

# Human Plight, Kantian *as if*, and Public Reasons in Korsgaard's Moral Theory

**Vitaly Kiryushchenko**  
York University, Toronto

DOI: 10.2478/disp-2020-0004

BIBLID [0873-626X (2020) 56; pp.99–119]

## Abstract

The conception of two kinds of practical identities, which Korsgaard introduces in the *Sources of Normativity*, helps her explain how universal categorical reasoning is compatible with the moral content of individual practical decisions. Based on this conception, she devises an interpretation of the Kantian *as if* principle amended by her argument for the public shareability of reasons. I suggest that, in doing so, Korsgaard steps too far away from Kant's architectonic approach to the question of why moral norms bind us, and that, as a result, the Korsgaardian explanation, as it stands, cannot be accomplished.

## Keywords

Aesthetic judgement; categorical imperative; practical identity; public reasons; reflective distance.

## 1 Introduction

In *The Sources of Normativity* (1996), Christine Korsgaard claims that the possibility of human agency presupposes what she calls a “reflective distance” from our immediate impulses. Reflective distance calls for deliberation, which means that a human agent is not unconditionally determined by her desires to act in a certain way, but is capable of rationally deciding which incentive to act on (Korsgaard 1996<sup>1</sup>: 129, 229–31). Thus, the reflective nature of human consciousness gives rise to normativity:

When you deliberate, it is as if there was something over and above all

<sup>1</sup> Henceforth SN.

your desires, something which is you, and which chooses which desire to act on. This means that the principle by which you determine your actions is one that you regard as expressive of *yourself*. To identify with such a principle or way of choosing is to be, in St. Paul's famous phrase, a law to yourself (SN: 100).

The same theme is restated later in *Self-Constitution: Agency, Identity, and Integrity* (2009) in stronger terms. Korsgaard's principal claim there is that we are *condemned* to choice and action, and that "it is our *plight*: the simple inexorable fact of the human condition" (Korsgaard 2009<sup>2</sup>: 2). In discussing the nature of moral necessitation brought about by the "plight," Korsgaard chooses to take a middle-ground position between sentimentalism and dogmatic rationalism. Normative force, she says, arises neither from mere sympathy towards others, nor from some objective properties out there in the world. She insists that this force is due to an ongoing struggle for integrity and unity of a person through the constitution of her practical identities (SC: 7-26).

In *The Sources*, the conception of practical identity is introduced as "a description under which you ... find your life to be worth living and your actions to be worth undertaking" (SN: 101). Identities come in two flavors. First, there are identities of a particular agent—a philosopher, a parent, a citizen, a homeowner. They provide an agent with reasons that are immediately expressive of her personality. On this view, simply being yourself—as a philosopher, a parent, a citizen, a homeowner—already has some serious normative implications. Second, there is an additional identity expressing our *humanity*. Though we may be born into a particular identity (as children or neighbors) or adopt it (as professionals or believers), we identify with roles and devote ourselves to causes circumstantially (SC: 23). In contrast, our identity as human beings is not contingent, but *necessary*. This identity constitutes a motivational core shared not just by some, but by all moral actors. The distinction between the two kinds of identities is central to Korsgaard's Kantian moral theory in that it forms a point of transition from her account of action to her solution to the problem of the authority of moral norms. At the same time, the distinction introduces a considerable difficulty into this

<sup>2</sup> Henceforth SC.

theory. According to Kant, in order to understand why moral norms bind us, we need to describe how universal categorical reasoning is compatible with the moral content of our ordinary everyday actions and practical decisions, and, more importantly, why each *requires* the other. In Korsgaard's terms, we need to show how to reconcile the identities that are dependent on the contingent circumstances of our lives, with the necessary identity that expresses our humanity.

The Korsgaardian reconciliation has been the subject of much attention since the publication of *The Sources* (Gowans 2002; O'Day 1998; Skorupski 1998; Wallace 2009). Of special importance for the present study is the role played in this reconciliation by Korsgaard's interpretation of three Kantian conceptions. First, it is the "as if," a conception of Kant's aesthetic theory, according to which, in aesthetic judgment, we treat the beauty we perceive in objects *as if* it were a property of those objects and not an idea that has its origins in our reason exclusively (Kant 1914 [1790]<sup>3</sup>: §§7–9). Second, it is Kant's ideas of public reason and the aesthetic *sensus communis* (Gregor 1996: 11–2; KU: 169–73), both of which Korsgaard uses in her argument for the public shareability of reasons. Korsgaard is a Kantian, and it might seem reasonable to assume that the role the *as if* and the argument for the public shareability of reasons play in Korsgaard's theory should be similar to the role aesthetic judgement plays in Kant's transcendental philosophy. Kant sees aesthetic judgement as a *mediator* between the otherwise irreconcilable theoretical and practical parts of his critical project (Pippin 1996; Henrich 1994; Guyer 1990). Prima facie, Korsgaard seems to adopt a similar approach. She uses the publicity of reasons and the Kantian *as if* as two complementary ways to reconcile the contingent and the necessary kinds of practical identities. However, to accept this similarity is to assume too much. As I will show, the purpose of both Korsgaard's *as if* thesis and her account of public reasons is to point at the *immediate* analogy between the normative force of moral reasons and the persuasiveness of the discursive logical forms. The publicity of reasons, on Korsgaard's view, does not so much mediate between the true and the moral, as it simply opens our eyes to the fact that their normativity springs from the same source.

