



Responsibility and the Unity of Self Variations on a Kantian Theme in Brandom and Korsgaard

Adam J. Graves

Metropolitan State University of Denver
agrave13@msudenver.edu

Abstract: The philosophical projects of Robert Brandom and Christine Korsgaard are not often associated with one another, even though both authors share a common interest in normativity and trace that interest back to a common source, namely, the work of Immanuel Kant. More remarkably still, both authors also appeal to the task-responsibility of self-integration (that is, the need to weed out incompatible commitments and/or desires) in their respective accounts of normativity. In this essay, I argue that this task-responsibility cannot on its own sufficiently account for normativity in the strong sense – in the sense that judgments are thought to be answerable to objects and actions are thought to be constrained by moral obligations. Nevertheless, I claim that combining elements of Brandom's pragmatist program with elements of Korsgaard's theory of practical identity allows one to offset gaps in their respective accounts and to develop a more satisfactory account of the nature and source of normativity.

Keywords: Kant, Immanuel; Brandom, Robert; Korsgaard, Christine; normativity; responsibility; selfhood; self-integration; pragmatism; identity; agency.

Kant and the Normative Turn

In the opening section of *Making it Explicit*, Robert Brandom claims that one of Kant's cardinal innovations was to show that "conceptually structured activity is distinguished by its *normative* character."¹ Whether these activities involve theoretical judgments or practical doings, they are best understood "in terms of the special ways [in which we make ourselves] responsible for them" (*ME* 9). In order to clarify precisely what made this "normative turn" so special, Brandom contrasts Kant's central concern to the central concern of Cartesian epistemology. He writes:

For Descartes, the question was how to think about our grip on our concepts, thoughts, or ideas (Is it clear? Is it distinct?). For Kant the question is rather how to understand their grip on us: the conditions of the intelligibility of our being bound by conceptual norms.²

Within the Cartesian framework, the most urgent task is to explain what representational success amounts to once the relation between an object and its representation can no longer rest upon a naive notion of resemblance. In other words, his challenge was to explain what it means for a judgement to be certain. Within the Kantian framework, at least on Brandom's reading of it, the most pressing task is to understand

¹ Robert B. Brandom, *Making It Explicit: Reasoning, Representing, and Discursive Commitment*, Cambridge, MA: Harvard University Press 1994, p. 8. [Henceforth cited as *ME*]

² Robert B. Brandom, *Reason in Philosophy: Animating Ideas*, Cambridge, MA: The Belknap Press of Harvard University Press 2009, p. 33. [Henceforth cited as *RP*]

the nature of normativity itself—"the bindingness or validity (*Verbindlichkeit, Gültigkeit*) of conceptual norms" (RP 33), the special ways we are made responsible (and thus liable to normative assessment) as a result of the rules (for example, of appropriate application) that are implicit in the concepts humans employ in their theoretical commitments and practical undertakings. In other words, Kant's challenge is to explain what it means for a judgement or action to be necessary.³

Brandom is not alone in characterizing Kant's legacy in terms of a normative turn, which places questions of necessity, obligation, and responsibility on center stage. Christine Korsgaard's work, though more narrowly focused upon moral psychology, represents a variation on the same theme. On her account, Kant's conception of autonomous agency provides the makings of a solution to an heretofore insolvable puzzle regarding the nature, source and justification of the authority implicit in one's experience of moral obligation, one's sense that something ought to be done or believed. Korsgaard is rightly dissatisfied with many of the classical ways of explaining normativity's source. On the one hand, the so-called voluntarists who attempt to trace obligation back to the commands of an authoritative legislature are no less doomed than, say, Plato's Euthyphro; while, on the other hand, the substantive realists who seek intrinsically normative entities are either faced with an infinite regress (for what makes those entities normative?) or must engage in a bit of alchemy, hoping in vain to transform "what is the case" into "what ought to be."⁴ The distinctive advantage of the Kantian approach, as understood by Korsgaard, is that it shows how normativity arises directly out of the very nature of practical selfhood itself—it shows that a unified self, one that acts as an autonomous agent rather than a mere heap of garbled impulses or desires, is one that makes itself genuinely responsible to principles which it imposes on itself, rather than to principles imposed upon it by a foreign legislative authority or by independently existing entities.

³ Brandom points out that necessity (*Notwendigkeit*) should be understood in terms of *deontic* rather than *alethic* modality, since what Kant has in mind is the manner in which judgement and action are bound by laws (ME 10).

