

Free belief

JOSEFA TORIBIO

Department of Philosophy, Indiana University, Bloomington, IN 47405, USA
(E-mail: jtoribio@indiana.edu)

Received 25 April 2002; received in revised version 4 March 2003

Abstract. The main goal of this paper is to show that Pettit and Smith's (1996) argument concerning the nature of free belief is importantly incomplete. I accept Pettit and Smith's emphasis upon normative constraints governing responsible believing and desiring, and their claim that the responsibly believing agent needs to possess an ability to believe (or desire) otherwise when believing (desiring) wrongly. But I argue that their characterization of these constraints does not do justice to one crucial factor, namely, the presence of an unreflective, sub-personally constituted, ability to spot the kind of situations in which the reflective critical abilities constitutive of responsible believing (and desiring) should be deployed.

Key words: belief, critical pop-out, evidential norms, recognitional capacity, responsibility

In "Freedom in Belief and Desire," Philip Pettit and Michael Smith (1996) advance a novel and interesting thesis about what it is to believe and desire freely. "To hold a belief or desire freely is to hold it in the presence of an ability, should the belief or desire be wrong, to get it right" (Pettit and Smith 1996, p. 445).¹ Only *responsible* believers and desirers possess this ability. They are what Pettit and Smith call orthonomous agents (see their 1990, 1993). An orthonomous agent is someone who is able to recognize the existence of normative constraints governing her beliefs and desires and, importantly, someone capable of responding appropriately to those constraints, by correcting mistakes when they arise. Pettit and Smith thus endorse – despite philosophical arguments to the contrary² – a certain version of the possibility of believing (and desiring) at will. Furthermore, they support this possibility by arguing that this ability is what constitutes free will.

The main goal of this paper is to show that Pettit and Smith's argument is importantly incomplete. I accept Pettit and Smith's emphasis upon normative constraints governing responsible believing and desiring. But I shall argue that

Castlefield Press
PHEN
PIPS No. 5151930
ART. 141
CP, TYPESET, DISC

their characterization of these constraints does not do justice to one crucial factor, namely, the unreflective, sub-personally constituted, ability to spot the kind of situations in which the reflective critical abilities constitutive of responsible believing (and desiring) should be deployed. The spirit of my paper is thus sympathetic, but aims at an important elaboration of the idea of believing responsibly. To this end, I develop a notion that I'd like to call "critical pop-out". The idea, in a nutshell, is that in order to believe responsibly, the agent should be *skilled at automatically detecting* the kind of situations in which the critical abilities crucial to responsible believing should be deployed. The process of *spotting* these situations must (on pain of regress) be automatic. The critically sensitive situations must simply "pop-out" to the doxastically responsible subject. To see how this proposal differs from, but ultimately complements, Pettit and Smith's account, let me first present some of their own considerations regarding free thought.

Evidential norms

Responsible believers are to be located within what Pettit and Smith call *the conversational stance*. It is within this stance, i.e. within the realm of intellectually oriented conversations, that people make certain assumptions about themselves and about one another. Central among these assumptions are the ideas that (a) the beliefs we form bear on common questions; (b) everyone has a certain authority regarding her own beliefs, and (c) when disagreement occurs, it would usually disappear as a consequence of reviewing the supporting evidence. Finally, people also assume that everyone who engages in conversation shares these ideas. The resulting picture has a certain Gricean flavor, in the sense that for rational conversation to work in the way it does, everyone has to believe that everyone else also believes these assumptions to be operative (Pettit and Smith 1996, pp. 430–431).

In order to adopt the conversational stance, an additional and important requirement is needed. The orthonomous, i.e., the responsibly believing agent also needs to possess a *self-starting* version of the aforementioned ability to believe (or desire) otherwise when believing (desiring) wrongly. That is to say, the agent should be able not only appropriately to respond to the normative requirement of e.g., modifying his belief/desire vis-à-vis the evidence or arguments presented by *others*, but should also be able to respond appropriately to evidence/arguments "he will often be in a position to produce himself" (Pettit and Smith 1996, p. 447). The adoption of the conversational stance is thus not restricted to the standard idea of exchange of reasons in conversation

among different people. *Intrapersonal* conversation, i.e., the kind of reflection a person alone can engage in when assessing matters or trying to make decisions also belongs, and importantly so, to the conversational stance. The ability to correct errors as a result of self-presented considerations is thus an important part of what makes an agent a responsible believer, i.e., an *orthonomous* agent.

