

Constitutivism and the Inescapability of Agency

Luca Ferrero

University of Wisconsin at Milwaukee

ferrero@uwm.edu

<http://www.uwm.edu/~ferrero>

to appear in *Oxford Studies in Metaethics*, vol. IV

Jan 12, 2009 -- Version 3.05

This is an uncorrected proof – subject to minor revisions

PLEASE QUOTE FROM PUBLISHED VERSION WHEN AVAILABLE

1 Introduction

1.1 The norms of rationality and morality have special authority; they are *categorically* binding. They bind agents regardless of their contingents motives, preferences, and intentions. By contrast, the norms of particular games, institutions, and practices, are only *conditionally* binding. They have normative force only for agents who have a good enough reason to participate in them. A statement that one ought to move the knight along the diagonals, for instance, expresses an ought-according-to-the-norms-of-chess. But the oughts of rationality and morality are not qualified with the clause ‘–according-to-the-norms-of-rationality/morality;’ they rather tell us, as Stephen Darwall writes, ‘what we ought to do *simpliciter*, *sans phrase*.’¹

How can we account for the categorical force of the norms of rationality and morality? Some philosophers have argued that the grounds of these unconditional oughts are to be found in the nature of agency.² In a rough outline, their basic claim is that the norms and requirements of practical rationality and morality can be derived from the constitutive features of agency. Hence, a systematic failure to be guided by these requirements amounts to a loss of agency. But there is a sense in which we *cannot but be* agents. It follows that we are necessarily bound by the oughts of rationality and morality, we are bound by them *sans phrase*.

1.2 The success of this argumentative strategy—which goes under the name of ‘constitutivism’—depends on establishing the following two claims. First, that the norms of rationality and morality can be derived from the constitutive features of agency. Second, that we cannot but be agents, that agency is non-optional.³

¹ Darwall (1992: 156).

² Constitutivist views are defended by Korsgaard (1996; 1997; 1999, 2002), Railton (1997), Millgram (1997: ch. 8), Shapiro (1999), Velleman (2000; 2004; forthcoming), and Rosati (2003).

³ A further problem with constitutivism concerns how it handles errors and imperfections in attempts at complying with the constitutive standards of agency. The worry is that constitutivism might implausibly imply that agents can only exist as perfect agent, which in turn would preclude the possibility of any

Constitutivism has been criticized on both counts. Some have argued that the constitutive features of agency offer too thin a basis for the derivation of substantive normative principles and requirements.⁴ Others have objected that agency does not have any special status vis-à-vis ordinary games and practices; that our participation in agency is *optional* in the same sense as our participation in ordinary games and practices.

1.3 In this paper, I will offer a partial defense of constitutivism. I will show that there is something special about agency that makes engagement in it significantly different from the participation in other ordinary enterprises (by which I mean games, practices, institutions, and the like). I will argue that agency is ‘inescapable’ in a way that could help explain its role in grounding unconditional oughts. My defense of constitutivism, however, is limited in scope since space restrictions prevent me from discussing the prospects of deriving substantive norms from the nature of agency.⁵

2 The Shmagency Objection

2.1 The constitutive standards of an ordinary enterprise *E* determine what the agent is to do in order to engage in it. If a subject systematically fails to abide by the standards of chess, say, she is not a chess player. The rules of chess are binding on anyone who intends to play that game. But their normative force is *optional*. An agent is not *actually* bound by them unless she has a good enough reason to play chess in the first place. Moreover, whether one has a reason to play chess is not something that can be derived from the constitutive standards of chess

genuine criticism for failures to abide by the standards of agency; see Cohen (1996: 177), Railton (1997: 309), Lavin (2004), Kolodny (2005), Fitzpatrick (2005), and Coleman (unpublished).

⁴ See Setiya (2003; 2007). Railton (1997: 299) hints at a similar worry in the attempt at deriving epistemic norms from the constitutive features of belief. A related concern (raised in conversation by Jacob Ross and Mark Schroeder) is that the constitutive features of agency might be necessary but insufficient for the derivation of specific normative principles.

alone.

If agency were like an ordinary enterprise, the same would be true of its constitutive standards. First, the standards of agency and what could be derived from them would be binding only on those subjects who have a good enough reason to be agents, to engage in the ‘enterprise of agency,’ as I will sometimes say. Second, whether one has reason to be an agent could not be derived from the constitutive standards of agency alone.