⁴ These are admittedly rather hasty dismissals, at least as I have articulated them here. For a more nuanced treatments of these classical accounts, see Christine M. Korsgaard, *The Sources of Normativity*, Cambridge University Press 1996, pp.21-47. [Henceforth cited as SN]

The fact that two prominent American philosophers cite Kant as the chief inspiration for their respective theories of normativity is perhaps not all that surprising. What is surprising is just how little attention has been paid to the seemingly obvious overlap between their respective projects. Despite the fact that they enter Kant's kingdom through different gates and for wholly different purposes, the similarities between Korsgaard and Brandom remain rather striking. I will explore here just one of these similarities—namely the connection between responsibility and selfhood, or more specifically, the indispensable role responsibility is said to play in the formation of a unified self.⁵ My interest in staging this imaginary dialogue between Brandom and Korsgaard goes beyond the merely academic enterprise of noting a resemblance. I argue that their respective stories can be told in a manner that reveals their complimentary nature. Korsgaard's understanding of the relation between practical identities and the formal requirements of autonomous selfhood supplies an explanation of the strong sense of obligatoriness in the practical domain which seems to be missing in Brandom; while Brandom's pragmatist framework provides, in turn, a metalanguage for understanding the social conditions that underwrite this kind of obligation.

Allow me to begin by clarifying what I mean by normativity. Wherever norms are concerned one finds three interdependent components: (a) a kind of necessity associated with binding obligations, that is, the sense of ought that genuinely obliges but does not compel; this kind of obligation implies (b) a standard of assessment, such as rules regarding the appropriate use of a concept or the correct formulation of a maxim; and this in turn suggests (c) the possibility of either success or failure. In other words, norms not only entail a kind of necessity, but they also imply standards and presuppose fallibility (that is to say, the possibility of failing to meet those standards). These are the three legs of normativity. In the absence of components (b) and (c), normative force would become indistinguishable from natural or causal forms of necessitation, and the idea of normativity would ultimately collapse upon itself.

⁵ Curiously, Kant himself never made this connection explicitly—and so drawing the connection requires one to attend to the spirit rather than the letter of Kant's corpus. Arguably this spirit animates both of their projects.

Responsibility and Self-Integration

According to Brandom, the kind of responsibility one undertakes in making judgements or forming maxims involves an obligation to integrate those new commitments with the whole that comprises one's previous commitments (RP 35). "Engaging in that activity" obliges one to "weed out materially incompatible commitments" and ultimately "produces, sustains, and develops a synthetic unity of apperception: a *self* or *subject*" (RP 36). This notion of self is not some extraneous feature of Brandom's semantic-pragmatist program, but rather it is essential to the very intelligibility of his account of the social practice of deontic scorekeeping that underwrites normativity in the first place:

The notion of one performer who is responsible for two different claims is implicit in the practical acknowledgement of relations of inferential consequences among claims. One interlocutor is not responsible for the inferential consequences of commitments undertaken by another; such consequence relations govern only the commitments of a single interlocutor. [ME 559]

In other words, the so-called scorekeeping practice of tracking normative statuses (or of assessing the compatibility of commitments and their inferential consequences), which is so central to Brandom's version of semantic pragmatism, initially arises only under the condition that certain attitudes or commitments are recognized as having been undertaken by individual selves. Responsibility is always bound up with the responsibility of being a unified self. Hence, the coherence of selfhood is correlative to that of social practice.⁶

My chief concern with Brandom's characterization of this integrative notion of responsibility has to do with its potential application to the practical (moral) dimension of human experience and the relatively weak form of necessity seemingly implied therein. In other words, I am concerned above all with the first leg of normativity mentioned above: and, specifically,

⁶ Apparently, this move in Brandom depends upon the existence of being capable of making (or expressing) such commitments, and also on a social order in which those commitments are recognized or at least, in principle, recognizable. One is responsible only for the inferential consequences of commitments one has made, or one is recognized by others as having made them.

whether the form of necessity implied by the task-responsibility of integration sufficiently obligates or binds an agent in the relevant sense. I suspect that it does not. Brandom's more nuanced theory about the objectivity of conceptual contents might be read, however, as providing a framework within which a stronger form of obligation can be advanced within certain contexts. In the following section I explain how Brandom's understanding of objectivity as a structural feature belonging to each individual scorekeeping perspective helps explain why judgements about objects (or judgments with representation purport) are bound by norms in the relevant sense. I will also argue that such an account cannot, on its own, explain how or why one's actions are similarly constrained by a sense of obligation (since moral obligations have little to do with objective correctness in the sense of representational purport).

