which is not functionally equivalent to the self-constitutivist strategy for vindicating moral authority. Section 6 argues that endorsement understood as an act of imaginative rehearsal fails to carry out the main insights of Kantian constructivism regarding normative discussion and the transformative potentiality of practical reasoning. Section 7, defends the importance of multiple stances to do justice to the complexity of moral normativity.

KEYWORDS

Morality; normativity; obligation; Kant; Korsgaard.

Skeptics have raised worries about the categorical authority of moral judgments and, especially, those judgements that establish the claims of others under the guise of obligations.¹ They have questioned not only the psychological force of such obligations to overcome the agent's attachments to the 'dear self', but also their normative authority. They have doubted their overridingness, that is, their authority to systematically defeat the competing claims of personal desires and concerns, and constrain practical deliberation so as to undercut the normative reasons for action that spring from non-moral projects. An implicit matter of contention is whether non-moral projects can be considered truly categorical insofar as they are self-defining; if they can, then they have rational authority in deliberation, and then it would be paradoxical to consider moral obligations overriding: the conflict between self-defining non-moral projects and moral obligations would become impossible to assess from a unified agential stance (Williams 1981, pp. 20-40, 114-132; Williams 1985, Chapter 4, 10, Frankfurt 2004, 2006).

¹I use the term "skepticism" to identify the meta-ethical positions which denies the objectivity and categorical authority of moral obligations, see e.g., Mackie 1977, Williams 1985, Frankfurt 2006.

In response to these challenges, some Kantian moral theorists have attempted to show that morality is rooted in some basic features of (rational) agency, and thus moral concerns permeate and spread through the entire architecture of rational deliberation, which designs the contours of the self. This is the “self-constitutivist view” of moral authority. On this view, moral obligations are normative rational requirements, and their authority is first-personal, because and insofar as humans are rational agents. If rationality is a constitutive feature of human agency, one can explain that moral obligations bind human agents with inescapable authority by arguing that they are rational requirements. In virtue of the unifying function of rational norms, the self is not merely a bundle of disparate perceptions and disorderly incentives but a unified whole. Insofar as they are rational requirements, moral obligations do not only serve as constraints in deliberation, but also and more fundamentally as unifying forces which shape the agent’s self. Thus, human agents are *internally* bound by moral obligations, i.e., in virtue of normative standards that pertain to their nature as rational agents, rather than externally, via sanctions and incentives. Correspondingly, moral norms are modes of self-organization and self-constitution.

The most sustained and ambitious project of this kind has been magisterially defended by Christine M. Korsgaard in *The Sources of Normativity* (1996), and further elaborated in her *Self-constitution* (2009). Korsgaard’s argument aims to vindicate both the categorical and first-personal authority of moral obligations by grounding them in agential features, which cannot be ignored without that humanity itself be destroyed. To this extent, being moral becomes a matter of personal integrity. The thesis that the self is in charge of its own constitution explains how and why moral action is not an external imposition, but it is anyone’s personal responsibility.² If the argument stands, the self-constitutivist view provides a definitive answer to the question why be moral: being moral is being one’s own self: “Your reasons for acting express your identity, your nature; your obligations spring from what that identity forbids” (Korsgaard 1996a, 101). At least, this is the promise of the self-constitutivist account.³

Many think that this promise is not fulfilled, or that it cannot be fulfilled unless the self-constitutivist commits to the reality of some values, e.g., the value of humanity. In absence of a realistic foundation of the process of self-constitution, it can be objected that the selves may not have an inalienable concern for being self-constituted agents; some may be happy to be just “shmagents”, hence failing Korsgaard’s standards of autonomous agency (Enoch 2006). The ensuing debate has focused on the meta-ethical implications of the “shmagency objection” for a thoroughly constructivist meta-ethics.⁴ This essay departs from this debate because it focuses on an under-discussed aspect of the self-constitutivist strategy in support of the objectivity and categorical authority of moral obligations, that is, the dominance of

²For Korsgaard (2019) this is a merit of her view in contrast to Aristotelian forms of constitutivism. For a survey of other variants of constitutivism, see Katsafanas 2018.

³A source of dissatisfaction concerns the characterization of bad or non-moral actions. Such actions are not completely autonomous insofar as they are not endorsed in full reflection. Ginsborg (1998) argues that Korsgaard (1996) treats actions based on reasons that spring from non-moral identities as “not free”. Tenenbaum (2011) reiterates the point against Korsgaard (2008). Cf. also Gobsch 2019.

⁴On the damage done by the shmagency objection cf. Tenenbaum 2011.

the deliberative perspective. Korsgaard herself insists that the issue of the objectivity of moral obligations arises in the first-person deliberative perspective (Korsgaard 1996a, Korsgaard 2021, 178). Realists and constructivists alike take for granted that the capacity to make sense of the normativity of moral obligations in the deliberative perspective represents a decisive test for any ethical theory; and critics of Korsgaard largely concede that it is a merit of the self-constitutivist view that it accounts for the inescapable authority of moral obligations, and their overriding status in deliberation.⁵

My aim is to challenge the view that the issue of normative authority is best accounted by focusing on the first-person deliberative perspective. I take this challenge to show that the strategy based on reflective endorsement does not support the self-constitutivist view.⁶ More broadly, the argument points out that different stances in addition to the first-person deliberative stance may be relevant to highlight the complex phenomena of moral normativity. In this regard, the unchallenged dominance of the first-person deliberative perspective on the authority of moral obligations is an unwise methodological choice, insofar as it conceals relevant aspects of the normative authority of morality, which are not intercepted by the current debate, and deserve to be clearly distinguished.⁷

1. THE SELF-CONSTITUTIVIST VIEW OF MORALITY

Morality names a wide cluster of phenomena, encompassing not only moral obligations but also ideals, conceptions of the good life, customs, habits, and norms that arise in specific circumstances at specific times and at specific cultural coordinates. In current debates, morality is understood narrowly, as the domain of what we owe to each other. But the question of what exactly we owe to one another is not completely insulated from normative ideals of rational agency, and so the attempt to draw a sharp line among different clusters of questions about morality – narrowly or broadly understood – may be objected as stipulative and question-begging. The self-constitutivist view argues for a normative ideal of agency that provides a direct link from moral obligations to normative reasons for action. It makes moral obligations normative in the peculiar sense that they are requirements of rationality allowing for a well-functioning agency. One central argument in support of such a constitutive relation invokes the value of humanity. However, ‘humanity’ is an evaluative and not purely descriptive concept, and the scope of moral obligations to humanity varies

⁵“Korsgaard’s view can explain why, in a Kantian view, human beings ought to follow the categorical imperative even when the categorical imperative commands them to act against strong inclinations. [...] because only when reason is heard, does it make sense to say that there is a whole and I have acted”, Tenenbaum 2011, 453-454. Shafer-Landau 2003; Enoch 2006, 195-96.