<sup>3</sup> Henceforth KU.

The aim of this paper is twofold. First, it is to show that there is a mismatch between Kant's expressively *architectonic* approach to the problem of the authority of moral norms and the way Korsgaard applies the key ideas of this approach in order to explain the relationship between contingent and necessary practical identities. Second, it is to demonstrate that, against the backdrop of this mismatch, Korsgaard's explanation fails. It is not my intention in this paper either to offer an alternative way to reconcile the two kinds of identities, or to claim that the reconciliation between the two necessarily *requires* a mediating role for aesthetic experience as Kant's *Critique of Judgment* implies. My intention is only to show that Korsgaard's reconstructive use of Kant's aesthetic terminology takes the transcendental arguments central to Kant's first two *Critiques* out of the context of Kant's architectonics. I will begin by describing how Korsgaard's conception of two practical identities is justified within her account of action. I will then show what challenges underlie Korsgaard's distinction between the two—the challenges that ultimately lead her to accepting a version of Kant's notions of the *as if* and public reason. I will proceed by outlining an interpretation of Kant's own solution to the problem of the bindingness of moral norms, and then argue that, on the condition that this interpretation is accepted, the Korsgaardian reconciliation, as it stands, cannot be accomplished.

## 2 Action, practical identities, and the *as if*

Korsgaard's first step away from Kant is evident in her account of action. In the very beginning of *Self-Constitution*, she describes her concept of action as a combination of Kantian and Aristotelian views. On the one hand, just as Kant's syntheses of understanding provide the unity of the manifold of experience in one consciousness (Kant 1929 [1781]<sup>4</sup>: B130–43), actions, as practical syntheses, provide the unified selfhood by putting together what is otherwise a mere heap of impulses, desires, and motives for action. This unifying activity, according to Korsgaard, helps an agent to serve her proper function because it finds its source in the agent's primary practical identity as a human being. On the other hand, Aristotle tells us that the form of

<sup>4</sup> Henceforth KRV.

a thing consists in arranging parts of a thing into a functional unity, in enabling it to serve its purpose (Sachs 1995: 49–59). Kant insists that an action is good if it has the property of lawfulness. According to Aristotle, a good action is performed for the sake of the noble. Korsgaard finds no contradiction in combining the Kantian acting-in-accordance-with-an-imperative with Aristotle’s virtue-based ethics. Her suggestion is that, when we think whether an action is done from duty or for the sake of the noble, we do not think about two alternative purposes for the action, but rather about “a kind of value that may be possessed by *a whole action*—an act done for the sake of an end” (SC: 20; emphasis added).

According to Korsgaard, we often—quite mistakenly—think about the reason for an action as something separated from the action itself. Meanwhile, what we call an action in this case appears to be only an *act* deprived of a context provided by the reason for it. To use Korsgaard’s own example, an act described as “Jack went to Chicago” is worthy if Jack went from Indianapolis to Chicago to visit his mother. If Jack took the journey to buy a box of paperclips, it is unlikely that anyone would see this as a good reason for him to go. This shows that the reason for an action is an important piece of an-action-taken-as-a-whole: we cannot fully understand and evaluate the action without explicating the reasons that endow it with the particular value that it has. A good action is a practical synthesis of different reasons, motives, desires, attitudes, appropriate circumstances, and bodily movements. A good action, taken as a whole and performed at the right time and in the right way, *embodies* its reason (SC: 14).

If this is correct, Korsgaard’s holistic account of action does not allow a clear distinction between hypothetical imperatives, which merely take us from the reasons we have to the choice of what we need to do, and categorical imperatives, which regulate actions performed for their own sake. Affected by an incentive, we formulate a maxim to act upon and then determine whether it contradicts the imperative. If no contradiction is involved, then the moral law expressed by the maxim is not a mere addition to the means-to-an-end structure of an action, but itself becomes a part of the incentive for performing this action. Approached holistically, an action itself becomes an end. Whereas a good action naturally presupposes the means-to-an-end structure, within a particular practical identity (a philosopher, a parent, a citizen,

a homeowner) it has an absolute meaning; and, insofar as we do not make a clear distinction between the two imperatives, we are safe to conclude that all our practical reasons are immediately motivational. Korsgaard makes the normative force of reasons heavily depend on our practical identities. If our maxims are compatible with the categorical imperative, and yet, in performing the actions, we contradict those maxims, then we *lose our identities* and, Korsgaard claims, *cease being ourselves*. Unless we are willingly guided by the two imperatives together, and unless particular maxims of our conduct conform to both of them, we are neither acting nor willing at all.