Theoretical and Practical Normativity in Kant

It might prove helpful to begin by contrasting the kind of responsibility that the deployment of concepts seemingly entails within a roughly Kantian framework to the kind of integrative task-responsibility envisioned by Brandom. Keeping within Brandom's idiom, we might say that the metaphysical deduction of the categories represents Kant's own effort to explicate what one is in fact already committed to insofar as one takes oneself as having any empirical cognition, that is, any representations of objects; whereas the transcendental deduction justifies that commitment by establishing the objective validity of the categories (that is, their application to objects) on the basis of their indispensable synthetic role in transcendental apperception, which might accompany any such representation. Given these *a priori* concepts, the task of integrating future theoretical endorsements (or judgements about empirical objects) with the whole of one's previous commitments or beliefs is always already circumscribed by antecedent commitments that are implicit in all such objective knowledge claims. In other words, it would appear that some norms are already present at the beginning – if not strictly before judgments are made, then at least from the very moment one enters into the game of making judgments. Brandom writes:

This is part of what Kant means by calling [the categories] "pure" concepts...and saying that our access to them is "a priori" – in the sense that the ability to deploy them is presupposed by the ability to deploy any concepts. [RP 55]

So, when it comes to theoretical judgments, Kant claims that normativity gets its grip on us by way of what is already conceptually implicit in all such judgment—namely, the *a priori* concepts of the understanding.

Kant might also be read as saying that the moral law performs a parallel function within the practical domain: whereas the concepts of the understanding represent necessary conditions for the cognition of objects, pure practical concepts (or the objective principles of volition) determine necessary conditions under which an autonomous will (and, thus, genuine agency) first becomes possible.⁷ Just as judgments are bound by the application of pure concepts of the understanding insofar as one is a knower of objects, so too one's actions must be circumscribed by the pure principles of practical reason insofar as one is a genuinely autonomous agent. In either case, the transcendental demands placed upon knowers and doers precede any of the particular demands issuing from the specific judgments and actions one makes precisely because these demands articulate necessary conditions for claiming representative purport or autonomous agency.

It is important to note, however, that these conditions, though constitutive for cognition and action, exhibit precisely the kind of necessity rightly associated to normativity since they not only imply the work of spontaneity but also certain standards of application, which may or may not be met. Though they oblige, they do not compel. One can, after all, fall short of grasping an object, as in the case of the theologian's transcendental misapplication of the categories in the cosmological argument,⁸ or fail to adopt a permissible maxim, as in the case of the dishonest person who makes a false promise in order to secure a loan.⁹ Within this Kantian picture, the responsibility to integrate or reconcile past and present commitments is itself always already bounded by requirements pertaining to the deployment of pure concepts—in both the theoretical

and practical domains. In the absence of this additional constraint, nothing aside from the rather weak requirement to maintain the internal coherence of one's commitments could underwrite the sense of obligation associated with normativity.

In case it is not already obvious why I claim that theories of integrative responsibility which lack additional support (that is, lack some further constraint such as provided by the pure concepts in Kant) are too weak to satisfy the first condition of normativity, consider the following: The judgment of the believer who is prepared to endorse whatever additional claims follow from the conclusion of his erroneous proof of God's existence and the action of a dishonest man who is prepared to carry out his deceptive plan with the outmost consistency could both be described as having achieved some kind of internal coherence; but neither the former's judgement nor the latter's action will satisfy the additional requirement imposed upon us by reason in its practical or theoretical forms. This means that one's obligation to integrate one's future commitments with the whole of one's past commitments is, for Kant, always already grounded in a stronger obligation involving the deployment of pure concepts—an obligation that is stronger because it is genuinely necessary, though not causally necessitating.

Self-Integration and One's Answerability to Objects

Brandom seems to agree that something beyond mere coherence is needed in order to explain the normativity of concepts, but he rejects any view according to which conceptual content is said to be fixed transcendently, or in advance of the empirical processes in which those very concepts are employed.¹⁰ Transposing the Kantian picture into a linguistic key, Brandom argues that it involves a two-staged theory in which semantic questions are thought to be fully settled before epistemological ones arise.¹¹ And he argues that

⁷ Immanuel Kant, *Critique of Practical Reason*, transl. Mary J. Gregor, Cambridge, UK: Cambridge University Press 1997, 5:19-20.