Pettit and Smith are very clear about the conditions that have to be met for an agent to be considered a proper conversational interlocutor (either with herself or with other agents). The conditions are threefold: “there are certain norms governing what that subject ought to believe, the subject is disposed to recognize those norms, and she is disposed to respond in the way required” (Pettit and Smith 1996, p. 436). The norms the authors have in mind when stating the first of these conditions are related to the kind of evidence that justifies holding particular beliefs. Given how the world is, it is right to believe that e.g., “Paris is the capital of France”, wrong to believe otherwise. If I believe that if p then q , and I also believe that p , then, *ceteris paribus*, I should believe that q , etc. Pettit and Smith call these norms *evidential* norms (although they are thinking not only of evidence provided by environmental conditions, but also evidence that accrues as the result of certain inferential steps).

The second condition, namely, that the subject be disposed to recognize those evidential norms, is probably the most important for present purposes. To recognize such norms is, according to Pettit and Smith, for the subject to have beliefs “with contents of the form: it is true that p or it is false that p ; the evidence supports the hypothesis that p or is against the hypothesis that p . . . To believe that a certain proposition . . . is true or is supported by the evidence or is entailed by something that is itself accepted is to believe, in effect, that it is right to believe the proposition, wrong to disbelieve it” (Pettit and Smith 1996, p. 434). Finally, the responsible believer should be able to act upon such a recognition of norms, i.e., she should “be disposed to maintain beliefs that comply with the norms and to reject beliefs that fail to do so” (p. 435). When these three conditions are met, Pettit and Smith claim, we can say of the agent that she enjoys free thought, because she possesses the ability to get her beliefs right, should they be wrong.

In thus discussing the possibility of believing responsibly (and hence freely), Pettit and Smith neatly sidestep the traditional question of whether or not an agent has control over possible alternatives concerning what to believe – an issue which has played a central role in most discussions of believing at will.³ The right question, it is argued, cannot be whether the agent has free choice over what to believe, since if the agent has done everything right, there is no value whatsoever in having the choice of being wrong! We do not believe re-

sponsibly because we freely choose to believe p instead of q . Rather, we are said to freely choose to believe p *because* we meet *other conditions* for responsible believing, namely, the conditions stated above.

This kind of move is, of course, not new in the literature. The idea that responsibility is a *precondition* of choice or voluntary control rather than the other way round echoes some of Strawson's claims in "Freedom and Resentment" (1974), even though Strawson is there discussing moral, and not, as if were, *doxastic* responsibility. The idea is also related to Dennett's take on the issue of control and the so-called "could have done otherwise" principle in discussions of free will (Dennett 1984). The "could have done otherwise" principle can only be sensibly interpreted, Dennett claims, as the possibility of a properly functioning agent modifying her actions in the future as the result of being presently prompted to corrections by the provision of training or feedback. Someone "could have done otherwise" only if she is able to learn from the current outcome of her actions; only if she is "cognitively tuneable" so as to act differently when facing the "same" situation in the future (Dennett 1984, pp. 139–144).

Importantly, Pettit and Smith distance themselves from this Dennettian point of view by stressing that on their account, unlike Dennett's, the *reason-laden* character of the requisite tuneability is of the essence.⁴ Their idea of revision due to recognition of relevant evidential norms requires the subject to be *vigilant* at the personal level, whereas Dennett's view, it is claimed, requires only a rather passive, *mechanical* agent. Given the Dennettian model of attitude formation, the kinds of revisions and adjustments involved in learning to cope with the demands of evidence and rationalization "may happen within them [the subjects] without any recognition of why they should happen and without any efforts on their behalf to help them happen" (Pettit and Smith 1996, p. 441). Pettit and Smith thus depict Dennett's notion of a subject's being "cognitively tuneable" by training or feedback as a potentially fully sub-personal phenomenon, and reject the relevance of any "adjustments" at that level for (personal level) questions about responsibly believing.