⁶This is not the objection pressed by Derek Parfit that Korsgaard’s argument trades on an equivocation between the motivational and the normative force of moral obligations, Parfit 2006, 351-80. Nor that the argument ignores the distinction between first-person practical deliberation and detached philosophical reflection, Smith 1999.

⁷Such a debate has been primarily concerned with whether constructivism has a theoretical advantage over realism regarding the account of normativity in the deliberative perspective. This is how the question is framed by Korsgaard, and engaged by realists, Shafer-Landau 2003; Enoch 2011; Stern 2013.

according to the bounds associated to it. Correspondingly, the view that moral obligations to humanity claim universal authority should answer a series of thorny issues: who are the relevant ‘others’, and why are individual subjects bound by their claims?⁸

The self-constitution view holds that the norms constitutive of rationality are also norms of *moral agency* that apply to and bind all agents as such. This view entails that moral obligations are rational requirements, but also that the demands of morality lie at the core of the self, rather than being external and contingent constraints which the self is at liberty of disregarding. Such constraints are inescapable, that is, not optional and dispensable, in that they can be occasionally violated but not completely and systematically disregarded without that the self collapses and any agential perspective goes lost (Korsgaard 1996a, pp. 100-101). Moral action is the perfect principled mode of self-constitution, and its violation is a variety of self-alienation. The constitutive and binding force of the norms of rationality explain at once the metaphysical and the moral integrity of the self. The self constituted by the exercise of rational norms is said to have a radical and unremitting concern with morality. This is not to say that the self is ‘necessitated’ by specific moral norms. In fact, such norms can be violated in different ways and degrees. However, insofar as human agents are guided by norms of rationality in their self-constitution, they are sensitive to the demands of morality. They can refuse to comply with such demands, but they would do so at the expense of themselves, since moral transgressions generate a loss in the unity and rational organization of the self, and thus count as self-alienation. To this extent, the self-constitutivist view purports to offer a definitive argument against *moral skepticism*.

But it also promises to account for the authority of moral obligations in a way that is alternative and superior to others objectivist accounts such as, specifically, *contractualism*. Contractualism treats moral obligations as ideally or hypothetically contracted, and identifies their normative source and authority with an agreement settled under specified conditions (i.e., unforced and informed). As a result, contractualism approaches morality in terms of relations of mutual accountability, which are based on reciprocity and mutual recognition. For instance, Thomas Scanlon holds that “An act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behavior that no one could reasonably reject as a basis for informed, unforced, general agreement” (Scanlon 1998, p. 153). This formula relates moral obligations to moral reasoning, and thus rests on larger claims about reasons and viable forms of reasoning. The authority of moral standards is explained in terms of the relation of mutual recognition; and wrongness consists in unjustifiability: wrongness *is* the property of being rationally unjustifiable. Scanlon’s theory of obligations vindicates a stable and robust relation between the domain of moral obligation and rationality. There is an important sense in which the account of morality as mutual accountability establishes an *intrinsic* relation between moral obligations and reasons for action. The normative force of moral obligations

⁸The question of the scope is faced upfront by O’Neill 1996. She holds that Kant’s claim about the scope of moral obligations ranging over humanity differs from the contractualist account of moral obligations, see O’Neill 2003. Korsgaard seems to hint at this distinction when she distinguishes between the categorical imperative, which imposes universality; and the moral law, which is meant to capture the idea that moral reasons are public, in that they address all others.

can be explained in terms of reasons for action that derive from the hypothetical stipulation of a contract with one another. But to say that agreements and contracts produce normative reasons for action (i.e., in fulfillment of the respective obligations) does not mean that moral obligations are requirements of rationality as such.

On this premise, the contractualist view allows for different sets of normative requirements and obligations, depending on the scope of the agreement.⁹ In the case of duties toward fellow-citizens, the source is social and specific, that is, relative to a specific historically given ‘civil society’. In the case of duties toward humans, the source is the value of humanity under a given characterization or an idealized and generalized conception of civil society. In all these cases, the interference of and to others is precisely what is regulated and governed by moral obligations. This may be considered an advantage of the contractualist formula, which appears to be flexible enough to capture different spheres of morality. But it then becomes problematic to understand how moral obligations protect humanity in itself. If the source of normativity is agreement, then the scope of obligations varies accordingly. If, instead, the agreement is meant to enforce respect and recognition of the value of humanity, then it is not the ultimate source of authority.

On the self-constitutivist view, moral authority has nothing to do with hypothetical agreement, but springs from a constitutive connection between agency, rationality, and morality. Skeptics cannot challenge the authority of moral obligations while holding on the deliberative stance (Korsgaard 2009, pp. 32-33). This inescapability is *deliberative*, not simply in the sense that it overrides other claims in deliberation, but also in the more basic sense which pertains to the organization of the stance of agency. If moral obligations rule out or curb some personal projects and aspirations, it must be because they are self-defeating, and thus cannot be transparently sustained by the agent’s self. While constitutive the norms of rationality are normative: they can be violated, but at risk of losing any agential perspective on what to do.