2.2 David Enoch has recently argued that agency is indeed optional like any ordinary enterprise, and that constitutivism is therefore untenable.⁶ It is impossible to ground unconditional obligations in the constitutive standards of an enterprise that is only binding if one has an independently given reason to engage in it.⁷ The normative force of the reason to be an agent, assuming that there is indeed such a reason, would elude the constitutivist account of normativity.

Enoch’s argument is based on what might be called the ‘shmagency objection.’ He asks us to imagine a subject—a ‘shmagent’—who is indifferent to the prospect of being an agent. The shmagent is unmoved by the constitutive standards of agency. For instance, in response to Korsgaard’s version of constitutivism—according to which agency is the capacity for self-constitution—the shmagent says:

Classify my bodily movements and indeed me as you like. Perhaps I cannot be classified as an agent without aiming at constituting myself. But why should I be an agent? Perhaps I can’t act without aiming at self-constitution, but why should I act? If your reasoning works, this just shows that I don’t care about agency and action. I am perfectly happy being a shmagent—a nonagent who is very similar to agents but who lacks the aim (constitutive of agency, but not shmagency) of self-constitution. I am perfectly happy performing shmactions—nonaction events

⁵ The aspiration of grounding unconditional oughts and deriving substantive normative principles are arguably the most ambitious aspirations of constitutivism but by no means the only ones; see Velleman (2004: 288-289).

⁶ Enoch (2006). See also Sharon (unpublished).

⁷ Millgram (2005) is the first one to have pointed out that constitutivism might be the target of a criticism of this kind, although he does not go as far as Enoch in objecting to the ultimate viability of constitutivism.

that are very similar to actions but that lack the aim (constitutive of actions, but not shmactions) of self-constitution.⁸

2.3 A shmagent is unmoved by the constitutive standards of agency in the same sense in which someone who is indifferent to the game of chess, let's call him a chess-shmayer, is unmoved by the standards of chess. A chess-shmayer could successfully challenge the force of the constitutive standards of chess by saying, 'I don't care about chess. I am perfectly happy being a chess-shmayer—a nonplayer who is very similar to chess-players but who lacks the aim of chess playing (say, making legal chess moves with the ultimate goal of checkmating my opponent). I am perfectly happy performing chess-shmoves—non-chess moves that are very similar to chess-moves but that lack the aim of chess playing.'

The challenge of the chess-shmayer is *external* to the game of chess. Attempts at convincing the chess-shmayer to care about chess cannot be made *within* the game of chess since he is neither moved nor bound by its rules.⁹ Likewise, a chess-player who is worried that her playing might not be justified has to get outside of the game in order to find out if it is. In the meantime, she might still continue to play chess, but figuring out whether she has reason to do so is not part of the game of chess.

Enoch's suggestion that there might be shmagents is supposed to show that the standards of agency can only be binding for those subjects who have an independently established reason to be agents, whether or not they are already participating in the agency-enterprise. This reason, if it exists, must in principle be accessible to shmagents and effective in moving them; that is, it must be both available and binding outside of the enterprise of agency.

2.4 The shmagency objection is targeted at all versions of constitutivism. Whatever standards

⁸ Enoch (2006: 179).

⁹ This is not to say that the constitutive standards of the game are *irrelevant* to its justification. The standards matter, for instance, for the individuation of the object of the justification. The point that I am

are held to be constitutive of agency, one could always imagine a shmagent who is indifferent to those standards.¹⁰ Hence, whether the objection succeeds or fails is something to be determined in abstraction from particular versions of constitutivism and their specific suggestions about the constitutive features of agency.¹¹ In this paper, I will argue for the viability of the constitutivist strategy on the face of the shmagency objection, but I will not try to defend any particular version of constitutivism. My argument appeals only to those general features of agency that are accepted by all constitutivist theories. The discussion requires nothing more than an agreement on a very general characterization of the concept of full-fledged intentional agency, on agency as the capacity to shape one's conduct in response to one's appreciation of reasons for action and to engage in the practice of giving and asking for these reasons (both about one's own conduct and that of others).

2.5 The shmagency objection is even more general in scope than it might appear at first. The objection can be extended to undermine all forms of constitutivism, even those that are not centered on agency. If Enoch is right that agency is optional, the same appears to hold of shmagency as well. The question whether there is reason to be an agent rather than a

making in the text is only that the process of justifying the playing of the game is not part of the playing itself, it is not a series of moves internal to the game.