⁸ Immanuel Kant, *Critique of Pure Reason*, transl. and ed. Paul Guyer and Allen W. Wood, New York, NY: Cambridge University Press 1998, A609/B637.

⁹ Immanuel Kant, *Groundwork of the Metaphysics of Morals*, ed. and transl. Mary Gregor, Cambridge, UK: Cambridge University Press 1998, 4:404. [Henceforth cited as *GMM*]

¹⁰ By empirical process I mean the complex social practices governing the application of concepts and the assessment of the commitments (and their various consequences) entailed by those applications. I do not (merely) refer to cooperation of the faculties of spontaneity (concept) and receptivity (intuition).

¹¹ Robert B. Brandom, *From Empiricism to Expressivism: Brandom reads Sellars*, Cambridge, MA: Harvard University Press 2015, p. 7. [Henceforth cited as *EE*] Indeed, the rejection of foundationalist epistemologies

this view, in turn, rests on an untenable conception of "semantic purity" that simply fails to explain the bindingness of concepts.¹² Brandom asks:

What holds fixed, in advance, the commitments one undertakes by applying [a concept], if its content is wholly up to the "spontaneous" activity of the subject? The Kantian division of semantic and epistemic labor seems unable to exclude the possibility that "whatever seems right to me is right" – in which case the issue of correctness does not get a grip...There is nothing in the picture to confer *determinate* contents on concepts, nor to hold them in place *as...determinate*.¹³

Brandom's own solution to the problem of establishing this sense of bindingness is to tug Kant in the direction of Hegel, and to suggest that what cannot be accomplished through a division of semantic and epistemic labor at the level of the individual subject may in fact be achieved "by a genuinely *social* division of labor" (RGH 13). In one sense, then, Brandom will claim that the objectivity of conceptual norms depends upon meanings fixed by social practices embodied by a community of speakers. So, to cite one of his routine examples, "it is wholly up to me whether I assert that the coin is copper – rather than manganese, say." But, he continues, it is

not up to me what else I have committed myself to by claiming that, and what would entitle me to that commitment. The metallurgical experts my community charges with the care and feeding of the concept "copper" will *hold* me responsible for having

which claims that certain privileged representations form an autonomous stratum whose conceptual contents are fixed in advance lies at the heart of the semantic pragmatism program that Brandom tries to generate out of insights supplied by Quine and Sellars. From Sellars, he adopts a healthy suspicion of empiricism's mythological beliefs in sensory datum as a privileged representation. Combined with Quine's lessons about aprioricity, this will also result in a suspicion of the privilege Kant grants to the pure concepts.

¹² Brandom sometimes refers to this more generally as a semantic "atomism," especially when he encounters the corresponding problem on the side of empiricism (EE 148-9).

¹³ Robert B. Brandom, "Reason, Genealogy, and the Hermeneutics of Magnanimity," p. 12. Online access July 4, 2021, http://www.pitt.edu/~rbrandom/Texts/Reason_Genealogy_and_the_Hermeneutics_of.pdf. [Henceforth cited as RGH]

committed myself to the coin's melting at 1084° C., and to have precluded myself from claiming that it is an electrical insulator. [RGH 13]

Here are the beginnings of an explanation of how the responsibility involved in integrating one's commitments into a coherent whole might actually entail the appropriate (that is, strong, rather than weak) sense of obligation that I am looking for, an obligation that is no longer anchored by *a priori* concepts, but rather by socially-constituted conceptual norms that place constraints on one's integrative efforts. Determining whether I have successfully rooted out materially incompatible commitments is not just a matter of my own attitudes and beliefs – about, for instance, what I happen to take the word copper to mean – but, rather, it is decided by the socially determined objective content of the concepts I employ, and that content in turn determines what follows from what, and thus what I am in fact responsible for insofar as I employ those concepts. Here, I am responsible in a genuine sense, in spite of the fact that the objective content of concepts is no longer given *a priori*.