2. SELF-CONSTITUTION AND THE VALUE OF HUMANITY

In order for the constitutive norms of rationality to play such a large role in fixing the stance of agency and, at the same time, in grounding moral obligations, they must be both formal and normative. On the Kantian view, they are norms of universality. But their relation to the value of humanity, and correspondingly the role of this value in the argument about self-constitution are hard to pinpoint. Korsgaard’s responses to these central issues have taken different shapes over the years.¹⁰ First, in commenting Kant’s formula of humanity, she holds that the rational nature of humanity commits rational agents to consider their ends as good. Since such ends are intrinsically valuable, they are also committed to a further ultimate value, and this value is humanity. Humanity is thus a value-conferring

⁹On the distinction between contractualism and constructivism, see O’Neill 2003.

¹⁰Korsgaard 1986, 1996a, 2021.

property, a capacity in virtue of which valuers attribute value to things and persons. Furthermore, valuers are committed by rationality to recognize a fundamental value in themselves insofar as they are setting their own ends. Korsgaard argues that the duties to humanity can be derived from the capacity for rational choice (Korsgaard 1996a, pp. 118-123; cf. Kant 4.437, 6.387, 6.392).

A major difficulty in this argument is that one cannot infer that humanity is valuable in itself merely because it is a condition of valuing (Barandalla & Ridge 2011). However, in a more recent essay, Korsgaard argues that this presumption is not necessary to her argument.¹¹ She now thinks that the argument should be framed in more modest terms, so as to say that “valuing ourselves is a presupposition of valuing anything else”, and in any act of valuing, “we express the value we necessarily place on ourselves”.¹² In defending this understanding of how humanity is implicated in valuing, Korsgaard openly appeals to Thomas Scanlon’s view that the chief way to value others is to value and respect their choice and interest (Korsgaard 2021, pp. 171, 202; cf. Scanlon 1998, p. 105). However, she rejects Scanlon’s (realist) account of reasons and moral reasoning, and holds that this view is better vindicated by a constructivist meta-ethics. To appreciate the contrast between these meta-ethical views is best to set aside the ontological commitments associated with them, and focus on their respective account of the nature of reasons and of moral reasoning. According to realism, substantive reasons are independent of the requirements of rationality, and work as premises in moral reasoning (Scanlon 1998, pp. 25-32). On the constructivist view, instead, rational requirements are constitutive, that is, they do not work as premises but as organizing norms (2009b, p. 67; 1996b, pp. 321-325; Bagnoli 2011, 2013, 2021).¹³ For Scanlon, reasons as “relational” (2014, p. 120); and they are practically significant insofar as they bring us into an agreement with others (Scanlon 2014, p. 2). Scanlon’s emphasis on relationality is supposed to avoid the caricatural charge that moral realism presupposes the existence of normative properties “out there” and that we encounter reasons “wafting by” (Mackie 1977; Korsgaard 1996a, p. 44; Scanlon 2014, p. 120). Korsgaard is unconvinced by this response (Korsgaard 2021, pp. 171, 175), and I think she is right because the mere introduction of relationality does not resolve the main issue between moral realism and constructivism, since also constructivism takes moral normativity to be

¹¹“I now think that in order to make that part of my argument, I do not need to defend the general thesis that if something is the condition of all value it must be taken to be valuable itself.” (Korsgaard 2021, p. 171).

¹²“At the time I wrote “Kant’s Formula of Humanity” I was inclined to think of the argument as establishing that humanity has something like intrinsic value. Later I decided that on Kant’s view, all value must be conferred by valuing agents, or, to put it in a different way, that valuing is prior to value. So as I now read the argument, its point is that valuing ourselves is presupposition of valuing anything else. As I would now put it, in taking the things that are good *for* us to be good *absolutely* – good in the way that makes them worthy of rational pursuit – we express the value we necessarily place on ourselves” (Korsgaard 2021, p. 188 fn 3; cf. Korsgaard 1996b, chapters 7 and 10).

¹³“The alternative to substantive realism about reasons is the view that all substantive reasons are identified by the application of the categorical imperative, or some other principle of formal rationality, to which an agent is committed by virtue of his rationality, regardless of his own beliefs and commitments. In that case, we need only relativize his subjective reasons to his knowledge of the non normative facts” (Korsgaard 2021, p. 188 fn 10, 172 fn 8).

relational (Korsgaard 1996b, cf. Korsgaard 1996b, p. 275, chapt. 7 and 10). The key question is *what kind of relationality* pertains to moral normativity. The interesting part of Korsgaard's objection against moral realism or realism about reasons is that it takes the relations to reason to be epistemic (Korsgaard 2021, p. 175). Thus, we should expect that the purported advantage of constructivism concerns a different rendering of relational normativity, which is practical rather than epistemic. To explicate this contrast, Korsgaard presses the view that the practical standpoint is to be understood as the first-person deliberative standpoint.

This is an important piece in the entire argumentation because it allows Korsgaard to establish that the normativity of morality, unlike other forms of normativity, should be expressive of individual autonomy. To function as proper reasons for action, considerations must be such that they represent the stance of agency. The question of normativity arises for and matters to agents who have a distinctive kind of subjectivity, that is, capable of self-consciousness. Thus, the source of normativity is the capacity for evaluating: "whether you have reason *at all* depends not on what is out there in the world, but what sort of subjectivity you have. Reasons exist in the first instance in the deliberative perspective itself. But the kind of subjectivity that is needed is not the capacity to form an objective conception of reason: it is rather the kind that Kant associated with autonomy, the capacity to make a law for yourself. The essential element of a reason is its normativity, and its normativity for you rests in the fact that you legislate acting on it as a law", (Korsgaard 2021, p. 178; cf. 1996a, §3.4).

This is a juncture in Korsgaard's argument which has come under severe scrutiny by her critics. If reasons exist in the deliberative perspective, and they originate in the individual subjectivity, how are they responsive to the claims of others? Why should one respect humanity in others? There are significant strands of Korsgaard's work that directly point to a robust and plausible answer to this question: "The primal scene of morality - argues Korsgaard in *Creating the Kingdom of Ends* - [...] is not one in which I do something to you or you do something to me, but one in which we do something together" (Korsgaard 1996b, p. 275). The very purpose of construction is to find reasons that people can share. This is the beginning of a relational account of moral normativity, which is alternative both to the realist and to the second-personal construals of such relationality: "As persons, we have a claim on one another's help when it can readily be given or desperately needed. It is the status of humanity, as the source of normative claims, that is the source of all value. The argument, in other words, has brought us back to Kant" (Korsgaard 1996b, p. 299).¹⁴

While the relational strategy is a strength of constructivism about reasons, it does not operate in full force in *The Sources of Normativity*, whose main argument concerns the rational grounding of practical identities.¹⁵ Roughly, the argument goes like this. Humanity

¹⁴Some accounts are based on the second-personal account of the value of humanity, but even in this sense there are different accounts of relationality, see e.g., Darwall 2006, Zylberman 2017, De Magt 2018.