This line of arguments provides a convincing (and quite elegant) explanation of how we assign meanings to our actions. Our incentives and our will always work together, and we are able to call an action good or bad only if we recognize the means-to-an-end structure of it as a constitutive principle of our autonomous will. The question is, though, whether this rule in fact allows no exceptions. I think it does allow some. If I begin to doubt that I ought to do what I have reasons to do according to one of my identities, this identity, according to Korsgaard, is immediately put under threat of being lost. In this case, my failing to feel the force of my reasons makes the threat *psychologically* very real for me. Does Korsgaard's holism about action guarantee that the threat of losing one's identity is itself a good enough reason to abide by the norms that arise within this identity? I think not; it would not be hard to imagine a case in which the threat means less to me than the reasons I might have to drop the identity. Korsgaard's holistic account of action seems to presuppose some sort of moral phenomenology that conflates psychological dispositions of agents with the norms that establish the authority of those psychological dispositions (cf. O'Hagan 2004: 53). Whatever the merits such phenomenology might have in other theoretical frameworks, its acceptance does not seem to be a good outcome for Korsgaard's account.

In Korsgaard's defence, it might be claimed that, if we admit that all agent's practical identities are tightly interconnected, and together uniquely constitute what the agent considers to be *herself*, a holistic account of action seems to be our only option. Indeed, unless we treat *any* one of our actions as an end in itself, pursuing which amounts to having our very humanity at stake, we cannot even begin to consider our actions in the context of reasons that motivate them. It is an "all

or nothing” kind of situation. This moral rigor, however, comes at a price. If there is absolutely no trade-off allowed between our short-term, routine practical goals and our absolute moral ends, it is not clear how much of this or that practical identity we can give up under serious threat in order to avoid complete self-destruction. As has been claimed above, Korsgaard shows successfully how her idea of two practical identities fits within her holistic account of action. However, her distinction between two kinds of identities (contingent and universal) is challenged in view of the way she describes the relationship between two Kant’s imperatives (hypothetical and categorical).

My conjecture is that it is in anticipation of this and similar objections that Korsgaard introduces the notion of the *as if*, which provides the first important link between her holistic account of action and her moral theory. Here is one way of seeing how this introduction is motivated. According to Kant, freedom is thinkable for an individual agent in the form of a universal restrictive law exclusively (Engstrom 2009; Kitcher 2004; Kong 2012; Paton 1947; Pelegrinis 1984). Only the categorical imperative can be such a law—and a moral law proper, for that matter (Kant 2002 [1785]<sup>5</sup>, 4:420). Contrary to this view, Korsgaard insists on the inseparability of the categorical and the hypothetical imperatives. Our necessary identity, which refers to our humanity, incorporates the means-to-an-end structure that governs our actions as expressions of our contingent identities. The contingent, defined by one’s self-determination relative to some end, and the necessary, which describes one’s self as a cause of this end, are inseparable. But the nature of this self-determination is yet to be clarified. We may certainly *agree* on something (e.g. our humanity) being the condition of all value, but how would this help? For instance, how would our agreement eliminate the difference between the normativity of Korsgaard’s claim that anyone would have to do the same as me if analogously situated, and the normativity of a much stronger *Kantian* claim that a chosen action is obligatory for any moral actor, regardless of their particular situation and practical identities? This poses a dilemma for Korsgaard. On the one hand, “[t]he intersubjectivity of reasons cannot be constructed out of individual acts of reflection and transcendental glue” (O’Hagan 2004: 58). We cannot simply supplement a summary

<sup>5</sup> Henceforth G.

of individual practical decisions with an utterly abstract, but unifying transcendental story. On the other hand, the categorical imperative as a formal rule abstracted from all sorts of subject matter of the will cannot serve as a proper object of some contingent belief. Accepting the dilemma as it stands leads either to Kant's own "aesthetic" solution (to be discussed below), or to a moral phenomenology, in accepting which I face an unsettling choice between living under permanent psychological threat and becoming a perfect moral machine. It is to tackle this dilemma that Korsgaard introduces the notion of the *as if*.

In endorsing moral maxims, Korsgaard says, an agent acts *as if* she is what she identifies herself with. The *as if* Korsgaard has in mind here is, of course, not about concocting fictional stories for the lack of a better choice. It is a discursive structure that makes the reconciliation between the two kinds of identities possible by allowing an agent to treat her contingent identities as if they were "the sources of absolute inviolable laws" (SC: 23). Thus understood, the *as if* may be considered a clarification of Korsgaard's reflectivity thesis, with the formulation of which our discussion started. Recall that, according to Korsgaard, due to the reflective nature of human consciousness, we are distanced from our immediate desires and, therefore, in our actions we are not determined by those desires—at least not determined unconditionally. Because of this, simply conforming to a rule is not enough for us in most circumstances; we must also take some reflective stance towards it. What matters is how we *take* the rule. On this view, the fact that something,  $x$ , is normative, supervenes on the fact that we have reasons to consider  $x$  as normative. Accordingly, a practice is normative just in case we have reasons to get involved in it *as if* it was such. Korsgaard needs this idea because it helps her to avoid a confusion between our psychological dispositions towards existential threats and moral norms that establish the authority of those dispositions. In treating my dispositions *as if* they were objectively valid, I immediately move, as it were, from the bottom to the top of Korsgaard's account, or ascend, in one giant leap, from reasons based on my contingent practical identities to reasons based on objectively valid norms. *Prima facie*, then, with the introduction of the *as if*, the objection from moral phenomenology fails.



