

Making up one's mind

Pascal Engel

to appear in J.L. Marion and R. Morrissey, eds

Le Soi/The Self

Cahiers parisiens The University of Chicago, Volume 3 in September/October (2007).

Dans cet article, l'auteur essaie de décrire le rapport entre savoir ce que l'on veut, qui est une manière de se connaître soi-même, et être quelqu'un de responsable. En replaçant ces réflexions à la fois dans un contexte néo-kantien récent et à travers les reformulations de ce problème opérées par la philosophie analytique, l'auteur évoque dans le détail deux manières d'être responsable de ses croyances, une manière forte et une manière faible. Ce faisant, l'article remet en question plusieurs tentatives de considérer la formation des croyances comme constituant une sorte d'action, et de considérer la connaissance de soi comme étant le fruit d'une capacité à réfléchir qui permettrait au sujet de créer pleinement ses propres croyances.

1. Introduction

A number of writers have recently rediscovered the Kantian idea that there is a close connection between the notions of having a thought, of being a reflexive being and having a self. One way in which this connection is often expressed is through the theme of the responsibility that a reflexive being has for his beliefs and judgments. Here are some expressions of this idea in recent writings:

(1) "Judging, making up one's mind what to think, is something for which we are, in principle, responsible – something we freely do, as opposed to something that merely happens in our lives. Of course a belief is not always, or even typically, the result of our exercising this freedom to do what we think. But even if a belief is not freely adopted, it is an actualization of capacities of a kind, the conceptual, whose paradigmatic mode of actualization is in the exercise of freedom that judging is. This freedom, exemplified in responsible acts of judging, is essentially a matter of being answerable to criticism in the light of rationally relevant considerations. So the realm of freedom, at least the realm of freedom of judging, can be identified with the space of reasons." (Mc Dowell 1998 : 434)

(2) "Kant's reconciliation of us as free in virtue of being rational, with us as bound by norms in virtue of being rational – and so of freedom as constraint by a special kind of norm, the norms of rationality – accordingly involves treating the normative status of moral obligation as instituted by normative attitudes." (Brandt 1994: 51)

(3) "To understand fully the fundamental notions associated with reason, including the notion of reasoning, judgment, change of mind, propositional attitude, point of view, one must have and employ a first person concept.... A subject, or a 'critical reasoner' in this full sense not only must understand the evaluative norms that provide standards that count reasoning good or bad, but also that a subject who is able to understand these norms of reasons must 'immediately be moved by reasons.' To understand reasons one must understand their force and application. So to be aware of these norms involves a tendency to be immediately motivated by them." (Burge 1995: 249, 252).

Many other instances of such neo-Kantian reasoning could be found in contemporary philosophy

(for instance see Korsgaard 1996). The gist of the reasoning, which these quotes illustrate, seems to be the following: one cannot have reasons for thinking that P, or doing that P, without being able to be moved by these reasons, and to take oneself as responsive to them, as able to implement them in one's epistemic and practical decisions. And this also requires that one is able to ascribe these reasons to oneself, and have the concept of a first person. Hence the concepts of reason, of subjectivity and of agency are intrinsically interconnected. In spite of the fact that the interconnection between these concepts seems to be obvious to some philosophers, it makes sense to ask: what is the relationship between the idea that one can know one's own mind, have a form of self-knowledge, and the idea that one is a responsible agent? My aim in this article is to try to spell out in which sense this is correct. I shall distinguish two accounts of responsibility for belief. A strong one, according to which reflectivity and responsiveness to reasons are *constitutive* of rational control of our belief, and a weak one, according to which it is merely a *precondition* upon such a control. Only the weak version seems to me correct. But both have recourse to a specific feature of our self-ascriptions of beliefs: transparency.

2. Transparency

There seems to be no relation at all between the idea of self-knowledge and the idea of being a responsible agent. For, if self-knowledge is possible at all, it is a cognitive achievement, which pertains to our cognitive nature, not a practical achievement, relevant to our agency. So where is the link? It seems to be associated to a certain conception of thought as something active and to a certain conception of reflectivity and of self-consciousness. In what sense? What is the relationship between the fact that I can come to know my thoughts and feelings and the fact that I am [] responsible for these thoughts? Am I responsible for my thoughts in the same ways as I am responsible for my actions?