We can call Pettit and Smith's criticism of such an image of cognitive change *the sub-personal worry*. The worry seems to stem from a deliberately intellectualist view of norm compliance and recognition. On this view, it is insufficient that the subject be responsive to what the relevant norms require. Rather the recognition of such normative demands must itself take place within the conversational stance, i.e., by the subject engaging in conversation with herself or with someone else (Pettit and Smith 1996, p. 442). The label "mechanical" applied to Dennett's account can be unpacked as characterizing a lack of personal level awareness on the part of the subject. The subject is not aware

of being governed or having to be governed by the rational demands of evidential norms, and without such cognitive vigilance, there is no room for responsible believing. It is thus not difficult to see why Dennett's view is presented as inadequate. Dennett's error is not to pay sufficient attention to the crucial personal level components involved in the requisite norm-sensitivity. Instead, Pettit and Smith end up presenting what might be called a *self-contained* model of the normative dimensions. It is self-contained because the model does not posit any *crucial* role for sub-personal states in the explanation of attitude formation. Their characterization of how to meet the normative demands is conceptually independent of any sub-personal level contributions.

In opting for such a *self-contained* model, Pettit and Smith succumb to a kind of mirror-image of Dennett's error. Where Dennett paid insufficient attention to personal level elements, Pettit and Smith, I shall now argue, pay insufficient attention to the equally crucial role of certain sub-personal components to responsible believing. In the rest of the paper I shall argue that unless we acknowledge a crucial sub-personal component allowing an agent automatically to recognize those situations in which she should actively deploy her personal level abilities to respect evidential norms, we haven't provided a viable characterization of responsible believing. An account of responsible belief is possible only if we opt for a *non self-contained* model of how we meet the normative demands.

I will elaborate these ideas in two stages. First, I shall analyze in more detail Pettit and Smith's own account of the recognitional capacities in virtue of which we are sensitive to certain evidential norms. Second, I shall introduce and characterize a specific sub-personal component as an additional, and necessary, condition for responsible believing.

Recognitional capacities

As I pointed out earlier, Pettit and Smith's account of responsible belief involves the adoption of a conversational stance, which, in turn, requires three conditions to be met: there need to be belief-relevant norms, the agent must be capable of recognizing this to be the case, and, finally, the agent must be able to act upon this recognition.

I want now to concentrate on the second of these three conditions, namely, the agent's capacity to recognize evidential norms. In Pettit and Smith's account, this recognitional capacity has a dispositional character. There are, of course, situations in which certain circumstances would prevent the agent from acting in accordance with such dispositions – inattentiveness, forgetfulness,

passion, etc. If these circumstances completely block the subject's abilities to reason, if they become – to use Pettit and Smith's terminology – “disabling obstacles,” then we may withdraw (what they term) our “authorization” of the subject as a conversational interlocutor. Otherwise, if the obstacles are not thus disabling, then there remains the possibility of bringing the agent to recognize that her beliefs do not meet relevant norms: ones that she, in general, is disposed to acknowledge. This is done through the exchange of arguments:

When you authorize someone in conversation, you hold her to the expectation that she will balk at discrepancies between the two of you, and do so in a way that invites the ascription of beliefs like the following: that you each have different belief attitudes toward the same content, that the evidence available may rule out one of those beliefs as unsupported or false . . . If an interlocutor failed to live up to that expectation, if she failed to manifest any notion of there being a common content of belief or a common fund of evidence, for example, then you would have no reason to take her attitudes seriously. (Pettit and Smith 1996, p. 434)

Since conversation is also *intrapersonal*, the strategy should work along the same lines in those cases in which the agent has to deliberate by herself, i.e., without being prompted by any other agent's presentation of reasons or arguments. How are we to understand this capacity of an agent to recognize normative standards when deliberating on her own? Is this kind of capacity based solely on the agent's disposition to make, albeit imperfectly, the kind of inferences prescribed by Pettit and Smith's notion of evidential norms?