¹⁰ Enoch (2006: 170 fn 1).

¹¹ Constitutivism is sometimes presented as being about the constitutive standards of action rather than agency. As far as the discussion of the grounds of categorical normativity is concerned, however, constitutivism is better formulated in terms of *agency* as the capacity to engage in intentional action. This is because the argument revolves around the comparison between the exercise of this capacity and the participation in ordinary enterprises. I think that versions of constitutivism originally cast in terms of action can be reformulated easily reformulated, for present purposes, in terms of agency. This is not deny that the agency/action distinction might be relevant—as argued by Setiya (2003), for instance—for the derivation of substantive norms and requirements. Particular versions of constitutivism might also be differentiated on the basis of the *kinds* of features that they hold to be constitutive of agency/action (which could be aims, motives, capacities, commitments, or principles) and on whether these features operate at the personal or subpersonal level. (For instance, Velleman puts the emphasis on aims—which up to Velleman (2004) he presented as operating at the subpersonal level; Korsgaard and Railton present constitutivism in terms of personal-level compliance with principles; Rosati talks in terms of (sub-personal?) constitutive motives and capacities of agency.) None of these differences are relevant, however, for the main topic of this paper, the discussion of the viability of the general constitutivist strategy.

shmagent is thus to be adjudicated outside of both agency and shmagency. This adjudication is a move in a distinct enterprise, one that provides a standpoint external to both agency and shmagency. Let's call it 'uberagency.'

Could constitutivism be relocated at the level of uberagency, of the more comprehensive enterprise that includes both agency and shmagency as optional sub-enterprises? The problem is that an Enoch-style objection could still be moved to this kind of constitutivism. Couldn't we always imagine the existence of shm-uberagents, subjects who are indifferent to the constitutive standards of uberagency? That is, subjects who would be bound by the standards of uberagency only if they had an independently established reason to be uberagents? The same move used to show that agency is optional can thus be used to show that uberagency is optional. Moving at an even higher level would not help because the move could be repeated ad infinitum. The possibility of this regress shows that, *pace* constitutivism, appeal to the constitutive standards of *any* enterprise (be it agency, uberagency, or what have you) could never account for any *categorical* ought.

3 *The Inescapability of Agency*

3.1 The initial appeal of the shmagency objection rests on the impression that there is a close analogy between agency and ordinary enterprises. If one can stand outside of chess and question whether there is any reason to play this game, why couldn't one stand outside of agency and wonder whether there is any reason to play the agency game? The problem with this suggestion is that the analogy does not hold. Agency is a very special enterprise. Agency is distinctively 'inescapable.' This is what sets agency apart from all other enterprises and explains why constitutivism is focused on it rather than on any other enterprise.

3.2 Agency is special under two respects. First, agency is the enterprise with the largest

jurisdiction.¹² All ordinary enterprises fall under it. To engage in any ordinary enterprise is *ipso facto* to engage in the enterprise of agency. In addition, there are instances of behavior that fall under no other enterprise but agency. First, intentional transitions in and out of particular enterprises might not count as moves within those enterprises, but they are still instances of intentional agency, of bare intentional agency, so to say. Second, agency is the locus where we adjudicate the merits and demerits of participating in any ordinary enterprise. Reasoning whether to participate in a particular enterprise is often conducted outside of that enterprise, even while one is otherwise engaged in it. Practical reflection is a manifestation of full-fledged intentional agency but it does not necessary belong to any other specific enterprise. Once again, it might be an instance of bare intentional agency. In the limiting case, agency is the only enterprise that would still keep a subject busy if she were to attempt a ‘radical re-evaluation’ of all of her engagements and at least temporarily suspend her participation in all ordinary enterprises.¹³

3.3 The second feature that makes agency stand apart from ordinary enterprises is agency’s *closure*. Agency is closed under the operation of reflective rational assessment. As the case of radical re-evaluations shows, ordinary enterprises are never fully closed under reflection. There is always the possibility of reflecting on their justification while standing outside of them. Not so for rational agency. The constitutive features of agency (no matter whether they are conceived as aims, motives, capacities, commitments, etc.) continue to operate even when the agent is assessing whether she is justified in her engagement in agency. One cannot put agency on hold while trying to determine whether agency is justified because this kind of practical reasoning is the exclusive job of intentional agency. This does not mean that agency falls outside of the

¹² For the idea of the jurisdiction of an enterprise, see Shafer-Landau (2003: 201).