In order to understand what McDowell, Brandom, Burge and others are up to, let us leave aside here a large chapter of the issue of self-knowledge: knowledge of own experiences, and let us concentrate on cognitive attitudes, such as beliefs. Typically, first-person authority is said to be a feature of our self-knowledge of our experiences: pains, sensations, phenomenal states and the like. The kind of knowledge of oneself at stake here is knowledge of our own beliefs. The interesting fact here is the one noted by La Rochefoucauld:

«Tout le monde se plaint de sa mémoire, et personne ne se plaint de son jugement.» [Maximes 89]

In other words, failures of our memory seem much more common than failures in judgment. To judge, and to know that one judges, seems to be much more secure than any sort of sensation of memory. But what is it to self-ascribe beliefs and thoughts to oneself?

Prima facie, self-ascriptions of belief are ascriptions of the form *I believe that P* where “P” is a first-order belief content, and where *I believe that P* thus expresses a second-order belief. When one says “I believe that P”, one does not simply ascribe to oneself a certain content, but also the *mental state* or *attitude* of belief. But is this *prima facie* analysis correct? Is it really the case that when one thinks *I am thinking that P* one attributes to oneself in addition to the content *P* the content that one *thinks* that P? In other words, is our inner eye, if such there be, in such thinking episodes directed at our *attitude* towards P? Gareth Evans denies this:

□ “In making a self-ascription of belief, one's eyes are, so to speak, or occasionally literally, directed outward – upon the world. If someone asks me: ‘Do you think there is going to be a third world war?’ I must attend, in answering him, to precisely the same outward phenomena as I would attend if

I were answering the question: ‘Will there be a third world war?’” (Gareth Evans, 1982: 225)

In our self-belief ascriptions, the *self-belief* element is, so to say, transparent, and is not present *within* the content at which our “eye” is directed. We only attend to the first-order content itself. This feature has come to be known under the name of *transparency*. When I ask myself the question:

(i) “Do I believe that P?”

I do not ask myself :

(ii) “Do I *believe* that P?”

□ but I ask :

(iii) “Is ‘P’ true?”

which gives way simply to asking :

(iv) P?

Let us call this the *transparency test for self-ascriptions* of beliefs. The test just says that a question of the form (i) is answered by just answering question (iv). The first-person question about one’s belief is answered by reference to the same reasons that would justify an answer to the corresponding question about the world. From within the first-person perspective, I treat the question of my belief about P as equivalent to the question of the truth of P.

What is the sense of the transparency feature? It is not unproblematic. There are several strands in the interpretation of this feature.

(1) A first, and somewhat obvious, strand is Kantian. It was explicit in Strawson’s *Individuals*, which Gareth Evans was commenting upon when he proposed the transparency test. The “I”, the self, is not part of my representations when I represent something. Translated in terms of belief and self-ascriptions, it means that neither the attitude of believing nor the subject of this attitude is explicitly represented in the content of a self-ascribed belief.

(2) A second strand is Wittgensteinian: what the transparency feature of self-ascriptions of belief show is that in so far as self-knowledge has to rely on ascriptions of beliefs to oneself, self-knowledge is not knowledge of something “inner”.

(3) A third strand is an externalism about thought content: the content of our intentional mental states is determined by what is external, by the outside environment of the speaker.

But neither of these views is unproblematic. Let us start with the Wittgensteinian view. Its main point is not simply that a self-ascription of belief does not point to something inner or to some Cartesian substantial self, but it is also epistemological. As Crispin Wright (1995) puts it, “there is no substantial epistemology” of self-knowledge. To ascribe a belief to oneself is not to report anything about oneself, it is not to track any sort of “real fact.” It is, rather, a form of expression, of *Ausserung*. This point figures prominently in Wittgenstein’s familiar remarks about Moore’s paradox: « P but I do not believe that P » is odd because to assert that P, and thereby to express one’s belief that P,

does not amount to describing oneself as believing that P. The problem which this view poses is that self-knowledge ceases to be a kind of *knowledge*, since it is not knowledge of anything. A similar objection can be addressed to the externalist view. If thought-content is determined by the external environment, the subject, in order to know his own mind, has to look at the external environment. But, as Boghossian (1989) says, this threatens to destroy the very idea of *self-knowledge*: “Self-knowledge disappears as a cognitive achievement.” The Kantian strand itself is not unproblematic. When I ask myself: “Do I believe that P” and set out to answer this question by wondering whether P is the case, I certainly do not attend explicitly to *my believing that P*, but in so far as the first question is supposed to yield *self-knowledge* about my own states, it seems hard to say that this knowledge reduces to a knowledge about the world. So the Kantian strand threatens to fail to recognize self-knowledge as a cognitive achievement. Most writers about self-knowledge, however, agree that it is neither a perceptual capacity nor the result of an inference. So how can we interpret the transparency feature in order to illuminate self-knowledge?