¹⁵Korsgaard (2021) marks the difference by noticing two conceptions of humanity, as a value property and as normative standing: "But I now think Kant means something a little different. He means that we regard humanity or the power to determine ends as a property that confers a certain kind of normative standing, and

gives an identity “simply as a human being, a reflective animal who needs reasons to act and to live” (Korsgaard 1996a, p. 121). Reflective agents need reasons in order to act or to live, and such reasons for action derive from particular practical identities, that is, conceptions of oneself under which humans value themselves (Korsgaard 1996a, p. 101). But any such particular practical identity is ultimately grounded in humanity (Korsgaard 1996a, p. 129). Thus, in order to value any specific practical identity, one must value one’s own humanity (Korsgaard 1996a, p. 123). To act at all (on the basis of reason), rational agents must value their own humanity (Korsgaard 1996a, p. 123). “To treat our human identity as normative, as the source of reasons and obligations, is to have... [a] ‘moral identity.’” (Korsgaard 1996a, p. 101). The resulting view is that “An obligation always takes the form of a reaction against a threat of a loss of identity”, (Korsgaard 1996a, p. 102).

3. THE CLAIMS OF OTHERS

Korsgaard’s critics have forcefully objected that the exclusive emphasis on the relation between that morality and agential integrity is misleading, in that it suggests that the normative source of moral obligations is the agent’s self, rather than the value of humanity (Cohen 1996; Williams 1996; Skorupski 2002; Wallace 2012, 2019, p. 46).¹⁶ They have urged that to properly value humanity requires us to identify the source of moral obligations elsewhere, in others, that is. The violation of humanity affects the agent’s integrity, of course; but the primary purpose of moral obligations is not to protect anyone’s integrity.

The matter is often put in terms of a radical and insuperable tension between different tasks of ethical theory. On the one hand, the view that moral obligations are self-constitutive rational requirements explains why they bind in the first-person and thus inescapable for individual agents deliberating about what to do. On the other hand, the competing accounts preserve the other-oriented aspect of moral obligations, by insisting that they are grounded on the value of humanity, quite independently of the agent’s appreciation of it in particular deliberative contexts. However, this way of presenting the contrast does not get us far enough. Korsgaard does not deny that humanity is the bearer of value. Indeed, she holds that “Valuing humanity in your own person rationally requires valuing it in the persons of others”, (Korsgaard 1996a, p. 121, cf. 2021). Her point seems to be that the value of humanity is explicable by examining what valuing entails, rather than by insisting that humanity is in itself valuable and should occupy the first premise of moral reasoning. The contrast between these two proposals is reflected in the kind of explanation offered: should the authority of moral obligations be explicated in terms of the value properties or in terms of the activity of valuing?

with it certain normative powers, on the being who has it”, (Korsgaard 2021, p. 183). This allows Korsgaard to consider humanity in relation to the membership in the Kingdom of ends (4.429). I doubt that Kant establishes this distinction, or that the distinction solves the problem that Korsgaard faces.

¹⁶A related concern is that Korsgaard’s account of moral obligations misunderstands the basic insights of Kant’s account of morality, which requires a re-orientation toward others, (Skorupski 1998; Bird-Pollan 2011, 377; cf. Bagnoli 2007).

My hunch is that the debate has not made much progress because of a failure to appreciate different *foci* of disagreement: the first concerns the roots of moral obligations, and the second their normative authority. Critics of Korsgaard are right that moral obligations toward others must take others as their ends, and so we cannot say that it is wrong to wrong others because wrongdoing undermines our integrity. However, once it is agreed that the root of moral obligations is not the preoccupation to preserve the moral integrity of any one particular agent,¹⁷ there is still an open question about why and how moral obligations to others become authoritative. Whether they are contracted or self-constitutive, why do moral obligations oblige us? The self-constitutive view purports to be at advantage in addressing this question because it traces moral obligations and the self to the same roots: they are a matter of construction rather than discovery, and they must be constructed under the normative guidance of the same constitutive norms of rationality.

On a Kantian view, it is certainly true that the wrongness of a wrong action – that is, an action that violates an obligation toward others – does not reside in the fact that it upsets the agent’s self or disfigures her integrity. But the fact that the transgressor profoundly disapproves of her own violation, and barely tolerates herself for having wronged someone, tells us something important about the subjective authority of moral obligations. It tells us that in violating others, disregarding duties toward them, rational agents do not only undermine and endanger others; they also undercut and undermine themselves. To this extent, the violation of moral obligation to humanity is ultimately a self-defeating move, though this is not to say that its wrongness consists in the fact that it is self-defeating.

Is the self-defeating feature part of the core meaning of wrongness? The wrongness of an action that wrongs others depends primarily on what it does to others. However, the impact that wronging someone has on the wrongdoer is not an accessory aspect of wrongness. This claim can be sustained and argued from a variety of philosophical perspectives, so it is neither peculiar nor distinctive to the self-constitutivist view. The contractualist can say that rational agents are mutually accountable in virtue of a hypothetical agreement based on mutual respect and recognition. In transgressing obligations toward others, wrongdoers undermine their relations to them. The latter consideration does not support the claim that the wrongdoer’s violation directly affects her own integrity, but it may suffice to say that wrongdoing affects the wrongdoer’s efficacy (and hence her agential normative capacities), by undermining the personal relations on which she thrives. The self-constitutivist view aims to establish a stronger connection, and say that wrongdoing is self-defeating. If the self-constitutivist argument is successful, then, any agent has a normative reason not to commit wrongdoing. Yet, this is not to say that the reason why wrongs are wrongs depend on the way they affect one’s agency and integrity. Likewise, on the view that moral obligations are demands that people legitimately make to one another, the fact that an act violates such a demand and so disrespects the authority that these demands presuppose *is* a normative reason not to perform the wrongful act; but it is additional to whatever features make the act wrong (Darwall 2013, p. 68). To conflate these two issues is to conflate the

¹⁷Cf. Korsgaard 1996a, p. 136, 1996b, p. 301, Korsgaard 1996b, pp. 188-224. Cf. Bagnoli 2017.

normative source of moral obligations with the subjective authority they gain in the first-person deliberative stance.