### 3 Aesthetic judgment and autonomy

I claimed in the beginning that there is a mismatch between Kant's architectonic approach to the problem of the authority of moral norms and the way Korsgaard applies the key ideas of this approach in order to explain the relationship between contingent and necessary practical identities. Before summarizing the mismatch, a comparison between the two approaches is needed. This will require a bit of exposition.

Kant's rational agent always *knows* how she ought to act, but the moral law she uses to assess her own practical maxims is abstracted from all sorts of concrete practical consideration in favor of its universal form. Because of this, whenever, in facing a particular practical choice, the agent questions this or that particular resolution to act, she can never answer the question *why* she has to follow the imperative based on her experience. Being able consistently to treat others as ends and being committed to duty for its own sake requires a conversion in which moral righteousness becomes the *nonnegotiable* condition of one's pursuits. This conversion cannot be accomplished by gradually improving one's understanding of the highest moral end. Meanwhile, the agent also knows that freedom through moral action is achievable in principle. It is a moral *ideal*, an a priori concept of which is imposed on us by reason. That is why the categorical imperative, which is an expression of the moral law, has only a *regulative* significance, and no other rule, motive, or act of understanding can effectively mediate between it and the vicissitudes of my particular practical decisions. According to Kant, to explain where the normative force of *moral* maxims comes from, we need to appeal to *aesthetic* experience, which can help us eliminate the gap between the acratia whims of our nature and our unconditional freedom.

The key issue that leads Kant to appeal to aesthetic judgment is that both in *Groundwork* and in the *Critique of Practical Reason* the hypothetical and the categorical imperatives work separately. Kant is clear on the fact that knowledge of a formal law as such does not constitute any action-guiding competence, and therefore, does not prevent a moral agent from heteronomous behavior in accordance with the hypothetical imperative that may well contradict such knowledge. Yet the agent may resort to an aesthetic judgement—a kind of judgement which,

although not itself based on knowledge of any explicitly formulated rule, helps the agent work her theoretical knowledge into a finite set of practical habits. In §59 of the *Critique of Judgment*, (“Of Beauty as the Symbol of Morality”) Kant claims that aesthetic judgments, just like moral judgments, are characterized by universality. But unlike the principles of moral judgment, the principles of judging about the beautiful, in Kant’s own words, are “not cognizable through any universal concept” (KU: 252). Aesthetic experience appeals instead to what Kant calls the “communal sense”—a sense that we are “compelled to postulate [...] to account for the agreement of men in their appreciation of beautiful objects” (KU: xxii). Simply put, an agent might not be able to rationalize why this particular something appears beautiful to her. However, the agent is unable to live in an aesthetically deficient world, and in the changes she makes she conforms to the law that she cannot formalize. As Ginsborg (2001) puts it, Kant’s aesthetic judgement

can function independently of the understanding in situations where the relevant rules or concepts are not already specified. To exercise judgement in this independent way is to judge particulars to be contained under rules or concepts which are, so to speak, not already in understanding, but rather made possible by those acts of judging, themselves (253).

Accordingly, Kant’s second and third *Critiques* together form two series of interdependent arguments: As a moral agent aware of the categorical imperative, I cannot work it into a finite set of practical habits; and conversely, as an aesthetic agent *not* aware of the law of aesthetic judgement, I always have an idea of how to rearrange my setting. It is precisely because the normative force of the principles of practical reason may not always work in the same way as the normative force of logical forms, in the *Critique of Judgment* Kant offers a re-interpretation of logically structured knowledge into ethically sound conduct by the power of aesthetic judgement. Although grounded on our personal emotions, aesthetic delight, just like moral reasoning, claims universal validity and allows us to make morally sound conduct thinkable. On the one hand, contrary to Korsgaard, no one can consistently adjust their behavior to a formal rule in making *every* choice—simply because each of us makes too many of them to be able

to tell which ones involve maxims important enough to be subject to the universalizability test. However, we are able to *think as if* the application of our maxims were logically consistent. On the other hand, though the categorical imperative, a formal rule as it is, cannot acquire any particular practical meaning, we can still *act as if* our behavior was morally meaningful by appealing to the aesthetic idea of a law that is to obey without asking any why. Thus, Kant's approach to the problem of normativity, as he himself insisted on many occasions, is *architectonic*; in order to solve the problem of the normative authority of moral reasons, it involves all three normative sciences: logic, ethics, and aesthetics.

Korsgaard does not accept Kant's interpretation of morality in terms of beauty. She makes no use of aesthetic judgement as a mediator between the theoretical and the practical. Contrary to Kant, she wants to make the notions of the practical and the normative *immediately* available. The authority of the principles of practical reason, she insists, should have the same self-evident source as the normative force of discursive logical forms (SC: 67). As one of Korsgaard's examples suggests, it is not a person's *belief* in modus ponens that leads her from "if A then B" and "A" to the conclusion that "B," but the modus ponens *itself*. Likewise, in moral reasoning, the rules are there not just to be learned by rote, but to define the essential features of the mental economy of a person attempting self-determination. In a few passages throughout *Self-Constitution*, Korsgaard claims that by the very fact of accepting a rule, I cause myself to act (SC: 22, 106). But what does *accepting* a rule mean? It indeed seems to imply something more than "knowing," "being aware of," or "understanding." At the same time, "accepting" might also mean different things. I may unwillingly accept something that I consider inevitable, or willingly accept something that I know is wrong. My own maxims *may* comply with the formal requirements of the moral law, but what seems to be lacking, in this case, is a conceptual link between the two kinds of identity that would explain how exactly my accepting a rule causes me to act.