3. Transparency and reflective control

A recent trend of interpretation of transparency emphasizes the deliberative and the practical meaning of this feature (Moran 2001). On this reading, the claim of transparency requires the deferral of the theoretical question:

“What do I believe?”

to the deliberative question

“What am I to believe?”

On this view self-knowledge rests upon rational agency. It's because we commit ourselves, through some sort of decision, to a given belief, that the belief is our own, and it through this kind of commitment that we come to know our own beliefs. Moran agrees with Wright and others that self-knowledge is not a theoretical achievement (not a perceptual capacity or evidence gathered from our observation of our own behavior, not an eye directed inwards). Otherwise we cannot explain how it can be immediate and imply first-person authority. If self-knowledge were the gathering of evidence about oneself, there would be a decoupling of the subject's commitment to the truth of the fact that he believes that P from the truth of his judgment that P. His evidence via an internal glance may give him evidence for the truth of “I believe that P” but it does not give him evidence to the same degree for the truth of P. The disjoint nature of the sets of reasons in favor of “I believe that P” and “P” leaves the possibility for the subject that his judgment that he believes that P and his judgment that P could diverge. But there cannot be any divergence between “I believe that P” and “P”. By definition my judgment about whether P *determines* my self-ascription of a belief that P. To understand beliefs as states which “aim at truth”, formed and maintained by agents of the basis of their reasons, one must meet the transparency test. A subject's *making up his mind* as to whether P gives him knowledge that he believes that P, if he does:

“A person is credited with first-person authority when we take the question of what he does believe to be settled by his decision as to what he is to believe.” (Moran 2001: 134)

In a nutshell, Moran's answer to the question: How do we know our own beliefs? is the following:

- (i) S knows, in virtue of understanding his own concepts, that his beliefs about P are what he holds true with respect to P.
- (ii) S knows by considering what is true of P, that he holds P to be true.
- (iii) so S knows that he believes that P.

But this leaves unanswered the question: how do we know what is reported by (ii)?

Moran's answer is that "One must see one's deliberation as the expression and development of one's belief and will." (2001: 94).

This, however, is quite problematic. Does that imply that to self-ascribe the belief that P is *deciding to believe that P*? This conception of self-knowledge seems to make belief an active attitude, and to imply a form of doxastic voluntarism. But can one control one's belief? There are familiar reasons for which it is not the case, given for instance by Bernard Williams:

"It is not a contingent fact that I cannot bring it about just like that that I believe something, just as it is a contingent fact that I cannot bring about, just like that, that I am blushing. Why is this? One reason is connected with the characteristic of beliefs that they aim at truth. If I could acquire a belief at will, I could acquire it whether it was true or not; moreover I would know that I could acquire it whether it was true or not. If in full consciousness I could acquire a belief irrespective of its truth, it is unclear that before the event I could seriously think of it as a belief, i.e. as something purporting to represent reality." (Williams 1973: 148)

If Moran meant that having the kind of self-knowledge about one's own mind which is expressed in the transparency feature were a form of voluntary control over one's beliefs, he would immediately fall into the difficulties of such a view. But that is not, apparently, what he means. He means rather that we have an immediate entitlement to our own beliefs which implies a form of commitment. But in what sense does this commitment *constitute* our self-knowledge? According to Moran, reflection implies this sort of immediate control and *constitutes* it. But it is still unclear in what sense self-knowledge could be a practical achievement, and imply a form of control on our beliefs which would imply a form of *agentive* responsibility. Moran seems to mix, or at least to bring close to each other, the question of control of action and the question of control of belief. But these are quite distinct matters.