4. SELF-CONSTITUTION BY REFLECTIVE ENDORSEMENT

The failure to distinguish the issue of the normative sources of moral obligations from their first-personal authority in deliberation may be due to – or at least encouraged by – Korsgaard’s emphasis on “reflective endorsement” (Korsgaard 1996a). Korsgaard holds that reflective endorsement is the basic mechanism of agential authority, and ultimately vindicates moral obligations.

The normative authority of moral obligations is the result of a process of first-person authorization (Korsgaard 1996a). Such a process is composed of two subsequent stages: the first stage marks a detachment from the proposal of action, while the second stage transforms a proposal of action into a reason for acting. This mechanism also explains why rational agents feel the pull of moral obligations toward others: “there is a sense in which these obligations are real – not psychologically, but normatively. And this is because it is the endorsement, not the explanations and the arguments that provide the materials for the endorsement, that does the normative work” (Korsgaard 1996a, p. 257).

Korsgaard seems to assume that the argument based on reflective endorsement and the argument from self-constitution are ultimately equivalent, at least functionally, in that they both aim to establish that moral obligations are inescapable.¹⁸ However, the argument from reflective endorsement does not establish by itself that moral obligations are constitutive of the self; it only establishes that they are inescapable in first-person deliberation, at least on two conditions (i) *if* reflection is governed by the norm of universality, and (ii) *if* the norm of universality affords moral obligations. By itself the individual capacity to fully endorse or reject a proposal of action does not assure that moral obligations should have a unifying or grounding role in the constitution of the self.

Furthermore, this particular capacity (of endorsement) is neither necessary nor sufficient to establish the normative authority of moral obligations. It is not sufficient because endorsement by itself does not work as a moral filter for reasons for action; it only transforms incentives into subjectively authoritative reasons for action. This mechanism may suffice to explain why particular subjects find moral obligations authoritative in the first-person deliberation, but it does not explain why such moral obligations should have a legitimate claim on them. In other words, the appeal to endorsement addresses the issue

¹⁸The matter is complicated by the fact that the method of reflective endorsement is deployed together with a transcendental argument: “but this reason for conforming to your particular practical identities is not a reason that springs from one of those particular practical identities. It is a reason that springs from your humanity itself, from your identity simply as a human being, a reflective animal who needs reasons to act and to live. And so it is a reason you have only if you treat your humanity as a practical, normative form of identity, that is, if you value yourself as a human being”, (Korsgaard 1996a, p. 121). There is a scholarly disagreement about whether the transcendental argument succeeds in its task, and whether it is already a commitment to moral realism, see Stern 2013, Schafer 2019.

of deliberative authority, but it leaves the issue of moral authority unaccounted. But endorsement is also unnecessary, if the self-constitutivist view shows that moral obligations are inescapable in the first-person deliberative perspective.

One may take these critical remarks to show that reflective endorsement tacitly presupposes the value of humanity, rather than showing that such a value sets the condition of possibility of any singular act of valuing; and hence that any attempt to derive moral obligations from considerations of bare agency is misguided and question-begging; and indeed, this is how the debate has proceeded thus far (Skorupski 2002, Kerstein 2020). But I think the correct conclusion to draw from the critical remarks offered above is that the argument of reflective endorsement is not *functionally equivalent* to the argument from self-constitution. To hold that it is amounts to missing the crucial distinction between the normative source of authority of moral obligations and the sort of subjective authority at play in first-person deliberation.

Differently than the issue of the normative source, the issue of subjective authority does arise in the stance of deliberation. But it is questionable that deliberating agents are bound by moral obligations only insofar as they are capable of reflective endorsement. That is, the process of endorsement is successful in warranting normative authority only because and insofar as the reflective stance is governed by norms respectful and protective of humanity. However, if the self-constitutivist argument does its job, then the mechanism of reflective endorsement is an idle addition. For the self-constitutivist view establishes that there are constitutive norms of rationality that justify moral obligations as grounded on agential features, and *by this means* it purports to explain their normative and subjective authority.

In the next two sections, I point to alternative accounts of endorsement, which do not provide any support to the subjective authority of moral obligations in the first-person deliberative stance.

5. ENDORSEMENT AS A MODE OF CONSTITUTION OF THE AGENT'S STANCE

Not all constitutivist accounts of agency afford self-constitution, and not all self-constitutive accounts embark in the project of grounding moral obligations, even though they are accounts of how to shape one's self that aim at some forms of correctness.¹⁹ Neither of these forms of constitutivism aspire to vindicate moral obligations against skepticism. Correspondingly, they do not bear the burden of showing how the claims of others have inescapable normative authority over any one particular agent. For instance, Harry Frankfurt holds: "I do not believe that anything is inherently important. In my judgment, normativity is not a feature of a reality that is independent of us. The standards of volitional rationality

¹⁹Korsgaard (2019) argues that Kantian constitutivism presents an advantage over the Aristotelian varieties because the former account for self-constitutions, while the latter cannot, and thus fail to vindicate agential responsibilities toward oneself.

and of practical reason are grounded, so far as I can see, only in ourselves. More particularly, they are grounded only in what we cannot help caring about and cannot help considering important,” (Frankfurt 2006, 363). But this is not because self-constituting agents do not recognize normative standards of agency, as Korsgaard holds.²⁰ On the contrary, they do recognize the distinction between craving and endorsing craving as one’s own. The form of guidance implicit in the notion of caring is inherently reflexive, though not the result of rational deliberation. Reflective endorsement is the paradigmatic way of authorizing action and fixing the bounds of the self, but it does not commit to morality, nor is it governed by rational norms.²¹ Correspondingly, it marks the boundaries between rational agency and irrationality in a radically different way than Korsgaard’s proposal. As Frankfurt notes “Our response to sinners is not the same as our response to fools”, (Frankfurt 2006, 258).