It is certainly the case that, especially in processes of decision making, we find ourselves – as isolated individuals – engaging in the sort of critical process that implies having already recognized a particular situation as one requiring the exercise of critical appraisal. However, this realization cannot be fully characterized in the way Pettit and Smith describe, i.e., as the subject's simply manifesting “beliefs involving notions like truth and support and entailment” (Pettit and Smith 1996, p. 434). To do that is to assume that the subject already has the cognitive ability to appreciate certain situations as requiring the application of a particular set of inferential principles. It is to assume that she has the ability to spot certain situations as governed by relevant evidential norms. In effect, Pettit and Smith just assume that the subject is able to *spot* when critical engagement is appropriate. The problem is less obvious when the conversational stance is understood *interpersonally*, since one role of the outside interlocutor is precisely to engage the subject in critical reflection. But when the individual is her own interlocutor, she needs to be able to *spot* those situations in which she ought to engage in critical reflection (similarly, in the interpersonal case, someone has to determine that critical reflec-

tion and engagement is appropriate). Since we cannot be constantly questioning our thoughts and beliefs, and since *spotting* when we should do so cannot, on pain of infinite regress, be a matter of actively entertaining or understanding additional inferential principles, we are forced to invoke some kind of automatic sub-personal component. Thus, Pettit and Smith's *self-contained* model of cognitive vigilance must be rejected.

Critical pop-out

I have argued that to believe responsibly we must not only meet the three conditions displayed by Pettit and Smith, but must also possess a specific kind of know-how or skill, namely, the ability to *automatically* recognize a situation as one requiring critical appraisal. This skill cannot itself – unlike the intellectual processes it triggers – be reason-based, on pain of infinite regress. Instead, the situations that require an exercise of the agent's critical obligations must simply *pop-out* to the agent in some non-reasoned, conceptually unmediated way. I shall call that necessary – but wholly sub-personal – component of believing responsibly “critical pop-out”. As befits a kind of know-how, critical pop-out operates in an automatic fashion and belongs squarely to the sub-personal level. However, I claim, the very notion of fulfilling one's intellectual obligations is unintelligible unless certain situations simply pop-out to the agent as situations requiring critical reflection. This notion of pop-out should thus be viewed as the *rule-stopper* that triggers rule-following activity conforming to Pettit and Smith's notion of “an agent's recognizing the demands of evidential norms.”⁵

It might perhaps be objected that a kind of superbeing, with unlimited mental resources and unlimited time, could be a responsible believer independently of possessing any capacity of critical pop-out. I am not convinced that this is so (wouldn't such a being just persist in an endless spiral of unending appraisal, meta-appraisal, and so on?) But more importantly, the interesting philosophical project surely is to understand how agents broadly speaking *like us* (i.e., time and resource limited material beings) can count as responsible believers. And this project, I want to say, requires us to recognize the cognitively crucial role of the kind of sub-personal mechanism I have described.⁶ The conduct essential to responsible believing thus occurs, as Pettit and Smith rightly insist, at the *personal* level, but it needs to be linked to the kind of sub-personal ability that I've called “critical pop-out”.

To take a mundane case, consider people who claim they *just don't see* (and hence don't form the belief) that the trash bag is full, and consequently (too)

seldom come to believe that the trash needs to be taken out. It is perhaps possible to imagine a training program for these subjects (some of whom I happen to know) such that they might learn better to detect full trash bags in their local environment. In one sense, clearly, the subjects were perfectly able to *see* full trash bags before the training. What the training does is to change the functional poise of certain inputs by engendering a new skill that allows those situations to automatically and unreflectively strike the subject as ones demanding further attention.⁷

The notion of critical pop-out is thus not to be assimilated to the much simpler notion of (if you like) “novelty pop-out.” Novel situations do, to be sure, tend to engage our critical and reflective faculties. But the kind of skill I am envisioning is quite naturally displayed in situations with which we are quite familiar. Consider, for example, the experienced air traffic controller who immediately (and just in time) recognizes that her command to drop 300 feet and turn 30° west would cause a head-on collision, or the logician who spots the elementary error in her own proof. Of course, we do not always succeed in identifying such errors. But insofar as we are *culpable* when we fail to do so, the preconditions of this culpability include both our possession of the personal level skills required to meet Pettit and Smith’s three conditions *and* of an automatic, sub-personally constituted, capacity to know when such critical engagement is appropriate. It is only courtesy of this dual constitution, I suggest, that the buck of responsibility can ever come to rest.

