Cognitive Virtue and Cognitive Self-determination
[Presented at "Engaging McDowell" conference, 13 July 2010]

Lissa Merritt

1. *Two sides to the idea of rational self-determination*

I take as my starting point an issue that surfaces here and there throughout much of John McDowell's work, but which asserts itself especially clearly in a paper called "Self-Determining Subjectivity and External Constraint". The question is how best to understand the notion of self-determination that lies at the heart of the rationalist tradition, conceived broadly to include its Kantian and post-Kantian developments. In raising this question, McDowell is particularly keen to give due acknowledgement to the nature of *empirical* cognition. This accounts for his abiding interest — from *Mind and World* onwards — in Kant's account of sensible intuition, especially as it develops under the lights of the second-edition Transcendental Deduction of the categories. In Kant, McDowell finds the resources to formulate the self-determination idea without doing violence to the obvious fact that in sensible experience the world has a hand in making our minds up for us.

In the "Self-Determining Subjectivity" paper, McDowell sets aside a common rendering of the self-determination idea on the grounds that it poorly suits the status of experience as empirical cognition: this is the conception of self-determination as "freely responsible cognitive activity, [or] making up one's mind". McDowell instead takes up the idea that the conceptual norms governing cognitive activity of *any* sort — be it experience or judgment — "must be capable of free acknowledgement by the subjects who engage in the activity" (96). Self-determination is then to be understood in terms of a required capacity to take a reflective — and presumably critical — attitude towards the norms themselves.

I take this basic rendering of the self-determination idea to be as much Kantian as it is Hegelian. McDowell also suggests as much when he points to two "sides" of this self-determination idea — two different, but ideally mutually supportive, renderings of a single

conception of rational self-determination. The Kantian version (which, it must be noted, is rooted in Leibniz's *New Essays*) points to certain fundamental norms as *constitutive* of any capacity for self-determined thought. To come into the use of one's reason is to have at least a tacit grasp of the relevant norms, which are principles allowing for coherent thought and experience of objects. It then follows that the subject is necessarily capable of free acknowledgement of the norms governing cognitive activity, which is the basic self-determination idea under consideration.

But this generically Kantian conception of rational self-determination tends to undermine itself, impairing the very critical ideal that it is presumably meant to express. The supposition that we can identify a set of norms wholly constitutive of any cognitive capacity encourages us to think of them as isolated from human history, and hence as fixed for all time. This, in turn, encourages their uncritical acceptance as givens that are beyond criticism themselves.[1] Hence McDowell suggests, in the conclusion of his paper, that Kantian one-sidedness gives way to a "pre-critical platonism" (107).

The needed corrective is to come from the other "side" of the self-determination idea — the Hegelian side. The norms governing cognitive activity can only be recognized as such from within a historically specific framework of concrete practices, and a shared form of life. But if this side assumes undue dominance, we are liable to infer — with recent neo-Hegelians, like Bob Brandom — that our *recognition* of fundamental norms is what *confers* authority onto them (105). These neo-Hegelians, McDowell complains, are committed to conceiving of this act of

---

[1] The self-determination idea is maintained through the identification of norms that are held to be *internal to* our cognitive capacity — but at the expense of admitting them as given dispensations, traceable perhaps to the benevolence of our creator. Kant's platonism about the forms of thought also infuses the sanguine promise of critical philosophy to bring metaphysics into a "permanent state" (Bxxiv). See also *P*366.

In *Mind and World*, McDowell refers to "rampant platonism" as the tendency to think of the fundamental principles of cognition in isolation from human life (see 77-8, 84, 88); in its place, McDowell advocates a "naturalized platonism": it is a platonism because it involves the idea that cognitive self-determination is to be understood in terms of a subject's capacity to reflect upon constitutive cognitive principles, but it is a "naturalized" platonism because it also supposes that such reflection can only take place within a concrete form of life, and thus through "second nature" (91-2).

recognition as taking place "in a normative void" (107). Moreover, such talk of instituting norms locates the source of their authority in *particular acts* of recognition — even if these acts of recognition are undertaken collectively, by an entire community. We then rule out the possibility of an entire community's going wrong about the norms; and this, in turn, opens the door to relativism. So it turns out that Hegelian one-sidedness undermines the critical ideal no less than Kantian one-sidedness does. Each side needs the other, McDowell suggests, if we are to hold onto a viable conception of rational self-determination at all.