6. ENDORSEMENT AS IMAGINATIVE REHEARSAL AND NORMATIVE DISCUSSION

It may be suggested that the notion of endorsement serves another purpose in Korsgaard’s account, akin to the expressivist model of normative acceptance, which helps elucidate the dynamics of rationality.²² According to Allan Gibbard, to call something “rational”, or to say that it is “the thing to be done”, is to express one’s acceptance of norms that permit it”, (Gibbard 1990, p. 7). His proposal is that we “fix on the dictum ‘To call a thing rational is to endorse it.’ And search for a sense of ‘endorse’ for which the dictum holds true”, (Gibbard 1990, p. 6). This expressivist notion of endorsement is governed by the requirement of consistency. Yet, interestingly, the requirement of consistency that bears on normative acceptance is not spelled out as a “mere” logical requirement. Rather, it is illustrated in terms of the constraints that it puts on a normative discussion: to accept

²⁰By contrast, according to Korsgaard, in absence of rationality, there would be no distinction between an incentive and the self, Korsgaard 2008.

²¹“Taking ourselves seriously means that we are not prepared to accept ourselves just as we come. We want our thoughts, our feelings, our choices, and our behavior to make sense. We are not satisfied to think that our ideas are formed haphazardly, or that our actions are driven by transient and opaque impulses or by mindless decisions. We need to direct ourselves—or at any rate to believe that we are directing ourselves—in thoughtful conformity to stable and appropriate norms. We want to get things right”, Frankfurt 2006, p. 74.

²²Korsgaard sometimes suggests that her activity of valuing shares some features with the expressivist agenda: “in taking the things that are good *for* us to be good *absolutely* – good in the way that makes them worthy of rational pursuit – we express the value we necessarily place on ourselves”, (Korsgaard 2021, p. 188 fn 3; 2008, p. 325). Some have suggested that expressivism may offer the appropriate semantics for constructivism. Constructivism agrees with expressivism about the naturalistic ontology, and about the view that practical deliberation is not about tracking ends that are real and mind independent objects. But there are notable differences between the two meta-ethics, which make them ultimately irreducible. Contrary to expressivists, constructivists defend the claim that there are normative facts or truths about the rational will and its activity: there is a normative realm, which is brought about through volitional activity. They share a practical conception of deliberative thought, but the constructivists hold that the normative claims are true in virtue of the commitments grounded in rational agency (cf. Wallace 2012).

a norm is to be prepared to avow it in a normative discussion. Normative discussion involves different parties, though such parties do not have to be real people. Gibbard construes normative discussion as a sort of “imaginative rehearsal”, which can be put forward by a solitary agent in a soliloquy about what to do, and serves the purpose of building the agential stance: “From this imaginative rehearsal, then, a kind of imaginative persona may emerge, an ‘I’ who develops a consistent position to take in normative discussion. It is then that we can speak most clearly about what the person accepts, he has a worked out normative position to take in unconstrained contexts” (Gibbard 1990, p. 75). This representation seems quite like Korsgaard’s reflective endorsement, and yet rather different than the universalizing process that aims at shared reasons (Korsgaard 1996b, p. 275).

Imaginative rehearsal is of practical relevance, since working out what to do, what to think, and how to feel in hypothetical contexts will affect what agents do, think, and feel in actual, like situations (Gibbard 1990, p. 72). The capacities involved in reflection and hypothetical reasoning are put to work more substantially in concrete cases of normative discussion, in which agents must find a way toward coordination. However, the normative discussion *with others* differs from imaginative rehearsal because others may engage in reasoning from different starting points (e.g., divergent values, different concerns, and beliefs), which affect the way they account for the context of choice and think of the options open to them. Thus, figuring out what to do with others, in interactional contexts, requires much more extensive operations than the check for internal consistency in considering what to do to them. As in the case of imaginative rehearsal, normative discussion is supposed to produce an agreement over a proposal of action among different people. But it is unlikely that such an agreement will be produced merely by observing the requirement of consistency. First of all, there might be courses of action that differ substantially even though they are not logically inconsistent. The requirement of logical consistency may discourage an individual from pursuing self-defeating ends, but it would not be of any help in selecting which proposal of action should be pursued *by* a group of people. The normative pressure toward reaching a consensus is not solely a logical pressure, and while logical consistency is a necessary constraint on any proposal for action, it is insufficient to provide any shared basis for joint action.

A second reason for rejecting the view that constructivism about reasons should take the form of imaginative rehearsal constrained by the logical requirement of consistency is that it fails the essential insight of constructivism, that is, that normative discussion is generative rather than conservative. To elect consistency as the sole norm of shared rationality is to assume that normative consensus should converge on a proposal that is already present in the first-person deliberative perspective. By contrast, it is plausible to presume that the normative discussion requires not only the check for internal and mutual consistency, but also the revision and comparative adjustment of the original individual proposals. Such revisions may be required to contain or eliminate local inconsistencies, but more interestingly they will also be revisions due by careful and respectful examination of the positions of others, or even demanded *by others*. In other words, normative discussion brings to the fore normative proposals, considerations, urgencies, and needs that may not be imaginable

before entering it. On the Kantian constructivist view, practical reasoning is both transformative and other-oriented.²³

7. THE STANCES OF PRACTICAL DELIBERATION

The debate on moral normativity inaugurated by Korsgaard assumes that these phenomena are captured in the first-person stance, because and insofar as this *is* the deliberative stance in which the agent is constituted. However, it is questionable that the whole problem of moral normativity is fully intelligible in terms of first-person stance as the stance of deliberation.

First, as the case of normative discussion shows, there is an interesting sense in which deliberation is about reaching the first-person plural stance, and reasoning is called for constructing the stance of a group-agent. Correspondingly, it is far from obvious that *all* moral obligations become salient and practically vivid in the first-person stance, that is, in the perspective of the individual deliberator.