¹³ On radical re-evaluation, see Taylor (1985: 40 ff.).

reach of reflection. But even reflection about agency is a manifestation of agency.¹⁴

Agency is not necessarily self-reflective but all instances of reflective assessment, including those directed at agency itself, fall under its jurisdiction; they are conducted in deference to the constitutive standards of agency. This kind of closure is unique to agency. What is at work in reflection is the distinctive operation of intentional agency in its discursive mode. What is at work is not simply the subject's capacity to shape her conduct in response to reasons for action but also her capacity both to ask for these reasons and to give them. Hence, agency's closure under reflective rational assessment is closure under agency's own distinctive operation: Agency is closed under itself.¹⁵

3.4 To sum up, agency is special because of two distinctive features. First, agency is not the only game in town, but it is the biggest possible one. In addition to instances of bare intentional agency, any engagement in an ordinary enterprise is *ipso facto* an engagement in the enterprise of agency. Second, agency is closed under rational reflection. It is closed under the self-directed application of its distinctive discursive operation, the asking for and the giving of reasons for action. The combination of these features is what makes agency *inescapable*. This is the kind of nonoptionality that supports the viability of constitutivism.

3.5 The inescapability of agency does not mean that there can be no entities that are utterly indifferent to it. It goes without saying that agency is *ontologically* optional. It is so even for us as biological organisms. Human animals are not necessarily rational agents. But this not the kind of optionality that is at stake in the debate on the grounds of normativity.

¹⁴ The clearest statements of what I call the 'closure' of agency under reflective rational assessment are found in Velleman (2000: 30-31; 142) and Velleman (2004: 290 ff.); see also Railton (1997: 317) and Rosati (2003: 522). For a similar closure in the theoretical domain, see Rysiew (2002:451) discussion of Thomas Reid's suggestion that the first principles of cognition are constitutive principles that operate as the fixed point of cognition.

In addition, the inescapability of agency does not imply the impossibility of dropping out of agency. First, there are brute and involuntary ways of both exiting from and entering into agency: one might nondeliberately fall asleep and wake up, lose and regain consciousness, die and (possibly) resurrect. Second, it is in principle always possible to opt out of agency in a deliberate and intentional manner; to act so as to bring about one's temporary or permanent exit from agency. An agent may commit suicide or, less dramatically, take the steps necessary to fall asleep, lose consciousness, or induce her temporary irrationality. But the subject who raises the question whether to commit suicide or interrupt her agency is not a shmagent. While she ponders whether to commit suicide, she is still living up to the standards of rational agency. For she is trying to figure out whether there is a good enough reason to leave agency. And if she decides to do so, she is still committed—as a rational agent—to sustaining her participation in agency as long as required to implement her intention to drop out of it (such as taking the necessary means to secure her successful suicide).¹⁶ The deliberate loss of rational agency—whether temporary or permanent—is supposed to be achieved as the culmination of an exercise of rational agency.¹⁷ The agent who contemplates the possibility of opting out of agency is not challenging the binding force of agency's standards. She is rather wondering whether there is *reason to continue* sustaining her participation in that enterprise in light of her particular circumstances. She is not professing an utter indifference to agency as such. She defers to and

¹⁵ Notice that closure under reflection is not to be confused with *stability* under reflection. The closure is not even a guarantee of this stability. As I discuss in § 7 below, there might be no guarantee that agency is able to validate itself.

¹⁶ See Velleman (2004: 291). Notice that a permanent exit from intentional agency might not coincide with biological death. The subject might go into a permanent coma, revert to a lesser kind of agent (a 'wanton', say), or turn into a 'weather-watcher' (see Strawson 1994). These entities are shmagents in the sense that they are indifferent to the constitutive standards of agency but, as I argue in the paper, they are not sources of troubles for constitutivism on account of their utter indifference to the standards of agency (which is not to say that some agents might find existence in the non-agential mode attractive and deliberately try to bring about their metamorphosis into a wanton or a weather-watcher).

¹⁷ See Velleman (2004) and Railton (1997: 313-317).

abides by the standards of agency in determining the fate of her future participation in it.¹⁸ In sum, agency can be inescapable in the sense required by constitutivism even if individual agents might deliberately opt out of it if they are offered a compelling reason to do so.