Hence McDowell concludes the "Self-Determining Subjectivity" paper with a gesture towards a further project: namely, to work out the details of an appropriately balanced, or "two-sided", account of rational self-determination.[2] Such a view would be stable, no longer prone to undermine the reflective and critical ideal that it is presumably meant to express, or at least make intelligible. Now, McDowell suggests that we ought to take up such a project so that we may better understand *Hegel* (107).[3] But shouldn't an appropriately balanced view of rational self-determination help us to understand Kant better, too? That, at any rate, is my concern here. Of course, it lies well beyond the scope of a short talk to sketch out a Kantian picture of two-sided self-determination. (Nor do I have the sketch tucked up my sleeve, I am sorry to say.) My work here is far more preliminary than that. I mostly want to make a fresh suggestion about where, in Kant, we might go looking for the resources needed to compile such a picture.

2. *Where should we look?*

The traditional place to go looking for Kant's views on self-determination is, of course, the Transcendental Deduction of the categories. And indeed, this chapter of the first *Critique* can profitably be read as the meditation on cognitive self-determination. But I doubt whether the

---

[2] Despite this concluding gesture, McDowell himself seems to tack towards the "generically Kantian" side of the self-determination idea (106) in this and other papers — though I think this may be largely attributable to his debate with Brandom and others about the institution of norms.

[3] In singling out Hegel in this way, I suppose he is simply following through on his shot against the relevant neo-Hegelians.

Deduction ought to be our sole text — or even our first point of reference — if our project is to draw up a "two-sided" picture of cognitive self-determination from Kantian resources. For while the Deduction is concerned throughout with the self-determination of the cognitive subject, it does so with a particular aim in mind: namely, to vindicate a claim about the categories as constitutive of our capacity for *theoretical* cognition. And this is where the Platonism seeps in: for Kant's sanguine hope to bring metaphysics into a "permanent state", as he says in the *Critique*'s Preface, rests on the promise of the Deduction.[4] Here Kant has in mind a metaphysics of *nature*; but he cherishes the same hope for a metaphysics of *morals*, the prospects of which depend upon the parallel clam that the moral law is constitutive of reason in its *practical* capacity. So while the principles drawn up from the categories tell us what it is to be a thing, or to figure in the domain of nature, the moral law tells us what it is to be a person, or to figure in the kingdom of ends. Metaphysics can be brought into a permanent state, because the principles in question are held to be timelessly constitutive of any cognitive capacity in its respective theoretical and practical employments.

My point here is simple. As long as we stick to those passages where the self-determination idea is bound up with an argument about constitutive cognitive principles, it will be difficult to make out whether Kant provides the resources for a two-sided picture. For this is precisely the angle from which the Hegelian complaint is lodged, that the Kantian self-determination idea easily falls into a bad Platonism. This provides preliminary grounds, perhaps, for rejecting the Deduction as our primary text on self-determination.

However, I would like to understand the problem of one-sidedness in a general way, without prejudice in favour of the Hegelian complaint. Is there some general error that we commit, whether we fall into the one trough or the other? One pitfall is an uncritical Platonism — that is to say: dogmatism. The lure of dogmatism, Kant suggests, is the promise of a resting place for inquiry. On the other side, the pitfall is relativism. And the lure of relativism can be

[4] Bxxiv; see also *P* 366.

glossed in the same general terms: it, too, promises a rest from further inquiry — only now in the form of an allowance to accept that one set of norms is, at least in principle, just as good as any other. The common error, regardless of which trough we fall into, is cognitive complacency.

This suggests that cognitive self-determination might be understood as a kind of humility. The properly self-determined cognitive subject has a certain strength of mind — and a readiness — to take a critical attitude towards the norms governing cognitive activity. And since these norms comprise one's cognitive capacity, it follows that to take a critical attitude towards the norms themselves is in turn to take a critical attitude towards one's own cognitive constitution. Thus it is, in turn, to take an appropriate interest in what we might call one's *character* as a judging subject, or cognitive agent.