In sum, the normative demands governing the personal level phenomenon of responsible believing can intelligibly be met only by (finite) beings possessing the automatic, sub-personal skill I have called “critical pop-out”. The key role played by conversation at the interpersonal level may seduce us into thinking that the possession of such a skill is merely an *enabling* condition, and one that should not be considered crucial to an account of responsible believing. But this argument loses its force when we look more closely at the *intrapersonal* aspect of the conversational stance. Our ability to correct errors as a result of self-presented considerations depends crucially upon the additional sub-personal skill of unreflectively spotting those situations which require our critical engagement.

There is, finally, a more general moral to be drawn from all this. For Pettit and Smith’s failure to give due philosophical weight to sub-personal activity is clearly part of a larger pattern. There is a strong tendency in the recent philosophical literature to view sub-personal factors as merely the *enabling* conditions that make certain personal level phenomena possible, but which play no philosophically interesting role in the accounts then developed.⁸ This kind of personal/sub-personal apartheid is supposed to turn on considerations

of rationality. Pettit and Smith's *sub-personal worry* is a case in point. It is thus an interesting result of the present analysis that at least one of the conceptually crucial conditions for responsible believing turns out to belong to the sub-personal level. A careful reconsideration of the relations between personal and sub-personal elements in philosophical account of mental phenomena is, I conclude, seriously overdue.

Notes

1. The idea of a link between freedom and/or rationality and the ability to correct errors is not itself new. Versions can be found in e.g., Dennett (1984), Frankfurt (1971), Haugeland (1998), and Strawson (1974).
2. See e.g., Alston (1986), Holton (1994), Hume (1958), James (1948), Montmarquet (1986), Scott-Kakures (1993), Williams (1973). Winters (1979) argues quite convincingly for the inconclusiveness of Williams' arguments.
3. See e.g., Alston (1988, p. 261): "...one has control over a given type of state only if one also has control over some field of incompatible alternatives. To have control over believing that p is to have control over whether one believes that p or not."
4. Pettit and Smith do not mention Dennett's 1984. Their reference is rather to *The Intentional Stance* (Dennett 1987).
5. As an anonymous referee (whom I thank kindly) reminds me, the discussion here seems to be related to the issue of *phronesis* in Aristotle. Aristotle discusses this special type of practical intelligence in vi 5, 1141 b 8–1142 a 30 of the *Nicomachean Ethics*. Practical intelligence or practical wisdom, *phronesis* is characterized as a perceptual ability to recognize, in particular cases, what is required of the subject for her to achieve virtue: "It calls for something like perception to recognize at what point one's action would become blameworthy" (Sorabji 1980, p. 206). To pursue this connection is, of course, beyond the limits of this paper. However, it should be noticed that a significant number of cognitive scientists have found this aspect of Aristotle's moral philosophy to be enormously fruitful for a defense of a more general, non-sentential and know-how oriented, model of knowledge (see e.g., Churchland 1996; Clark 2000; Flanagan 1996).
6. For some related ideas concerning the proper objects of epistemological analysis, see Cherniak's (1986) treatment of "minimal rationality".
7. It would be interesting to see how the claim that this subpersonal capacity can be *tuned* via learning relates to different species of nativist and non-nativist views in the developmental literature. It may, of course, be the case that the learning mechanism required to accomplish this tuning is innate, at least in the sense of not being the outcome of another learning mechanism previously in place. It would thus turned out to be true that some subjects are just *better* at this forward looking tuneability than others. However, nothing in the notion of critical pop-out defended here commits us to a particular view on this matter.
8. See e.g., Davidson (1980), Dennett (1969), Pylyshyn (1984), Taylor (1964), and especially McDowell (1994). Bermúdez (2000) is an excellent, critical, discussion of this tendency to marginalize the sub-personal, whose origins are traced to the works of Wittgenstein and Ryle.