4. Two kinds of belief control

We need to distinguish different ways in which one can have "control" of one's beliefs. Beliefs, like intentions, answer to *reasons*: there are reasons to believe, and reasons to intend. But the structure of reasons in each case is quite distinct.

Intrinsically or constitutively, reasons to believe are epistemic reasons: they consist in the existence of evidence for the truth of a belief. For a belief to be correct, and for having reasons to have it, one must attend to some sort of evidence. This means that to believe that P for reasons which fall short of being evidence for the truth of a belief, is not to believe that P with good reasons (I am not concerned here in trying to specify more precisely the nature of this correctness condition, see Millar 2004, Engel 2005). Now, intrinsically or constitutively, reasons to intend to (where " " denotes an action) are practical, not epistemic reasons. They are reasons to the effect that something is good, or in some sense desirable.

Now besides the intrinsic or constitutive reasons that one can have to believe or to intend there can be *extrinsic* reasons to believe or to intend. Thus one can have reasons to *want* to believe something (for instance because the belief would be comforting) and reasons to *want* to intend something (for instance because one feels that one's intentions to are not strong enough to secure one's -ing). For instance I may want to believe that I am not too late to catch the 5 p.m. train, in spite of my having good reasons to believe that I am actually too late. Or I may want to intend to buy an opera ticket, in spite of the fact that I know that this would involve queuing for several hours, which I don't like.

The fact that there are, both for the case of action and for the case of belief, intrinsic or constitutive reasons should not occult the fact that there is a strong *asymmetry* between these reasons. For a belief it is *essential* that it be true and justified by appropriate evidence, otherwise it cannot qualify as a belief. If I ask you: "Why do you believe this?" and you answer: "For no particular reason," there is something odd. But if I ask you the same question for an action: "Why do you do this?" and if your answer is the same as in the belief case, it can be acceptable. Why? Because belief is "serious" or subject to normative requirements in a way in which action is not.

In the light of this distinction between intrinsic and extrinsic reasons for believing and intending, one can distinguish two sorts of control over one's beliefs, or over one's intentions (I take this distinction from Hieronymi 2006):

- a) *evaluative* control: control under intrinsic theoretical or practical reason
- b) *manipulative* control: control under extrinsic reasons

One can certainly want to believe (or not to believe) that P for extrinsic reasons (say, P is pleasurable, or in the case where it is a belief we want to avoid, painful) and succeed in believing (not believing) that P through manipulative control (for instance by taking drugs, or undergoing hypnosis or brainwashing). In other words *bring it about* that one believes that P, but it's not *believing* that P, for in this case the reason is necessarily extrinsic. Similarly one can *bring about* that one intends to, but it is not *intending* to. Evaluative control is no control at all, for it does not involve the agent *making it the case* that he or she believes that P. When one believes for good epistemic reasons, one just takes it true that P, for these reasons. That does not involve in any sense being the cause of one's believing that P. One can, however, by becoming conscious of the fact that one believes that P, or by grasping reflectively the content that P, that is by *believing that one believes that P*, have *some sort of evaluation of one's belief* which may be called, in this sense, "control." Moving from first-order beliefs (which are not necessarily conscious) to reflexive beliefs gives us an awareness of what we believe through which they can be "up to us," but this very awareness does not in any way create the belief in us. There may be cases where becoming conscious of a belief and realizing that it is something that one believes involves a form of commitment, but even in such cases, committing oneself to the belief does not mean that it is willed by us. I may, for instance, realize that I believe that Ségolène Royal is incompetent, but it is not a belief which is up to me in the sense that I desired to believe this, and succeeded in installing it into me. No: I had the belief that Ségolène Royal is incompetent, but I did not fully think about it, or did not consider the matter. When it comes to my mind, and when I realize that the belief is one that I have, I have control over it. So if there is any "control" over our beliefs it is not manipulative control, but *evaluative* control. Moran and other writers, like Hieronymi, who defend this kind of account, want to say that we can have an evaluative control over our beliefs, which comes from reflection, and over our awareness of the kind of commitment to truth that genuine believing implies.