This is, I think, simply a variant on the self-determination idea that McDowell urges us to take up. This particular spin on it allows us to see, now from another angle, why the Deduction might not be our most appropriate primary text. While the Deduction surely has much to do with cognitive self-determination, it addresses the issue in highly abstract terms that make no reference to our ground-level cognitive practices. And when we sever this connection, we lose sight of the fact that self-determination is not so much a *fact about us* as rather a *problem for us*. We often comport ourselves quite badly in our cognitive practices, and fail to be "self-determined" at all. We often give our minds over to prejudice, which Kant glosses as a "tendency toward heteronomy of reason" (5:294). To give our minds over in this way is to fail to take an appropriate interest in one's own cognitive agency. (There is much more to be said here, both about Kant's view of prejudice, which is subtler than it may appear, and about the possibility of excessive self-examination — but I cannot take up these issues now.[5]) The idea is that self-

[5] McDowell reminds us of Gadamer's rejection of the standard enlightenment view about prejudice — wherein prejudice is always a bad thing, a giving of one's mind over to an alien force (*M&W*, 81 n.14, noting *Truth and Method* 277-85). I think that Kant's views about prejudice are more complex than is sometimes recognized. He points to three sources of prejudice are *imitation, inclination* and *custom* (see 9:76). The complexity stems from Kant's clear recognition that imitation, inclination, and custom are operational facts about human cognitive psychology: it is simply a fact about us, for instance, that inclination "gets the first word" (*KpV* 147). What distinguishes the enlightened or reflective individual is

determination should be conceived as an ideal that we are called upon to realize, or at least approximate, in our cognitive practices. Indeed, as I will suggest, one's ability to participate in a cognitive practice at all depends upon one's having this ideal in view as a point of orientation.

Now, if self-determination is a normative ideal, rather than a given dispensation, then perhaps we should consider it along the lines of virtue. So I will begin by drawing out some basic points from Kant's account of moral virtue — but only what is absolutely needed to take up his conception of *cognitive* virtue, which is my concern here.[6]

3. *Virtue*

Kant speaks of moral virtue in a puzzling variety of ways: sometimes as *health* (6:384; 409; 419); sometimes as *strength* (6:384; 409); sometimes as a cultivated excellence or *perfection* (6:419); and sometimes as an *ideal* (6:383; 6:409). Presumably we can simplify this under two headings: health and perfection. On the face of it, though, the two models seem to be at odds. Health is ordinary and common (at least we tend to think so when we enjoy it), whereas perfection is extraordinary, if it is even attainable at all. But Kant presents cognitive virtue under both models. So we will need to examine the two models, and consider their relation to one another, in order to understand Kant's view of cognitive virtue.

It might seem that the purpose of the two models of virtue is simply taxonomical, so that some virtues would be understood according to the model of health, and others according to the model of perfection. This is suggested, perhaps, by Kant's division between "negative" and "positive" duties to oneself. Negative duties are explained according to the model of health:

---

her propensity to recognize this "first word" for what it is — a merely "subjective" determination and not in itself sufficient grounds for a cognitive determination.

On the issue of excessive self-examination: despite the centrality of reflective self-examination to Kant's view of the ideal of enlightenment, I take it that even Kant recognizes that self-examination can be excessive, or at least in some manner inappropriate or overweening. This is a difficult issue, both philosophically (since it belongs to the difficult topic of the role of reflection in judgment), and historically (since I think it may be related to Kant's ambivalence about Pietism). At any rate, I haven't fully thought through the issue yet.

[6] (In fact, Kant's moral cognitivism implies that cognitive virtue could be thought of as the genus of which moral virtue is a species — but this is a matter that I cannot take up here.)

negative duties, Kant says, "*forbid* a human being to act contrary to the **end** of his nature and so have to do merely with his moral *self-preservation*"; such duties, he continues, "belong to the moral **health** [...] of a human being" (6:419). Positive duties, on the other hand, command one to *perfect* oneself, and so are understood according to the model of perfection, or cultivated excellence. Kant says that they "belong to [one's] moral *prosperity* ([...] *opulentia moralis*)" (6:419). So it can seem that the two models of virtue simply help us to sort the particular virtues under two heads.