3.6 The inescapability of agency shows that the analogy between ordinary enterprises and agency on which the shmagency objection rests cannot be sustained. The idea of a ‘shmagent’ is introduced by Enoch to show that there might be subjects who are indifferent to agency and would therefore need a reason available *outside* of agency to be convinced to take part in it. The inescapability of agency, however, shows that there is no standpoint external to agency that the shmagent could occupy and from whence he could launch his challenge. If the shmagent is supposed to be an actual interlocutor in a rational argumentation, his professions of utter indifference to the standards of agency are self-undermining. Professing one’s indifference, challenging the force of the constitutive standards of agency, and engaging in a rational argumentation are all instances of intentional agency. The subject who genuinely participates in this sort of philosophical exchange is not truly indifferent to the standards of the practice of giving and asking for reasons. However, if he is already inside that enterprise, he cannot be pictured as asking to be offered a reason to opt into it. He might ask about reasons to continue staying inside but this would make the shmagent indistinguishable from a genuine agent, although one that might be contemplating the possibility of committing suicide.

Finally, the ontological optionality of agency, allows for the existence of genuine shmagents in the sense of beings who are truly and completely indifferent to the standards of agency. But these are not the kinds of beings that can raise philosophical challenges to constitutivism. We might even imagine running into a genuinely indifferent shmagent that makes sounds indistinguishable from the alleged professions of indifference like the one previously quoted

¹⁸ This is not to be confused with the sort of unacceptable conditional commitment to one’s agential

(‘Classify my bodily movements and indeed me as you like...,’ see § 2.2). But this encounter would be only a bizarre coincidence of no philosophical significance. It would pose no more of a threat to constitutivism than a parrot that has been taught to recite a ‘shmagency mantra.’

3.7 It is only under extraordinary circumstances that entities that are truly indifferent to the constitutive standards of agency might appear to be engaged in anything that resembles genuine intentional agency for sufficiently long stretches of time. Hence, there is something puzzling about one feature of Enoch’s description of the shmagent. He presents the shmagent as ‘being perfectly happy performing shmactions—nonaction events that are *very similar to actions* but that lack the aim (constitutive of actions, but not of shmactions) of self-constitution.’¹⁹ Why does it matter that shmactions are supposed to be *very similar* to actions? Given the shmagent’s utter indifference to agency, there is no basis to expect a systematic non-accidental similarity between the conduct of agents and that of shmagents. There is no reason to believe that the lives of shmagents could be very much like those of agents *but for* the shmagents’ indifference to the constitutive standards of agency.²⁰ Insisting on the similarity might make it easier to persuade us to think that the jurisdiction of agency is not as encompassing as it might initially appear, and that we should regard the shmagent as able to raise actual philosophical challenges. But the expectation of this similarity is unwarranted.²¹

unity that characterizes some *defective* forms of agency, as discussed by Korsgaard (1999: 22-23).

¹⁹ Enoch (2006: 179, my emphasis).

²⁰ The suggestion that the lives of shmagents might be just like those of agents but for the shmagents’ indifference to the standards of agency is similar to the explanation of the working of a radio offered in the following philosophical joke (which I first heard in Warren Goldfarb’s lectures on Wittgenstein at Harvard University): ‘X asks Y: How does a telegraph work? Y: Think of it this way. There’s a large, long dog with his head in Boston and his tail in Springfield. When you pat him on the head in Boston, he wags his tail in Springfield; and when you tweak his tail in Springfield, he barks in Boston. X: OK. But tell me: How does a radio work? Y: Just the same, but without the dog.’

²¹ The similarity would matter if the shmagency objection were interpreted as making a much weaker point against constitutivism; if it were interpreted as raising issues with the specific conception of agency adopted by particular constitutivist theories rather than with the constitutivist strategy in general, as I discuss more extensively in § 6 below.

4 *Alienated Participation*

4.1 Can the shmagency objection be reformulated so as to circumvent the inescapability of agency? Enoch suggests that, if agency is indeed inescapable, the shmagent should be conceived not as standing outside of agency but as an *alienated participant*. This alienated shmagent is introduced as someone who claims: 'I cannot opt out of the game of agency, but I can certainly play it half-heartedly, indeed under protest, without accepting the aims purportedly constitutive of it as mine.'²²

What kind of objection to constitutivism is raised by alienated participation? Presumably, an alienated participant still needs to be given a good enough reason to be an agent, although not in order to participate (given that she is already in) but rather to overcome her alienation, to wholeheartedly embrace agency and internalize its constitutive standards. And this reason cannot be produced simply as a result of her inescapable although alienated participation.