References

- Alston, W. P. 1988. The deontological conception of epistemic justification. In: J. E. Tomberlin (ed), *Philosophical Perspectives*, Vol. 2, pp. 257–299. Atascadero, CA: Ridgeview Press.
- Bennett, J. 1990. Why is belief involuntary? *Analysis* 50: 87–107.
- Bermúdez, J. L. 2000. Personal and subpersonal: A difference without a distinction, *Philosophical Explorations* 3 (1): 63–82.
- Cherniak, C. 1986. *Minimal Rationality*. Cambridge, MA: MIT Press.
- Churchland, P. M. 1996. The neural representation of the social world. In: L. May, M. Friedman, and A. Clark (eds), *Mind and Morals. Essays on Ethics and Cognitive Science*, pp. 91–108. Cambridge, MA: MIT Press.
- Clark, A. 2000. Word and action: Reconciling rules and know-how in moral cognition. In: R. Campbell and B. Hunter (eds), *Moral Epistemology Naturalized. Canadian Journal of Philosophy* Suppl., Vol. 26, pp. 267–290. Alberta: University of Calgary Press.
- Davidson, D. 1980. *Essays on Action and Events*. Oxford: Oxford University Press.
- Dennett, D. 1969. *Content and Consciousness*. London: Routledge, Kegan and Paul.
- Dennett, D. 1984. *Elbow Room. The Varieties of Free Will Worth Wanting*. Cambridge, MA: MIT Press.
- Dennett, D. 1987. *The Intentional Stance*. Cambridge, MA: MIT Press.
- Flanagan, O. 1996. Ethics naturalized: Ethics as human ecology. In: L. May, M. Friedman, and A. Clark (eds), *Mind and Morals. Essays on Ethics and Cognitive Science*, pp. 19–44. Cambridge, MA: MIT Press.
- Frankfurt, H. 1971. Freedom of the will and the concept of a person. *The Journal of Philosophy* 68 (1): 5–20.
- Haugeland, J. (ed) 1998. Truth and rule-following. In: *Having Thought. Essays in the Metaphysics of Mind*, pp. 305–361. Cambridge, MA: Harvard University Press.
- Hume, D. 1958. In: L. A. Selby-Bigge (ed), *A Treatise of Human Nature* (Orig. 1739). Oxford: Oxford University Press.
- James, W. 1948. The will to believe. In: *Essays in Pragmatism*, pp. 88–109. New York: Hafner Press, London: Collier MacMillan.
- McDowell, J. 1994. *Mind and World*. Cambridge, MA: Harvard University Press.
- Montmarquet, J. 1986. The voluntariness of belief. *Analysis* 46: 49–53.
- Pettit, P. and Smith, M. 1990. Backgrounding desire. *The Philosophical Review* 99 (4): 565–592.
- Pettit, P. and Smith, M. 1993. Practical unreason. *Mind* 102 (405): 53–79.
- Pettit, P. and Smith, M. 1996. Freedom in belief and desire. *The Journal of Philosophy* 93 (9): 429–449.
- Pylyshyn, Z. 1984. *Computation and Cognition*. Cambridge, MA: MIT Press.
- Scott-Kakures, D. 1993. On belief and the captivity of the will. *Philosophy and Phenomenological Research* 53 (4): 77–103.
- Sorabji, R. 1980. Aristotle on the role of intellect in virtue. In: A. O. Rorty (ed), *Essays on Aristotle's Ethics*, pp. 201–219. Berkeley: University of California Press.
- Strawson, P. F. 1974. *Freedom and Resentment and Other Essays*. London: Methuen.
- Taylor, C. 1964. *The Explanation of Behavior*. London: Routledge, Kegan and Paul.
- Williams, B. 1973. *Problems of the Self. Philosophical Papers 1956–1972*. Cambridge: Cambridge University Press.
- Winters, B. 1979. Believing at will. *The Journal of Philosophy* 76 (5): 243–256.