True, I never accept anything in a vacuum, and this seems to be an important point of departure for Korsgaard's "as if" thesis in the context of her idea of self-constitution. There is a set of moral and psychological relations, which have to be in place and which are constitutive for me being *myself*. As Rovane (1998) puts it, those are

the relations that “must include some substantive practical commitments that serve as unifying projects,” whose execution “depends upon achieving ... overall rational unity within the self” (31). The condition of personal identity is thus the same condition that gives rise to the normative commitment to achieve the unity. Given this, whenever I act on a sincere belief grounded in one of my identities, my whole self is put in motion. But let us pause here for a moment. As Korsgaard’s self-constitution is “the struggle to be, in the face of psychic complexity, a single unified agent” (SC: 7), it is a *process*. Given the way this process is described, my will, in being involved in it, is determined by a set of purposes that have a hierarchical order: some of them I pursue for the sake of other purposes, some of those for others still, etc. At the very top, there is the highest purpose which is not conditioned by anything else and which commands that all my acts are to aim only at my own virtue as well as happiness for every rational agent. But again, the categorical imperative, which conditions this highest purpose and which I use as a standard in the process of my self-constitution, is an a priori rational principle. It tells me always to treat others as ends, not as means, and is applied simply because I am in possession of rational will, without reference to any particular lower-level purpose that I might have. The principle itself, as described in the *Critique of Practical Reason*, represents my *rational hope*. The whole point of calling the principle “regulative” is that there are no *schemata* that can effectively mediate between it and my actions in the way such schemata mediate between the a priori categories and experience in the *Critique of Pure Reason*. In the first *Critique*, Kantian categories of the understanding, in order to be applied to the manifold of experience, require the unifying activity of the imagination. The imagination operates through schemata that share the generality of the categories and the temporality of experience. By the same logic, in the second *Critique*, a mere heap of impulses, desires, and motives for action, to be presented as a unified selfhood, need something that combines the universality of the moral law and the continuity of my motives and desires, the continuity that I identify as *myself*. From the Kantian point of view, this something is my aesthetic experience, which helps me convert my knowledge of the law into my capacity to act on this knowledge. From the Korsgaardian perspective, it is a mere leap that I have to make from

reasons originated within my contingent identities to reasons based on the objectivity of the moral law. According to Korsgaard, the unifying activity is supposed to be performed by actions themselves. The whole point of her holistic account of action, as discussed above, is to justify this thesis. But again, what is lacking is an explanation of *why I have to act in the first place*.

Korsgaard's another attempt to show that the normative force of moral reasoning and the shareability of the a priori forms of theoretical understanding are immediately connected is to demonstrate that they rest on the same foundation of human autonomy. Recall that in the previous section we considered the Korsgaardian "as if" thesis as extension of her argument about the reflectivity of human consciousness. The idea of the reflective distance—the capacity to get one's own practical attitudes into reflective view—requires us to take an autonomous practical stance towards an incentive we have. According to Korsgaard, our capacity to *reflect* immediately brings the plight of moral choice of how to *act*. It is reflectivity, not mutual human sympathy or objective moral facts, that underlies our capacity to make truly moral choices. Only when we act from the distance, *as if* our practical identities had an absolute value, as the idea of autonomy suggests, the practical decisions based on our own will give rise to the obligations of normativity. This constitutes an important similarity between Korsgaard and Kant.

With this similarity in view, while both Korsgaard and Kant focus on autonomy, it is important that Kant's autonomy thesis may be interpreted in two different ways. The first, *epistemic* way consists in the overarching claim of the *Critique of Pure Reason* that the cognizing subject is autonomous as far as it prescribes to nature the rules and forms according to which the subject can have knowledge of it. In this sense, the autonomy of the cognizing subject follows from the fact that sensibility and understanding contain a set of constitutive principles a priori for the faculty of cognition. What Kant argues throughout the *Critique of Pure Reason* is not just that objects must conform to the conditions of our knowledge, but that, if we are to have any success in coming to know those objects, our reason has to *impose* such conformity upon them. Kant insists that,

however exaggerated and contradictory it may sound, to say that the

understanding is itself the source of the laws of nature, and so of its formal unity, such an assertion is none the less correct, and is in keeping with the object to which it refers, namely, experience” (KRV: A 127).

The second, *practical* way to posit our autonomy is to define it as “the property of the will by which it is a law to itself (independently of any property of the objects of volition)” (G 4:440). All maxims of conduct that pass the test of universalizability are laws that have grounds in Reason itself—and so are self-imposed. As Guyer (2014) puts it, according to Kant, “the only way for your will to be free or autonomous is for it to be governed by a law that it gives itself rather than to allow itself to act on whatever mere inclination happens to be alluring at the moment” (234).