Second, another cluster of interesting moral phenomena become salient if one switches from the first-person deliberative stance to the second-person mode that is distinctive of inter-personal relations. The second-person mode is often invoked as an approach to moral obligations alternative to the first-person approach, in that it explicates the authority of moral obligations in terms of irreducible normative relations among persons. To explain what a moral obligation is amounts to explicating how people make claims to one another, and thus reciprocally constrain their deliberation. This is the mode of personal interaction and mutual accountability (Thompson 2004, Darwall 2006). The difference is not just in terms of emphasis: in the first-person stance, the perspective from which moral obligations bind is individual. In the second-person stance, instead, that the bindingness of the obligation depends on its being mutual and mutually recognized.

By distinguishing the issue of the sources of moral normativity from the claim that the question of normative authority arises in the first-person deliberative perspective, we are better positioned to address a concern that threatens to undermine the self-constitutivist view of moral obligations, that is, the place of others. The claims of others seem better accounted in a second-person approach to moral obligations: members of the moral community in general (and not just those whom we owe certain duties) are entitled to demand that we honor our rights-entailing obligations. Members of the moral community have the authority entitling them to claim, if we fail, that we have done wrong even if we have not wronged them personally (Darwall 2007, pp. 60–63).²⁴

²³Bagnoli 2013.

²⁴“Moral obligations simpliciter are what we have the authority to demand of one another, not as individuals, but as representative persons—as representatives, as we might say, of the moral community. So wrongdoing also always involves a failure to respect legitimate demands and the authority to make them that is additional to the individual authorities that are involved in violations of rights and wrongings. Wrongdoing violates our authority as representative persons to demand that people not act in that way. This means that it also involves a failure of respect (of this *representative authority*) and thus a failure to appreciate a further reason not to perform the act that is additional to the act’s wrong-making features, including, indeed, even the fact that it

The reference to the moral community is crucial in that it points to a regulative normative ideal, which does not coincide with anyone's first-person deliberative stance in particular. Yet, for the argument to be convincing, one needs to justify how the bounds of the community are settled. In this respect, the second-person mode of practical deliberation allows us to uncover another cluster of philosophical problems that arise concerning the scope of moral obligations. The second-person mode is distinctive of normative discussion, in which you and I have not yet settled about what to do. In this case, to say that moral obligations are constraints on deliberation requires that you and I have agreed on their content and their authority. But suppose you and I do not agree about the measure of reciprocity that is obligatory. While both parties feel the pull of the other's claim, we also feel that the boundaries must be drawn somewhere, and we do not agree just about where, exactly. On some occasions, the normative discussion takes the form of a negotiation about the content of specific moral obligations. For instance, in the case that you and I deliberate about what to do tonight, as a result of a promise we made to one another last week to see each other next Friday, we may diverge about how exactly to fulfil that promise today that is Friday. Should I come by, or should you? In case you cannot at the time I propose, should we try for another time? We may envision different ways to go by, but it is likely that you and I may find an agreement that settles this particular issue. Once we settled, specific obligations follow, which both of us must honor; if we do not, reactive attitudes such as blame or reproach are called for. In other kinds of occasions, however, people may find hard to compromise. The normative discussion, bounded by reciprocal and legitimate expectations, may uncover differences that bring the normative discussion to a stop. In this case, each of the parties feels that they have reached a line which should not be crossed. You cannot demand that I leave my work behind in order to play tennis with you on Friday. This is where I draw the line. This is not just an occasional and momentary friction; rather, it reveals a ranking of values by which I have ordered my life, which you find unreasonably rigid. You have organized your life so as to play tennis once a week no matter what. I find your policy silly, irresponsible and morally reprehensible. Consequently, I do not recognize your claim as legitimate. Suddenly, the differences between us are manifest, and so severe that you have become alien to me: you really are one of *Them*. The issue about the content of moral obligations has become a more fundamental issue about the boundaries of the normative community, and our respecting standing from within it.

To wit, the second-person stance (I/Thou) is relevant to uncover the normative relations of mutual accountability that hold inside the relevant normative community. But when differences are severe, the third-person stance becomes appropriate to track and voice the depth of the disagreements. The third-person stance is relevant to approach the normative domain as closed under the norm of recognition; it captures the divide between domains of agents on which different norms range. The practice of normative discussion could bind

might violate someone's right (her *individual authority* to demand that one not so treat *her*) and so wrong her", (Darwall 2013, p. 68).

people together by grounding jointly justified and mutually binding reasons, but it might also reveal insuperable divides. Different stances on normative discussions may help us articulate these different forms of disagreement, and thus negotiate not only the content but also the scope of obligations. Perhaps, these divisions will be ultimately overcome, if the negotiations can be guided by sufficiently robust constitutive norms of rationality. But it is quite unlikely that they could be overcome by relying on repeated exercises of imaginative rehearsal and subsequent stretches of reflective endorsement.