But this cannot be right. At any rate, Kant speaks of cognitive virtue according to both models. In the *Metaphysics of Morals*, Kant says that we have a positive duty to cultivate our faculties, the "highest of which", he says, is the *understanding* (6:386-7). This would seem to imply that cognitive virtue is a cultivated perfection, a bauble of "moral opulence", rather than the fibre of moral health. But in another passage, from the *Critique of Judgment*, Kant speaks of cognitive virtue according to the model of health. There Kant introduces three maxims by which one would aim to think in a manner that is free from prejudice, broadminded, and internally consistent (5:294-5). He calls them the maxims of "common human understanding" which he explains is a "merely healthy (not yet cultivated) understanding" (5:293). Here Kant invokes *both* models of virtue, and seems to suggest that the basic cognitive virtue at issue would be understood as health, not perfection.

The two models of virtue, I will suggest, are interdependent: they belong together as part of a single picture of virtue. I will not try to argue this point in a general way, but rather will explain it only with regard to cognitive virtue. So let us begin by taking a closer look at the passage just mentioned, about the three maxims. These maxims are:

To think for oneself;

(1) *To think in the position of everyone else*;

(2) *To think always in accord with oneself*.

There is much to say about the content of these maxims, and their relation to one another as a unified package. I will mostly need to confine myself to a general treatment here. Let us begin with the already noted fact that they are maxims of a "merely healthy" understanding. This, Kant insists, is "the least that can be expected from anyone who lays claim to the name of a human being" (5:293). Now, this is striking, since Kant goes on to emphasize that it is actually quite difficult to conduct oneself in accordance with these maxims. Indeed, there doesn't seem to be anything common at all — at least not in the sense of ordinary — about the cognitive conduct of someone who is governed by these maxims. Most of us, for example, lack the broadmindedness specified in the second maxim: most of us are, as Kant says, "bracketed" by our own prejudices, or the "subjective private conditions" of our judgments (5:295) — at least some of the time.

In order to understand the passage, we will need to examine the concept of health more closely. True health cannot be understood as the mere absence of disease. Consider the boy in the bubble, whose story I remember vividly from my childhood. He might have been kept free from disease, but he certainly wasn't healthy. Health is the expression of an active principle, even if it is one that for the most part response to challenges from without. This is why Kant explains the model of virtue as health as the "*preservation* of [one's] nature in its perfection (as *receptivity*)" (6:419): the "as receptivity", I take it, points out that the active principle in question responds to challenges *from without*. But what exactly is it that gets preserved? Here it is important to remember that Kant's doctrine of virtue belongs to what he broadly refers to as a "moral anthropology": the moral anthropology goes beyond the establishment of the moral law as the constitutive principle of practical reason, and takes into account certain facts about human nature, or reason's embodiment in us.[7] In the specifically moral case, this plays out in terms of the idea that we have some natural predisposition to the good — a rudimentary orientation

[7] (Thus the project of a moral anthropology begins already in the third chapter of the *Critique of Practical Reason*, when Kant introduces moral feeling.)

towards it through moral feeling, the bare capacity for which *is* a natural endowment. But we have an obligation to *develop* this natural endowment, to cultivate moral feeling. In doing so, we strengthen this orientation towards the good. This is the basic idea of moral health, or a sound moral disposition. Presumably a parallel point can be made for the case of *cognitive* virtue. We have, as a natural endowment, some rudimentary orientation towards the end of knowledge. Thus Kant presumably admits the uncontroversial Aristotelian given, that we all by nature desire to understand. But we do not by sheer natural endowment have an adequate pragmatic grasp of *what it is* to aim at cognition. This is where the three maxims come in, and the point about cognitive health.

While Kant holds that anyone who has come into the use of his reason cannot help but to have some basic orientation towards the end of knowledge (in the general epistemic case), or to the good (in the specifically moral case), he also emphasizes that none of us is ever perfectly or infallibly rational. This is equally a claim about human nature. And it effectively means that a person does not generally do a very good job of keeping his sights on what Kant refers to as the "end of his nature" (6:419). Virtue as health is the basic acknowledgement of one's nature; it can be understood as the stable and resolute orientation towards the relevant end. This is why Kant also refers to virtue as *strength*: it is, he says, the "*strength* of resolution in a human being endowed with freedom, [...] his strength insofar as he is in control of himself [...] and so in a state of *health* proper to a human being" (6:384). It is a strength that "is recognized only through the obstacles it can overcome" (6:394).[8]