Nevertheless, Brandom is well aware of the limitations of intersubjective theories of objectivity that simply privilege the global perspective of the community over and above that of any one of its individual members, or in his words an "I-We" construal of intersubjectivity (ME 598-601). Conflating objectivity with what the community as a whole takes to be correct is hardly any better than conflating objectivity with whatever the individual takes it to be. For the possibility that the entire community might itself be wrong in its assessments remains perfectly intelligible, and that possibility cannot be accommodated within a theory that simply dissolves objectivity into a tyranny of the majority, for such tyranny does little more to capture one's sense of objectivity than would a tyranny of subjectivity.¹⁴

¹⁴ Brandom criticizes Crispin Wright's view about the incorrigibility of communally endorsed applications of concepts. He notes, "There clearly are socially instituted norms of this sort. Whatever the Kwakiutl treat as an appropriate greeting gesture for their tribe...is one; it makes no sense to suppose that they could collectively be wrong about this sort of thing. The question is whether conceptual norms ought to be understood as being of this type" (ME 53). And, of course, Brandom thinks there are good reasons to think that they are not.

It is in accounting for a notion of objective correctness that is neither semantically naïve (that is to say, settled by reference to *a priori* concepts) nor reducible to the endorsements of the community as a whole (that is to say, based on an I-We model of intersubjectivity) that Brandom's own theory of semantic pragmatism shows both promise and originality. His account, as laid out in *Making It Explicit*, relies upon the distinction between *de dicto* and *de re* specifications of propositional attitudes in order to show how the notion of objectivity of conceptual content precipitates out of the scorekeeping enterprise as a kind of structural or formal feature of the scorekeeping perspective itself. Suppose Colonel Mustard is shooting a game of pool in the billiard room with the man who helped him carry off the perfect murder.¹⁵ Suppose further that Inspector Japp knows the Colonel is in the billiard room, but he has not the slightest clue that the Colonel is in fact the murderer he has been pursuing all along. A *de dicto* specification of the Inspector's belief that "Colonel Mustard is in the billiard room" contains nothing about the murderer's whereabouts. But the *de re* ascription of the Inspector's belief by his scorekeeping interlocutor (by, say, the accomplice in this case) captures what the Inspector is objectively committed to without even knowing it: namely, that "the Inspector believes 'of the murderer' that he is" (ME 550) in the billiard room. In ascribing this belief to the Inspector, the accomplice is in fact keeping two distinct, but correlated, sets of books: one tracking what the Inspector acknowledges (his deontic attitude), and the other tracking what the Inspector is objectively committed to by way of what he acknowledges (his deontic status). According to Brandom, "Deontic statuses are just consequentially expanded deontic attitudes" (ME 596). The key idea here is the that distinction between what is objectively true and what is simply taken to be true, between statuses and attitudes, is an essential structural feature of the perspective of scorekeeping interlocutors.

Our practical grasp of the objective dimension of conceptual norms — normative assessments of the objective truth of claims and objective correctness of applications of concepts — consists in the capacity to coordinate in our scorekeeping the significance a remark has from the perspective of the one to whom

the commitment it expresses is attributed and its significance from the perspective of the one attributing it. [ME 598]

The objectivity of conceptual norms does not result from social consensus or the linguistic conventions established by the community as a whole, rather it is born out of the complex intersubjective interplay between each individual scorekeeper. In fact, it is a formal feature of the scorekeeping perspective. Insofar as objectivity is said to consist "in a kind of perspectival form, rather than in a not perspectival or cross-perspectival content" (ME 600), Brandom's account bears at least a superficial resemblance to Kant's own, though this formal feature is accounted for dialogically, as a necessary aspect of social linguistic practice, rather than monologically, as a necessary structure of rational being itself.

In any case, it is clear that Brandom has a story to tell about how judgments become answerable to the objects that they purport to be about, and thus how a judger comes to be bound (in the strong sense of the term) by conceptual norms. The kind of responsibility that it brings into view lies upstream from the merely integrative task-responsibility of weeding out incompatibilities between commitments, which, on its own, as I argued earlier, fails to capture the kind of obligatoriness normativity requires. This additional constraint is, as it were, anterior to that task-responsibility in much the same way that the deployment of the pure categories is anterior to the development of empirical concepts in Kant: it ranges over all of my commitments insofar as they involve claims about objects, claims that I take to be answerable to objective facts — or, to put it another way, insofar as they are responsive to my conception of objective correctness. On the one hand, the demands of integrative task-responsibility arise only on the condition that I have made some commitment or another, and the responsibility is then to reconcile future commitments with that earlier commitment and its various inferential consequences. But so long as those commitments can seem to be contingent and thus arbitrary, my obligation to satisfy that demand can also seem arbitrary and contingent rather than necessary. On the other hand, if a conception of objectivity belongs as an essential feature to the scorekeeping perspective (for it is, after all, presupposed by the discursive deontic practice of keeping two sets of books), then human beings, insofar as they are scorekeepers who themselves make judgments about the objective world,

¹⁵ The remainder of this paragraph consists of a modified and greatly simplified version of an example Brandom gives in Chapter 8 (ME 595).

are responsible not only to their past commitments, but also to that objective world, that is, to how things actually are, as that is being understood from within their individual perspectives.