4.2 The initial appeal of this response to the inescapability of agency comes, once again, from drawing an analogy between agency and ordinary enterprises. Alienated participation seems to be unproblematic in the case of ordinary enterprises. For instance, one might play chess half-heartedly, without internalizing its aim. This alienated chess-player would simply go through the motions of chess; she would just *pretend* to be playing chess. She moves the chess pieces in ways that externally match the legal moves of chess. Perhaps, she even moves them in ways that externally match the strategically deft moves of someone who genuinely intends to win the game. Because of her alienation, however, she is not truly playing chess. She is not making an earnest attempt either at winning or even at making legal chess moves. If she is presented with the opportunity to terminate her alienated participation or to make an illegal move, she is ready to take immediate advantage of this opportunity if it helps her to advance whatever ulterior goal

²² Enoch (2006: 188).

motivates her pretense. This is because the constitutive aim of her alienated playing is not the same as the constitutive aim of chess; it is only parasitic on it.

Under special circumstances, a simulation or a pretense might be ‘inescapable’ in the sense that the agent might be forced to sustain it until the game is over (say, she might be forced to ‘play’ it at gunpoint). When such circumstances obtain, all the moves that the agent makes as part of her sham playing might *look* exactly like those of a genuine chess player, given that she might find that it is better for her to continue her sham playing through the end of the game. The apparent completion of the game, however, does not make her into a genuine player, since she continues to be moved by a different constitutive aim.

4.3 Alienated participation in ordinary enterprises is a genuine possibility but not one that can be used to show that there is a problem with constitutivism. Alienated participation in ordinary enterprises is not a good model for the alleged alienated participation in agency. In the absence of a plausible analogy with ordinary alienated participation, however, I do not know what to make of the suggestion that there could be an alienated participation in inescapable agency. To begin with, as we have just seen, pretending to participate in an enterprise is not a genuine instance of participation in that enterprise, not even when one is forced to sustain the pretense until the simulated enterprise is over. This means that no ordinary enterprise is strictly speaking inescapable. One is *not* playing chess when one is just pretending to.

In addition, the ways in which an ordinary enterprise might be said to be inescapable have nothing to do with the inescapability of agency. Agency is not inescapable in the sense of being coerced or forced to participate in it, which are the ways in which ordinary enterprises can be said to be inescapable. Agency is inescapable in the sense that it has the biggest jurisdiction and it is closed under its distinctive operation.

4.4 Ordinary examples of alienated participation, such as pretending, playacting, and simulating,

are still instances of intentional agency, no less than the genuine participation in the simulated enterprise. This is another manifestation of the inescapability of agency. This means that any kind of alienated participation in agency, if modeled on this kind of pretending, would have to count as an instance of genuine participation in agency. ‘Pretending to be an agent’ or ‘going through the motions of agency,’ if they are to be understood on the only plausible model of alienated participation that we have, are ultimately instances of non-alienated intentional agency. One can playact or simulate any particular action and activity, including particular instances of playacting and of simulation, but playacting and simulating are still instances of genuine intentional acting. What about pretending to be an agent *tout court*? If this is something that is done *outside* of agency, it offers no example of alienated *participation*, which is what Enoch is after. Instead, if the pretense is carried out within agency, it cannot be an instance of genuine alienated participation.²³ One cannot pretend to be an agent as such without genuinely being an agent at least as far as one’s intentional pretense is concerned.²⁴

4.5 Are there other possible interpretations of alienated participation? I could think of two, but neither helps Enoch’s case against constitutivism. First, one might think of inescapability in

²³ In a footnote, Enoch (2006: 190 fn 47) appears to concede this point but he mentions it almost in passing, which suggests that he does not think of it as especially damaging to this overall position.