REFERENCES

- Bagnoli, Carla. (2021). "The Practical Significance of the Categorical Imperative", *Oxford Studies in Normative Ethics*, 11, forthcoming.
- Bagnoli, Carla. (2017). "Kant in Metaethics: The Paradox of Autonomy, Solved by Publicity", *The Palgrave Kant Handbook/Palgrave Handbooks in German Idealism*, Matthew Altman Ed. London: Palgrave Macmillan, 355-377.
- Bagnoli, Carla. (2013). "Constructivism about Practical Knowledge", in *Constructivism in Ethics*, Bagnoli ed., Cambridge University Press, 2013, 153-182
- Barandalla, A., & Ridge, M. (2011). Function and Self-Constitution: How to make something of yourself without being all that you can be. A commentary on Christine Korsgaard's The Constitution of Agency and Self-Constitution. *Analysis*, 71(2), 364-380
- Bird-Pollan, S. (2011). Some normative implications of Korsgaard's theory of the intersubjectivity of reason. *Metaphilosophy*, 42(4): 376-80.
- Cholbi, Michael J. (1999). Egoism and the Publicity of Reason: A Reply to Korsgaard. *Social Theory and Practice* 25 (3): 491– 517.
- Cohen, G.A. (1996). Reason, humanity, and the moral law. In G. A. Cohen, Raymond Geuss, Thomas Nagel, and Bernard Williams, eds., *The Sources of Normativity*. Cambridge University Press, pp. 167-88.
- Darwall Stephen. 2006. *The Second Person Standpoint*, Cambridge MA: Harvard University Press.
- Darwall Stephen. 2013. *Morality, Authority, and Law: Essays in Second-Personal Ethics*. Oxford University Press.
- Dean, Richard, Sensen, Oliver eds. (2021). *Respect*. Oxford University Press.
- Enoch, David. (2006). Agency, Shmagency: Why normativity won't come from what is constitutive of agency. *Philosophical Review*, 115: 169-98.
- Frankfurt, Harry, (2006). *Taking Ourselves Seriously and Getting it Right*. (Stanford University Press).
- Gibbard, Allan. (1990). *Wise Choices, Apt Feelings*, Cambridge: Harvard University Press.
- Gibbard, Allan. (1999). Morality as consistency in living. *Ethics*, 110: 140-64.
- Ginsborg, Hannah. (1998). "Korsgaard on Choosing Non-Moral Ends," *Ethics*, vol. 109, No. 1, (1998): 5-21.
- Gobsch, Wolfram, (2019). Autonomy and radical evil: a Kantian challenge to constitutivism. *Philosophical Explorations*, 22(2): 194-207.
- Katsafanas, Paul (2018). Constitutivism about Practical Reasons. In Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*. Oxford: Oxford University Press. pp. 367-394.

Katsafanas, Paul (2018). Constitutivism about Practical Reasons. In Daniel Star (ed.), *The Oxford Handbook of Reasons and Normativity*. Oxford: Oxford University Press. pp. 367-394.

Korsgaard, Christine M. (1989). Personal identity and the unity of agency: A Kantian response to Parfit. *Philosophy and Public Affairs*, 18(2): 103-31.

Korsgaard, Christine M. (1996a). *Sources of Normativity*, Cambridge University Press.

Korsgaard, Christine M. (1996b). *Creating the Kingdom of Ends*, Cambridge University Press.

Korsgaard, Christine M. (1997). The Normativity of Instrumental Reason”, in *Ethics and Practical Reason*, G. Cullity and B. Gaut (eds.), Oxford: Clarendon Press. Reprinted in Korsgaard 2008, 27-68.

Korsgaard, Christine M. (1998). Motivation, metaphysics, and the value of self. *Ethics*, 109: 49–66.

Korsgaard, Christine M. (1999). Self-constitution in the ethics of Plato and Kant. *The Journal of Ethics* 3 (1): 1-29. Reprinted in Korsgaard 2008, 100-126.

Korsgaard, Christine M. (2008). *The Constitution of Agency*, Oxford University Press.

Korsgaard, Christine M. (2009). *Self-Constitution: Action, Identity and Integrity*, Oxford University Press.

Korsgaard, Christine M. (2011). Natural goodness, rightness, and the intersubjectivity of reason: Reply to Arroyo, Cummiskey, Moland, and Bird-Pollan. *Metaphilosophy*, 42(4): 381– 94.

Korsgaard, Christine M. (2019). Constitutivism and the virtues. *Philosophical Explorations*, 22(2): 98-116.

Korsgaard, Christine M. (2021). Valuing Our Humanity. In Dean and Sensen 2021, 171-191.

Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin

Moland, Lydia. (2011). Agency and practical identity: A Hegelian response to Korsgaard. *Metaphilosophy*, 42(4): 368-75.

O’Neill, Onora. (1989). *Constructions of Reason*, Cambridge University Press.

O’Neill, Onora. (1996). *Toward Justice and Virtue: A Constructivist Account of Practical Reasoning*, Cambridge University Press.

O’Neill, Onora. (2003). Constructivism in Rawls and Kant. In S. Freeman, ed., *The Cambridge Companion to Rawls*. Cambridge University Press, pp. 347–67.

Samuel J. Kerstein 2020. Korsgaard's Kantian Arguments for the Value of Humanity, *Canadian Journal of Philosophy*, 31(1): 23-52.

Scanlon, Thomas. M. (2007). Structural irrationality. In Geoffrey Brennan, Robert E. Goodin, Frank Jackson, and Michael Smith, eds., *Common Minds: Themes from the Philosophy of Philip Pettit*. Oxford University Press, pp. 84-103.

Scanlon, Thomas. M. (1998). *What We Owe To Each Other*, Cambridge: Belknap Press.

Sem de Maagt (2018). Korsgaard’s Other Argument for Interpersonal Morality: the Argument from the Sufficiency of Agency, *Ethical Theory and Moral Practice*, 21 (4):887-902

Shafer-Landau, Russ. (2003). *Moral Realism*, Oxford: Clarendon Press.

Skorupski John. 2010. “Moral Obligation, Blame, and Self-governance”, *Social Philosophy and Policy* 27 (2): 158-180

Skorupski John. 2010. Rescuing Moral Obligation. *European Journal of Philosophy*, *European Journal of Philosophy*, 6: 335-355.

Stern, Robert. (2011). "The value of humanity: reflections on Korsgaard's transcendental argument," in J. Smith & P. Sullivan (eds.), *Transcendental philosophy and naturalism*, Oxford: Oxford University Press.

Wallace, R. Jay (2012). Constructivism about normativity: some pitfalls. In Jimmy Lenman & Yonatan Shemmer (eds.), *Constructivism in Practical Philosophy*. Oxford University Press.

Wallace, R. Jay (2019). *The Moral Nexus*. Princeton University Press.

William, B. (1981). *Moral Luck*. Cambridge University Press.

William, B. (1996). History, morality, and the test of reflection. In C. M. Korsgaard and O. O'Neill, eds., *The Sources of Normativity*. Cambridge University Press, 210-8.

Zylberman, Ariel (2017). Two Second-Personal Conceptions of the Dignity of Persons, *European Journal of Philosophy* 25:4, 921-943.