In view of this double interpretation, we are facing the same problem all over again. In Kant’s case, the two versions of autonomy thesis stem from two different parts of Kant’s theoretical edifice, which have to be reconciled by means of an appeal to aesthetic experience. On Korsgaard’s view, when we act in accordance with our contingent practical identities, we ascend to the level of the universal by treating those identities as if they had a universal value. Agents must make choices as if they are free even though they struggle to believe themselves to be such, because, in order to do anything at all, an agent must ignore the cultural and psychological reality that programs her, and decide what to do just as if she were free. Kant does use the notion of the “as if” in a similar manner in the *Groundwork*, in one of the formulations of the categorical imperative: “Act as if the maxim of your action were to become by your will a universal law of nature” (G 4:421). In Kant’s case, however, the notion receives its full justification only in the third *Critique*, as discussed above. Without this justification, within the framework of Kant’s ethics exclusively, the use of it remains underexplained.

#### 4 The publicity of reasons and *sensus communis*

Another attempt to reconcile the two kinds of identities can be found in Chapter 9 of *Self-Constitution*. This attempt may be seen as complementary to the bottom-up “as if” approach outlined earlier in the book. In this, new version, Korsgaard shows how we can move in

the opposite direction, from top to bottom, or from the universal reasoning to the moral content of our particular practical choices. We can do it, she says, through the appeal to the *public* character of our reasons. One of the examples she uses is that of promising (SC: 89-90). Suppose *A* offers a promise to *B*. If we consider this fact as comprising two successive acts—*A* giving a promise and *B* accepting it—we will not be able to understand a promise as an action taken as a whole in the sense outlined in section 2 above. Until *B* accepts the promise, *A* at any moment can take it back. There is no commitment prior to acceptance in this case due to the temporal gap between the two acts that prevents the promise from counting as an action proper: *A*'s act has no purpose and, therefore, *A* does not act at all. This is why actions like promises “must result from the formation of the single common will, from a moment of unity between us” (SC: 190).

In introducing the idea of the publicity of our reasons, Korsgaard makes three important points. First, in making a promise, I do not have an exclusive right over my own decision to act in a certain manner. Second, I cannot simply take it for granted that my having a good reason to *q* means that you would also have a good reason to *q* were you in circumstances similar to mine. If I have reasons to suppose that *q*-ing will make me happy, these are my reasons and my happiness, not yours. Meanwhile, according to Korsgaard's public conception of reasons, I cannot simply grant, but should *will* the outcome of your action—which is not possible in case our private reasons compete. If I think I have a reason to shoot you, I can accept the fact that you have a reason to do the same to me, but I cannot possibly *will* that you shoot me. Korsgaard's third point is that a true action results only from our reciprocal acts of will, which yields a *unified* action of me-agreeing-to-offer-what-you-agree-to-accept. These three points ensure that “it is only on the public conception of reasons that an universalizability requirement is going to get us into moral territory” (SC: 192). In performing actions based on public reasons, we enter into a special kind of agreement, where we are bound by mutual moral obligations. And when we make such agreements, “we *meet* in the noumenal world” (SC: 197). Meeting in the noumenal world means that we neither treat one another's reasons instrumentally, nor act out of sheer sympathy for one another. We treat one another as sources of reasons that have legislative power over us.

In appealing to the public character of reasons, Korsgaard aims to present the normative force that binds you and me, in the form of an immediate communion between us, which emerges when we think and act having one another in view as ends in themselves. It might seem that the idea of such communion serves in Korsgaard's moral theory the same role the aesthetic idea of *sensus communis*—the idea of an unconditional agreement in our shared aesthetic-based appreciation of beautiful objects—plays in Kant's architectonics. However, there is a difference between the two. According to Kant, this agreement expresses a necessary connection between beauty and goodness, and is an aesthetic prerequisite for the feeling of the sublime character of our moral goals, as well for as our capacity to form moral judgements. Feeling interest in the beauty of nature, according to Kant, is a definite mark of a moral disposition. Kant views the communion to be a result of an aesthetic idea justifying the principles of practical reason by mediating between the normative force of those principles and the normative force of discursive forms. In sharing some characteristics of both, it helps us convert logically structured knowledge into ethically sound behavior (KU: §40).

For Korsgaard, the possibility of such communion has an entirely different source. It is based on a somewhat peculiar idea about human communication; namely, on our being capable of a non-manipulative use of speech, or a "straight talk" in Korsgaard's jargon (SC: 196). In her polemics with Henry Sidgwick (SC: 177-206), Korsgaard acknowledges that an agent's motives expressing her self-interest are never fully and utterly transparent to her:

You can't determine your thoughts by determining what would be best for you to think, *not all the way down*. Thinking, after all, is just talking to yourself. And since meanings are public, talking is just thinking in the company of others (SC: 197).

In thinking privately, I am confused more often than not, and the transparency of my own private reasons is achievable only in a specifically organized communication with others. Reasons, says Korsgaard, are grounded not only in my personal desires and projects, but also (and more importantly) in my public reasoning that grounds my social positions.