One can extend Moran's account and accept the idea that in the sense of evaluative control we can be *responsible* for our beliefs or responsible believers. A responsible believer is not responsible in the

sense in which an agent is responsible for his acts, and liable to be praised or blamed for them. A responsible believer is rather a subject who is aware, in the reflective sense, of the ordinary normative requirements of beliefs, and of the fact that beliefs are responsible to certain kinds of intrinsic reasons – evidential reasons. Philip Pettit and Michael Smith characterize the relevant sense of responsibility for belief when they say:

“Responsible believers and desirers are orthonomous subjects, in the sense that they recognize certain yardsticks of right belief and right desire and can respond to the demands of the right in their own case.” (Pettit and Smith 1996)

So one can formulate a Principle of responsibility for beliefs:

(PR) *An agent is a responsible believer, when*

- i) he is rational or responsive to reason and to the norms of belief*
- ii) he is responsive to reason when he is able to reflect on his own reasons for believing*

The locus of responsibility is in second-order judgments. An agent is reflectively responsible of his beliefs if he satisfies the following condition of rationality:

If an agent believes that he has (most) reason to believe that P then he rationally should believe that P

If rationality can be imparted to responsible believers in this sense, breakdown of rationality should also be so imparted. But is it the case that one can be irrational in one’s beliefs in the same sense as one can be irrational in one’s actions. Can there be, for instance, *theoretical akrasia* in the same sense as there is practical *akrasia*? It is instructive to raise quickly this question, for it confirms the fundamental asymmetry of reasons for belief and reasons for action noted above. In practical *akrasia* the agent feels a conflict between his first-order desires and his second-order desires (the agent acts against his own better judgment). Now theoretical *akrasia*, or epistemic *akrasia*, would be a situation where the agent believes against what he takes to be rational to believe. But that seems to be quite hard to conceive, for it would amount to an open contradiction within the subject *at the conscious level*, and in this sense it would be quite different from cases like self-deception

As Pettit and Smith (*ibidem*) point out, the fundamental asymmetry between thought and action is further revealed by the following feature: Failures to exercise free will are matters of everyday experience, but failures to exercise free thought are elusive. In the practical sphere, we can say: “I am willing to but I ought not to.” My reasons not to do not lose their probative force because I have decided to. But in theoretical sphere, there is something paradoxical in “I believe that P, but I ought not to.” I am necessarily conscious of the conflict in the action case, whereas I am, most of the time, not conscious of the conflict in the second case.

5. Weak and strong responsibility for belief

Let us take stock. I have described two versions of the neo-Kantian account of freedom of belief:

- (i) *Weak version of PR*: reflection is a *precondition* for the rational control of

beliefs,

(i) *Strong version of PR*: reflection is *constitutive* of rational control.

Moran, McDowell and others take the strong version. They hold that not only rational control is constituted by the power of reflection, but also that there is a form of active element in the self which puts on a par responsibility for belief and responsibility for actions and autonomous agency. Burge too seems to endorse this strong version when he says:

“One must be susceptible to the force and implement normative evaluations in guiding thought and other acts that fall under such those evaluations and (to understand reasoning), one must regard reasons as effective in one’s judgments. Doing so amounts to acknowledgement of one agency. ... in recognizing the effect of reasons on one’s judgments and inference, one cannot think of oneself as powerless.” (Burge 1995, op cit)

But this strong version is wrong. We have first-order beliefs, which are rational or not. We do not have any control over them. They are just forced upon us. Neither do we have control over the norms of rationality. We do not become rational just by being aware of our beliefs and of the norms which govern them. Certainly, if a conflict arises between our first-order beliefs and what we believe that we ought to believe, then we should try to change our beliefs. But we do not do that because we are aware of these norms and of our beliefs. One does this because one has to comply with these norms, and is forced to do that. This essential passivity of belief and the asymmetry between the two kinds of reasons – for belief and for action – forbid any sort of assimilation of autonomy of thought and autonomy of agency.