Obviously, this is not the objective world of naïve realism. Nor is it an empirical world grounded in the sensory given or a phenomenal world conditioned by Kantian categories. Rather, it is a rich and conceptually contentful world that is ripe with inferential implications. And it is in virtue of the normative implications of this conceptually determined, intersubjective understanding of objectivity, which is implicit within discursive practice, that the further obligation to reconcile one's various commitments into a unified whole ultimately gets its grip on a judge. The answerability of one's judgements to the domain of objects (understood in this pragmatic sense) cannot be understood as a separate form of responsibility that is totally independent of the task-responsibility of self-integration; on the contrary, the two go hand in hand. It is best understood as the bottom jaw of a two-jawed framework, in the absence of which normativity has no real bite.

The Limits of a Representationalist Account of Normative Force

The theory all-too-briefly surveyed above was originally meant to supply a pragmatist meta-vocabulary that would capture what is expressed by representational vocabulary in terms of social practice. In other words, it was designed to explain what one's assertions about the objective world mean by explaining what one is doing insofar as one is making such assertions. Bearing in mind this fairly narrow focus, it will come as no surprise that the notion of objectivity it advances can only help clarify the nature of normativity insofar as it concerns fundamentally representational locutions, that is, judgments that claim to represent how things are.¹⁶ The problem, at least insofar as one is concerned with the practical dimension of human experience, is that moral norms might turn out to be indifferent to facts about how things are, or they might only be indirectly related to such facts, and therefore it is hardly clear that this story will be of any use when it comes to giving

an account of the nature and source of normativity within the practical domain. So, to the extent that one is interested in the moral dimension of human experience, the work that Brandom's theory of objectivity performs in helping to secure the normative force of the integrative task-responsibility will not pay off. Moral norms get no such grip. Barring some further account that also captures what gets expressed by moral and practical vocabulary, Brandom's semantic pragmatism can do no better than to account for moral normativity in terms of the weak responsibility of integrating one's past, present, and future commitments.

In order to see why some additional explanation—comparable to the one Brandom provides with respect to judging—is needed in order to secure normativity in the field of action, consider the limits of integrative task-responsibility when taken on its own: whatever my past practical undertakings may be, so long as I successfully integrate my future undertakings with them, I have, as it were, passed the test, I have met my obligation. Consistency supplies sole standard of assessment. In the absence of something like Kant's categorical imperative, which would provide some further normative boundary on what is permissible (in much the same way that Brandom's theory of objectivity sets boundaries for our theoretical endorsements), there is no way of distinguishing between, say, Aristotle's *phronimos*, on the one hand, and a merely clever crook, on the other. Just like the practically wise, the cleverest crooks unwaveringly pursue the means necessary for achieving their devious ends. Both remain consistent in their commitment to their respective ends.

This is why the practical task-responsibility of self-integration will need to be accompanied by some further constraint, a bottom jaw that could give it a bite. This would naturally come in the form of a story about which actions are permissible and which are not—just as the theoretical task of weeding out incompatible commitments needed to be delimited by a story about representations that are in themselves objectively correct. While integrative task-responsibility can explain, on its own, why I am responsible for X as a result of pursuing Y, that kind of obligation is merely conditional—it merely says that a commitment to doing X is instrumentally (but not morally) necessary for maintaining or pursuing one's commitment to Y. Kant, of course, had a name for the kind of conditional obligation that such commitments engender. He called them hypothetical imperatives, in contradistinction to categorical ones. From the point of view of Kantian

¹⁶ Brandon writes, "The major explanatory challenge for inferentialists is...to explain the representational dimension of semantic content—to construe *referential* relations in terms of *inferential* ones" (p. xvi).