On the interpretation I would like to offer here, this argument



from *Self-Constitution* runs parallel to Korsgaard's earlier claim in *The Sources*, where she urges that, in beginning with the motives that grow from our contingent identities, we have to build a step-by-step conceptual ladder all the way up to our highest moral ideal. In doing so, however, we have no way of showing conclusively why an agent would be inclined to take reasons of others into account. The only way up the ladder is to act *as if* these reasons were valuable. In beginning with publicity, on the contrary, we immediately grasp the key characteristics of reasons themselves, which can be traced all the way down to our particular motives and desires. The publicity of reasons is reflected in their shareability, in the fact that they are *social* (SN: 130-135; O'Day 1998: 63–5). At the same time, the publicity of reasons, just as Korsgaard's earlier "as if" argument, marks our autonomy. It shows that whenever we get involved in reasoning practices, there are certain conditions of cognition and action resulting from those practices, which are authored by us and should always already be in place in order for the practices to be possible. Given this link to Kant's autonomy thesis, it is quite surprising that, to make her final point, Korsgaard abandons Kant and instead uses Wittgenstein's private language argument. According to this argument, one cannot have a language that is in principle unintelligible to others. Wittgenstein shows that, for the norms of language to bind us, these norms should be publicly shared, and that a language entirely unintelligible to others would be unintelligible even for the person who supposedly has it. In Korsgaard's words, "[y]ou can no more take the reasons of another to be mere pressure than you can take the language of another to be mere noise" (SN: 143). As summarized in Norman (2000), "just as we cannot help responding to other people's utterances as meaningful, so we cannot help taking other people's reasons as reasons" (Norman 2000: 294).

The important fact to pay attention to is that language, which externalizes the reasons we share, is capable of making sense of disagreements just as much as of agreements between different perspectives. It is not, therefore, immediately clear whether the public character of reasons helps us all that much to come to terms with each other on every occasion. To put the matter differently, there is quite a distance to cover between perceiving one's words as mere noise and being a fully committed member of the Kingdom of Ends. The

choice between the two extremes, on which Korsgaard insists, might seem quite a bit too rigid (Geuss 2009: 198). Our agreements often-times already work perfectly well on the level of local social practices. For instance, you and I can share the view that my brother has committed a felony. However, beliefs that constitute our respective networks of collateral commitments, and the way we act on those beliefs on occasion, might be incompatible *in spite* of this agreement. From my perspective, the fact that I believe that my brother committed a felony might not necessarily presuppose the legal consequences it would from your perspective if you were a judge. And yet, it is quite possible that you and I keep using the term without a slightest trace of disagreement on a number of occasions when my particular attitude toward the relative is not explicitly involved. In fact, as our communication implies a potentially unlimited number of contexts in which the term “felony” can be applied, it is theoretically possible that we spend a lifetime discussing/acting on our beliefs about “felony” without ever contradicting each other, given that our claims/actions do not involve the difference in our commitments. But even more importantly, the normative force that will ultimately make me change my mind about my brother may result not from my assessing my own beliefs in universal terms, as Korsgaard would suggest, but from my rethinking the meaning of an appropriate article of the criminal code. Instead of appealing to self-constitution and the descriptions under which I value my humanity, it might be enough for the case simply to change my perspective on a particular set of local institutional practices. The reasons that led to an agreement between us in this case would be just as good and just as public as they would be if I were to universalize the resulting maxims. More to the point, even if we simply took the direct analogy between the normativity of meaning and the normativity of morality for a fact (as Korsgaard wants us to), then, given that language always animates a particular set of cultural norms, we would have to admit that the normativity of morality is a contingent outcome of conceptual schemes we share within a particular linguistic reality. Does this relativistic picture represent the result at which Korsgaard’s publicity thesis aims? Most likely not.

On a side note, even a complete egoist can be quite capable of sharing reasons. I may understand you perfectly well and yet ignore

what you say (Cholby 1999: 508–11). To reject egoism, we must demonstrate that practical reasons are shareable in a *substantive* sense, i.e. to demonstrate that we can obligate one another. What Wittgenstein’s private language argument aims to show is the impossibility of meanings that cannot be in principle shared by others. What Korsgaard aims to show *in using* Wittgenstein’s private language argument, is how the normative force of reasons is intersubjective (Wallace 2009: 483). And this is a larger problem. Consequently, an unresolved issue here is the relationship between the *semantically* grounded availability of meanings, which Wittgenstein’s argument demonstrates, and the *pragmatically* grounded normativity of moral judgement, for the justification of which, according to Korsgaard, this argument allegedly can be used.