[8] I think that this is why Kant says, in a nearby footnote, that "enlightenment is easy *in thesi*, [but] *in hypothesi* it is a difficult matter that can only be accomplished slowly" (5:294n.). I take this to mean that it is easy to defend enlightenment as a good, but difficult to keep one's sights on it so that it may figure as a basis for action. Kant goes on to say that enlightenment is "very easy for the person who would only be adequate to his essential end" — which presumably means that enlightenment *would* be easy for one who never attempted to go beyond the bounds of our cognitive capacity. Of course, Kant is famous for maintaining that it belongs to the "peculiar fate" of human reason that it cannot help asking questions that it is unable to answer: and so presumably there can be no "easy" enlightenment, at least not for us. Hence Kant concludes by saying that "it must be very difficult to maintain or establish the merely negative element (which constitutes genuine enlightenment) in the way of thinking" (5:294n.).

Our passage on the three maxims gives us some idea of what the relevant obstacles must be. The first maxim, *to think for oneself*, directs one not to be passive — or "heteronomous" as Kant says here — in the employment of one's reason (5:294). The second, *to think in the position of everyone else*, directs one not to confuse "the subjective private conditions of judgment" with objective grounds for judgment (5:295); it warns against arbitrariness, or being insensitive to the fact that whatever can figure as an object of cognition must be accessible (at least in principle) to *any* judging subject. Since the third maxim is supposed to arise from "the combination of the first two" (5:295), it should tell us something about the core principle at issue. It concerns a requirement *to think always consistently with oneself*. Inconsistent representations can be maintained by a subject only to the extent that she is either in the dark about their inconsistency or about their being *her own* representations. So the third maxim concerns the interest one ought to take in oneself as the coherent *source* of one's own cognitive determinations — however much may simply be given to one from without, whether through habit, instruction, testimony, or sensible appearance. This suggests, then, that the complete battery of maxims has to do with the preservation of one's own cognitive agency. This we can only preserve, Kant thereby implies, through the practice of judging in community with others.

But what about the other model of virtue, that of a cultivated excellence, or a self-developed perfection? As we have seen, Kant uses this model for the "positive" duties one has to oneself. Such duties involve taking discretionary ends — choosing, for example, to cultivate a talent for music or mathematics, rather than painting or medicine. So the positive duty is to *cultivate* one's cognitive capacity *in some particular way or another*. However, in the passage on the three maxims Kant seems to set aside the idea of cultivated perfection altogether, to focus entirely on the idea of basic cognitive health. However, this appearance is deceiving — though it will take a bit of legwork to see why.

When Kant says that virtue is an ideal, he points to it as a standard which we are called upon to "continually approximate". An ideal cannot be met with in the flesh, or found manifest

in empirical introspection. So the model of virtue as an ideal is related to Kant's view that humility is the foundation for any sound moral disposition (*KpV* 161). Humility, Kant says, is a "sublime state of mind" (*KU* 264). It is sublime because we look up to — or are literally elevated by — an ideal that can only be conceived in pure thought; yet for Kant a sublime state of mind must also involve some empirically introspective awareness of how we necessarily fall short of any such ideal. Still, it would be arrogant to suppose that we can even ascertain our distance from the relevant ideal. So the ideal does not exactly figure as a standard of measure, but rather as a point of orientation, and in that sense as a governing principle of our cognitive practices.

Although there is much more to be said about humility — a sadly neglected and much misunderstood topic, I think — we need to keep our sights on the parallel case of cognitive virtue. Is there an *ideal* at issue in the passage on the three maxims? Since Kant says that the third maxim arises from the "combination" of the first two, we might look there for the fundamental characteristic of cognitive virtue. The third maxim is *to think always consistently with oneself*. It is constitutive of the ideal in question that we *strive for*, and to some extent *actually achieve*, unity and coherence in light of both our own experience and the claims of other judging subjects. To have this as one's end, I suggested, can be understood as taking the appropriate interest in being the coherent *source* of one's own cognitive determinations. Thus the third maxim allows us to see how cognitive virtue relates to the problem of self-determination. It is not a given *that* we are the coherent source of our own cognitive determinations: any number of things that we think and believe and want may put pressure on such a self-conception. But a reasonably "healthy" cognitive subject appreciates this as an abiding problem: it is the problem of self-constitution, or the cultivation of character.