morality, hypothetical imperatives lack teeth, since they merely represent "the practical necessity of a possible action as a means to achieving something else that one wills" (*GMM* 4:414). Kant believed that unless it can be additionally shown that I ought to have endorsed X in the first place, or that a commitment to X is necessary regardless of what other commitments I undertake (and thus universal), then I am not bound by it absolutely, that is to say, I am not bound in the strong sense associated with normativity.¹⁷ In the absence of something that could do the work of the categorical imperative, even if I had some obligation to root out incompatible practical commitments, I could still go about my life as a kind of moral chameleon, perpetually endorsing ends and then immediately abandoning those ends as soon as I recognize the unwanted commitments they in turn require. One might call this the instrumentalist problem. Barring an account of this kind of moral necessity, the merely integrative theory of responsibility – which might also be expressed in terms of the hypothetical imperative to pursue the means to one's noncompulsory ends – fails to get any grip.¹⁸ If all that were expressed by moral vocabulary was the obligation to pursue the means of the ends I have chosen, and nothing bars me from selecting new ends, or nothing compels me to stick with the ends I had previously selected, then that vocabulary would not be genuinely normative. So, it seems as though an account of normativity grounded in the task-responsibility

of self-integration lacks the ability to do justice to the genuinely normative character of moral discourse.

Self-integration and Self

I believe that Korsgaard's work may help supply the missing element that would allow Brandom's pragmatist approach to more fully account for normativity in the practical domain. Her theory of self-constitution is able to answer the instrumentalist problem in a way that would be acceptable to Brandom, that is to say, in a way that does not require her to endorse the type of two-staged semantic theory that Brandom had accused Kant of employing. She does this by showing that the categorical and hypothetical imperatives are really two sides of the same coin – they are both principles of practical reason that serve to unify the will. This move not only has the effect of placing a kind of downward pressure on the principles of practical reason (showing how their contents emerge from within the socially determined space of reasons rather than out of the thin air of a priori reasoning) but it also reveals how the weak form of responsibility associated with self-integration can be joined to a much stronger form of responsibility – the kind of responsibility that Kant has associated with an *a priori* foundation, namely, the pure practical reason, but which Korsgaard will now associate with the role of concrete practical identities.

At first sight, Korsgaard seems to pursue a line of thought that would only yield the weak obligation associated with the task-responsibility of self-integration. She argues that "action is self-constitution" and that "what makes actions good or bad is how well they constitute you."¹⁹

The task of self-constitution involves finding some roles and fulfilling them with integrity and dedication. It involves integrating those roles into a single identity, into a coherent life. [SN 25]

So far, so good. But we have seen time and again that this kind integrative requirement, when taken on its own, lacks normativity in the strong sense. Korsgaard argues, though, that this act of self-integration actually involves an additional requirement – not one that is extraneous to the first requirement, but rather one that is implied or presupposed by it. And that is the recognizably Kantian requirement that "we can only

¹⁷ For Kant, even the existence of a universally shared end (namely, happiness) proves insufficient for establishing an absolute command of the means to that end, since there can be no surefire principle for determining once and for all what means would actually bring that particular end about. Thus, even a hypothetical imperative that incorporates a universal end "cannot, to speak precisely, command at all, that is, present actions objectively as practically necessary" (*GMM* 4:418). Therefore, Kant claims, "the question of how the imperative of morality is possible is undoubtedly the only one needing a solution" (*GMM* 4:419).

¹⁸ Or, if it tries to gain traction by insisting that such chameleon-like behavior is just the sort of violation that prevents one from becoming a self (that is, a unified agent), then it faces a different, but equally troubling, consequence: namely, it would then merely require that, for instance, corrupt politicians carry out their treachery in a ruthlessly consistent fashion, and that they do not cave to political pressures when justice comes calling.

¹⁹ Christine M. Korsgaard, *Self-Constitution: Agency, Identity, and Integrity*, New York, NY: Oxford University Press 2009, p. 25. [Henceforth cited as SC]

attach the 'I will' to our choices if we will our maxims as universal laws" (SC 76). In other words, the test of self-integration (which, a moment ago, I had characterized in terms of the merely weak form of obligation connected with hypothetical imperatives) is itself bound up with the test of universalizability that is more typically associated with the categorical imperative. This claim — namely, that hypothetical imperatives reflect an aspect of the categorical imperative — is Korsgaard's distinctive contribution (distinctive in that it represents a rather dramatic departure from what Kant is generally taken to have claimed).