## 5 Conclusion

Kant’s architectonic approach suggests that aesthetic normativity serves as a keystone that holds together his entire critical project. Moral maxims do not prevent an agent from behavior that contradicts her beliefs, and epistemic normativity by itself does not constitute any practical, action-guiding competence. Aesthetic judgement fixes the problem by mediating between the requirements of morality and the requirements of theoretical knowledge. What it shares with moral judgement is its disinterestedness, its universality, and its freedom. Just as the moral law is abstracted from all sorts of concrete practical consideration in favor of its universal form, the judgement of taste is disinterested because the aesthetic delight it expresses is not bound up specifically with any of the particulars of its object. The judgement of taste is free in the sense that it is a result of a creative play of imagination; and it is universal in that, not relying on any concept, it represents the beautiful as the object of a universal satisfaction. According to Kant, it is for no other reason than this that we “speak of the beautiful, as if beauty were a characteristic of the object and the judgement logical . . . although it is only aesthetical and involves merely a reference of the representation of the object to the subject” (KU: 56). Kant poses a gap between theoretical understanding and practical motivation—the gap which is inherent in moral consciousness and which stands in need of a mediation between the two.

Korsgaard recognizes the gap and uses Kant's key terms, but abandons Kant's aesthetic-based approach. Instead, she considers her *as if* thesis and her publicity argument as two complementary ways to argue for the *immediate* connection between truth and morality. Just as moral decisions are practical syntheses that put together what is otherwise a manifold of impulses, desires, and motives for action, our acts of understanding are theoretical syntheses that put together what is otherwise a manifold of mere empirical intuition. Both of Korsgaard's arguments aim to show that the normative force of moral reasoning and the a priori forms of theoretical understanding rest on the same foundation of human autonomy. Following Kant's logic, without the resort to aesthetic reflective judgment, we have no proper way to show that the true and the moral are mutually consistent. Meanwhile, in Korsgaard's case, Kant's third transcendental argument appears to be broken into two parts, none of which achieves the goal.

Vitaly Kiryushchenko  
 Department of Philosophy  
 York University, Toronto  
 kvitaly@yorku.ca

### References

- Cholby, Michael. 1999. Egoism and the publicity of reason: a reply to Korsgaard. *Social Theory and Practice* 25(3): 491–517.
- Engstrom, Stephen. 2009. *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Cambridge: Harvard University Press.
- Geuss, Raymond. 1996. Morality and identity. In *The Sources of Normativity*, by C. Korsgaard et al. Cambridge: Harvard University Press.
- Ginsborg, Hannah. 2011. Primitive normativity and skepticism about rules. *The Journal of Philosophy* 108(5): 227–54.
- Gowans, Christopher. 2002. Practical identities and autonomy: Korsgaard's reformation of Kant's moral philosophy. *Philosophy and Phenomenological Research* 64(3): 546–70.
- Gregor, Mary (ed.) 1996. *Practical Philosophy: Immanuel Kant*. Cambridge: Cambridge University Press.
- Guyer, Paul. 2014. *Kant*. New York: Routledge.
- Guyer, Paul. 1990. Feeling and freedom: Kant on aesthetics and morality. *The Journal of Aesthetics and Art Criticism* 48(2): 137–46.
- Henrich, Dieter. 1994. *Aesthetic Judgment and the Moral Image of the World: Studies in Kant*. Stanford: Stanford University Press.

- Kant, Immanuel. 1785. *Groundwork for the Metaphysics of Morals*. Translated by Allen Wood. New Haven: Yale University Press, 2002.
- Kant, Immanuel. 1781. *Critique of Pure Reason*. Translated by Norman Kemp-Smith. London: Macmillan, 1929.
- Kant, Immanuel. 1790. *Critique of Judgement*. Translated by John Bernard. London: MacMillan, 1914.
- Kitcher, Patricia. 2004. Kant's argument for the categorical imperative. *Noûs* 38(4): 555–84.
- Kong, Camillia. 2012. The normative source of Kantian hypothetical imperatives. *International Journal of Philosophical Studies* 20(5): 661–90.
- Korsgaard, Christine. 2009. *Self-Constitution: Agency, Identity, and Integrity*. Cambridge: Harvard University Press.
- Korsgaard, Christine. 1999. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Korsgaard, Christine. 1996. *The Sources of Normativity*. Cambridge: Harvard University Press.
- Norman, Richard. 2000. Public reasons and the “private language” argument. *Philosophical Investigations* 23(4): 292–314.
- O’Day, Ken. 1998. Normativity and interpersonal reasons. *Ethical Theory and Moral Practice* 1: 61–87.
- O’Hagan, Emer. 2004. Practical identity and the constitution of agency. *The Journal of Value Inquiry* 38: 49–59.
- Paton, Herbert James. 1947. *The Categorical Imperative: A Study in Kant’s Moral Philosophy*. Philadelphia: University of Pennsylvania Press.
- Pelegriinis, Theodosius. 1984. *Kant’s Conception of the Categorical Imperative and the Will*. London: Zeno.
- Pippin, Robert. 1996. The significance of taste: Kant, aesthetic and reflective judgment. *The Journal of the History of Philosophy* 34(4): 549–69.
- Rovane, Carol. 1998. *The Bonds of Agency. An Essay in Revisionary Metaphysics*. Princeton: Princeton University Press.
- Sachs, Joe (ed.) 1995. *Aristotle’s Physics. A Guided Study*. New Brunswick: Rutgers University Press.
- Skorupski, John. 1998. Rescuing moral obligation. *European Journal of Philosophy* 6: 335–55.
- Wallace, Jay. 2009. The publicity of reasons. *Philosophical Perspectives* 23: 471–97.