The weak version of PR, however, seems to be much more plausible. It says that awareness of our own beliefs and of the norms of reasons is a precondition of rational control, and that *if* we have this awareness, then we can be in a position to accept certain beliefs and reject others. But it does not say that control is effectuated *by* this very awareness. It says that it is a necessary, but not a sufficient condition. Certainly to be able to maintain a certain belief, to stick to it, or to reject it, I must be conscious of it. And to assess it, it seems that I must be aware of the norms which govern it

It is quite compatible with the weak version to deny that there is no active control at all on thought. The role of action in thought is at best indirect. It is prefatory, but not constitutive of thought. James already remarked in his *Principles of Psychology* that attention creates no idea. There is no reason to withdraw the traditional Humean empiricist motto that “Belief belongs to the sensitive, rather than to the cogitative part of our natures.” We can further doubt that reflection is a necessary condition for thought. The question to ask here is whether the epistemology of self-knowledge is internalist or externalist, this time not in the sense of the dependence of thought contents upon the external world, but of the dependence of thought contents upon our capacity to reflectively consider it. If there is knowledge of own minds, is it a knowledge of our knowledge? In other words, in order to know the content of our beliefs, do we really need to know that we know these? Is self-knowledge subject to the principle that to know one needs to know that one knows? It seems to be the paradigmatic reflective and internalist capacity. But if externalism about knowledge is right, and if we can know something without being in condition to know that we know (Williamson 2000), an externalist account of self-knowledge might well be correct. This is not, however, something for which I can argue here, and I must leave it as only a suggestion.

6. Conclusion

In opposition to the neo-Kantian account of the autonomy of thought, I have held that self-knowledge is not based on a reflexive capacity which would confer to the thinking subject a form of constitutive control over his own beliefs. Belief formation is not a matter of action, even though there can be manipulative or indirect control over one's beliefs. Neither is belief evaluation, which is a different capacity, is not a matter of taking responsibility of one's thoughts. We can be reflective agents, and are responsible to the norms and reasons of beliefs, but that does not make us free believers in any sort of deliberative sense. If ascribing beliefs to oneself is a form of knowledge, self-knowledge might well be a cognitive and passive, not a practical and active feature of our mental lives

REFERENCES

- Barnes, J. "Belief is up to us." *Proceedings of the Aristotelian Society*.
Volume 106, Number 2, March 2006, pp. 187-204 (18).
- Boghossian P. "Content and self knowledge." *Philosophical Topics* . 1989: 17: 5-26.
- Brandom R. *Making it Explicit*. Harvard University Press, 1994.
- Burge T. "Reason and the First Person." in Mc Donald, C. And Wright, C.
Knowing our own minds. Oxford University Press, 1995.
- Engel, P. *Believing and Accepting*. Dordrech, Kluwer. 2000.
"Free Believers." *Manuscripto*. Vol. XXV-Special Number-2002,
Mental Causation. André Leclerc, Giovanni da Silva de Queiroz et
Michael B. Wrigley (eds), pp. 155-175.
- "Truth and the Aim of Belief." D. Gillies, ed., *Laws and Models in
science*. London: King's College, 2005, pp. 77-97.
- Evans, G. *The Varieties of Reference*. Oxford University Press, 1982.
- Frankfurt, H. "Freedom of the Will and the Concept of a Person." *The Journal
Of Philosophy*. Vol. 68, No. 1, 5-20. Jan. 14, 1971.
- Goldstein, L. "Moore's paradox." in Engel 2000.
- Hieronymi P. "Controlling attitudes." *Pacific Philosophical Quarterly*.
Volume 87, Number 1, March 2006, pp. 45-74(30).
- Korsgaard, C. *The sources of normativity*. Cambridge University Press, 1996.
- McDowell, J. 1998 "Having the World in View: Sellars, Kant, and Intentionality."
Journal of Philosophy. 1998. XCV/9 : 431-491.
- Millar, A. *Understanding People*. Oxford University Press, 2004.
- Moran, R. *Authority and Estrangement*. Princeton University Press, 2001.
- Owens, D. "Epistemic Akrasia." *Monist*. 2002, Vol. 85, No. 3
- Pettit, P. and Smith M. "Freedom in Belief and Desire." *Journal of
Philosophy*. 1996. XCIII, 9, pp. 429-449.
- Shoemaker, S. *The first person perspective and other essays*.
Cambridge University Press, 1996.
- Williams B. "Deciding to believe." in *Problems of the Self*
Cambridge University Press, 1970.
- Wittgenstein , L. *Philosophical Investigations*. Blackwell, Oxford, 1952.
- Wright, C. "Self- knowledge, the Wittgensteinian legacy." in Smith, Mc
Donald, & Wright, eds. *Knowing one's own mind*, Oxford, Oxford University Press 1995,
pp.13-46.