Korsgaard defends her position by contrasting it with the possibility of what she calls "particularist willing." To see why hypothetical imperatives require willing one's maxims as universal laws (as is the case with the categorical imperative)

we need only consider what happens if we try to deny it. If our reasons did not have to be universal, then they could be completely particular — it would be possible to have a reason that applies only to the case before you, and has no implications for any other case. [SC 72]

A particularist will would therefore be one determined by immediate inclinations or impulses, rather than principles. This, it must be stressed, is not the same thing as acting on the principle "I shall do things I am inclined to do, simply because I am inclined to do them" — for that would actually amount to acting on a universal principle, namely the principle to treat inclinations as such as reasons. A truly particularist will would "eradicate the distinction between a person and the incentives on which he acts." Instead of unifying the self, a particularist will would reduce one to "a series, a mere heap of unrelated impulses, and as a result, nothing resembling a self would remain" (SC 76). So, Korsgaard writes, "the reason that I must conform to the hypothetical imperative is that if I don't conform to it, if I *always* allow myself to be derailed by timidity, idleness, or depression, then I never really *will* an end" in the first place — the very idea of willing an end would remain unintelligible (SC 69).

It is worth noting that Korsgaard is addressing the same worry I had expressed about theories that appeal solely to the task-responsibility of integration when trying to explain the normativity of moral discourse. The worry (which I called the instrumentalist problem) was that this task-responsibility alone could not justify the stability of my commitments. That is to say, it could not explain why I must maintain any of my

commitments over time, unless those commitments were themselves anchored in something else that was genuinely normative (in the way that our judgements are thought to be answerable to objects). Korsgaard's argument against the possibility of a particularist will demonstrates that one needs to act on reasons or principles in order to be an autonomous self (rather than a mere heap of impulses). That, in turn, implies that the (principled) stability of one's commitments is a necessary condition of agency itself. So, Korsgaard ultimately avoids the instrumentalist problem and explains the source of normativity within the practical domain without having to anchor her theory in a transcendental story about pure practical concepts, or a theory about objective normative facts.

Moreover, it is precisely because Korsgaard conceives of the reasons or principles upon which individuals must act as being supplied by an individual's contingent practical identities (rather than transcendental concepts) that her theory is ultimately of a piece with Brandom's own pragmatist strategy. In other words, she too might be read as turning Kant's transcendental approach on its head and giving it a social-pragmatist twist. According to Korsgaard, it is not the operations of pure practical reason but rather the "practical conceptions of our identity" as teachers, artists, or public health experts that are said to "determine which of our impulses we will count as reasons" (SN 129). Thus, moral obligations arise, at least in part, from one's practical identity, taken in a fairly concrete sense:

The conception of one's identity in question here is not a theoretical one, a view about what as a matter of inescapable scientific fact you are. It is better understood as a description under which you value yourself. [SN 101]

The necessity to act on universal principles, even when engaged in merely hypothetical imperatives, explains how we can in fact "be bound by obligations that spring from conceptions of our identities which are not in themselves necessary" but are rather contingent, which is to say, honed through social practice (SN 129).

Nevertheless, in order to complete her story about the nature of these identities, I would argue that Korsgaard would need to follow Brandom in taking Kant one step further in the direction of Hegel. Just as Brandom had shown how the normative force of objectivity precipitates out of the soup of intersubjective scorekeeping, I think Korsgaard could show that the

normative force of these identities emerges in a similar way out of the social nexus of reciprocal, recognitive attitudes and statuses. After all, whether or not I am really a teacher, an artist, or a public health expert is not simply up to me; it is not, as Korsgaard sometimes seems to suggest, merely a matter of my own self-conception. If it were, then these would not be roles at all—for they would not be accompanied by rules that one could fail to follow or standards one could fail to meet—and in that case Korsgaard's normative theory would come apart at the seams. Rather, such statuses can only be conferred on me by those who I in turn recognize as having the authority to confer those statuses. As Brandom puts it, "To be, say, a formidable

[chess] club player, I must be recognized as such by those I recognize as such" (*RP* 71). As a consequence, my ability to act on principles that serve to unify and constitute my will is equally dependent upon the reciprocal, recognitive attitudes within a community, which determine what roles there are to play and the rules by which we can play them. Therefore, while Korsgaard's theory about the role practical identities play in constituting selfhood can help secure the kind of normative force missing in Brandom's account of self-integration, Brandom's pragmatist program can in turn provide a useful meta-vocabulary in terms of which we can better understand how these identities come to be the sources of normativity that they are.