How can we understand this last point a bit more concretely — and in a way that makes contact with the self-determination idea with which we began? In McDowell's formulation, the self-determination idea has to do with a subject's capacity to freely acknowledge the norms governing cognitive activity. On my variation on the same theme, the self-determination idea

needs to be understood in terms of a certain readiness to take a critical attitude towards the norms governing cognitive activity. Although these are norms that, for the most part, we simply find ourselves with, nevertheless they are (as it were) the material components of our cognitive self-constitution; therefore, to take a critical attitude towards these norms is at the same time to take the appropriate interest in one's character as a cognitive agent.

Kant's account of the three maxims points to this critical ideal in the following way. Since Kant presents the third maxim (*to think always consistently with oneself*) as the synthetic unity of the first two, we can look to it to find the fundamental conception of cognitive virtue. This basic idea, I wish to suggest, is *originality*. The idea of originality emerges in the first maxim, *to think for oneself* — but it is incomplete there, since the bare idea of thinking for oneself is compatible with perfectly arbitrary and meaningless thought. The second maxim, *to think in the position of everyone else*, introduces a point about publicity or communicability that is required to bring the possibility of cognition into view. The third maxim is supposed to arise from the combination of the first two. With the third maxim, we return to the core idea of originality: but now the idea of originality is made complete through its relation to the idea of communicability. This allows us to conceive of originality — or of the critical ideal — as taking the proper concern to be the source of one's own cognitive determinations.

An ideal can be manifest only by degree. We can see how originality admits of degree if we consider the case of judgment. The most rudimentary expression of originality, in the sense under consideration here, is to recognize the relevance of a rule. For the most part, the rules in question are the cognitive norms that one simply finds oneself with, in having been brought up in a particular way, in a particular place, with a particular language, and so forth. When we learn, we often begin by aping the judgments of others; this is why Kant says that "learning is nothing but imitation" (5:308; cf. 7:225). He does not mean that this is all there is to our cognitive development; he simply means that our engagement with cognitive norms is heavily guided from without, at least initially. This guidance lets up once we can recognize for ourselves the relevance

of the rules that we simply find ourselves with. And this is how we can begin to cultivate a reflective attitude towards these norms — this is how the rules can cease to be *merely given*. For it is only then that we can begin to consider whether the norms in question are good ones, whether they really put us onto objects.

It is presumably also the case that the cultivation of the critical attitude requires one's robust involvement in *particular* cognitive practices. And this requires that one take discretionary ends in developing one's cognitive capacity in particular ways — say by pursuing music and mathematics, rather than painting and medicine. For norms are relevant, or not, in the life of a particular practice, and with regard to particular cognitive ends. And this, I think, explains Kant's insistence upon the positive duty of cognitive self-cultivation, and not merely the negative duty of cognitive self-preservation (or maintenance of basic cognitive health).

4. *Self-determination*

I have asked us to treat Kant's account of cognitive virtue as a starting point for working out a two-sided view of self-determination. My claim is simply that overemphasis on the putatively "Kantian" side of the self-determination idea — that is, the idea that the possibility of self-determination can only be explained by invoking timelessly constitutive cognitive principles — obscures an important point, namely that our orientation towards the end of cognition is something that must be strengthened and cultivated. So, while it may be the case that we all by nature desire to understand, we do not by sheer natural endowment have an adequate pragmatic grasp of what this involves. And while Kant may be right to suppose that we necessarily have some rudimentary grasp of the relevant constitutive principles simply by coming into the use of our reason, nevertheless there is something further that each person needs to work out for herself. This is how to conduct oneself in one's cognitive practices, in sorting out here and now what is and is not the case, and what is and is not to be done. Obviously there is much more to be said about the health and perfection of our cognitive lives. But Kant's three maxims provide

the basic framework for understanding how rudimentary orientation towards the end of cognition can be strengthened and cultivated. For this reason, I think, they provide the starting point for thinking about cognitive self-determination. The core idea is that self-determination ought to be conceived as a *problem* for us, rather than as a *fact* about us. We need to recognize how deeply Kantian this idea really is. Our understanding of Kant, and of the normative aspirations of his critical philosophy, depends upon it.