²⁴ It is only in the context of the *development* of agency that a being that is not yet a full-fledged agent might genuinely pretend to be such an agent. This is what might happen, for instance, in certain forms of child play (see Shapiro 1999). This possibility, however, does not offer any support for the shmagency objection. There is nothing in Enoch’s presentation of shmagents that suggests that they are like children, that they perform less than full-fledged intentional actions as part of a process of maturation into adult rational agents. (This is not to deny that constitutivism faces some intriguing philosophical questions about the nature of developmental transitions into full-fledged agency, given that we come to adulthood not via abrupt and brute transitions but as a result of an extended and gradual process that includes browbeating, manipulative inducement, and simpler forms of rational argumentation.)

terms of some kind of psychological compulsion. This suggestion does not work, however, because the very possibility of being dissociated from the springs of one's conduct, which is the kind of alienation that accompanies this kind of compulsive behavior, is incompatible with the identification required by the very notion of full-blooded intentional agency.²⁵

Second, couldn't we think of alienated participation as a sort of reluctance to abide by the constitutive standards of agency? There is no denying that being an agent can be hard work. It is not unusual to balk at the prospect that we are expected to satisfy all the demands of rational agency. There might be times when we wish that the job of agency were easier, and we might therefore meet its demands with some 'reluctance.' Those agents who are especially sensitive to temptation, more prone to *akrasia*, or lacking in resolve might exhibit considerable recalcitrance in meeting the standards of agency and not be as wholehearted at it as an Aristotelian *phronemos*. But these familiar psychological phenomena do not raise any objection to constitutivism. The existence of imperfect and defective agents, and the half-heartedness that might be experienced by enkratic ones are not evidence that participation in agency is normatively optional. They are only evidence that this participation might be psychologically arduous.

5 *Shmagency and Skepticism*

5.1 In the previous sections, I have argued that the shmagency objection fails because it rests on untenable analogies between agency and ordinary enterprises. Both the original version of the objection and its restatement in terms of alienated participation fail to acknowledge properly the distinctive inescapability of agency. The failure of the shmagency objection, however, offers

²⁵ This is the point missed by Marmor (2001:38-39)'s presentation of the idea of estranged and alienated participation, which is one of the sources of Enoch's discussion of the alienated shmagent. The fact that one might think of oneself as alienated from the springs of action does not prove that one can be estranged from one's intentional agency. The argument rather runs in the opposite direction. Those

only indirect support for constitutivism. It does not eliminate the possibility of other challenges and objections.

5.2 One worry is that the strategy used to reject the shmagency objection exposes a troubling inherent weakness of constitutivism. Constitutivism responds to the shmagency objection by denying the possibility of shmagents as rational interlocutors who could launch a genuine philosophical challenge. Entities that are utterly indifferent to agency do exist but they pose no threat to constitutivism since they raise no rational challenges or objections. This means that constitutivism succeeds at *defusing* the shmagency objection by showing that there can be no shmagents. As a result, however, constitutivism is unable to *defeat* the shmagent by refutation.

According to Enoch, this shows a serious limitation of constitutivism. The problem arises because of the anti-skeptical aspirations expressed by some constitutivists. If constitutivism is expected to offer a refutation of skepticism about normativity, the appeal to the inescapability of agency might backfire. Constitutivism could only show that the skeptic is impossible but could not prove that he is wrong.²⁶

5.3 This problem does not arise if we are dealing with the shmagent rather than with the skeptic. The shmagent is not necessarily skeptical about the categorical force of the norms of practical rationality and morality. The shmagent only rejects the suggestion that the ultimate grounds of normativity lie in the constitutive standards of agency. The shmagent does not necessarily deny that those grounds could be found elsewhere. He might even accept the suggestion that the constitutive standards of agency play a crucial role in the derivation of the norms of practical rationality and morality. Even so, he would claim that their categorical force ultimately depends on the existence of a conclusive reason for us to be agents; a reason which cannot be provided,

aspects of one's psychology from which one could be alienated or dissociated are, because of the very possibility of alienation, inadequate to account for intentional agency (see Velleman 2000: chs.1 and 6).

²⁶ Cf. Enoch (190 fn 44).

however, by the constitutive standards of agency.

Although the shmagent does not have to be a skeptic about normativity, a skeptic might try to argue for his position by taking the shmagency route. This skeptic-as-shmagent would grant the relevance of the standards of agency for the derivation of substantive norms but argue that the possibility of shmagents shows that there is no categorical reason to be agents.

5.4 Against this kind of skepticism, constitutivism could effectively use the strategy already deployed against the shmagent. If there can be no space for the shmagent as a rational interlocutor, *a fortiori* there can be no space for the skeptic-as-shmagent. This kind of skepticism is defused by being disarmed rather than defeated by being refuted.

This conclusion is troublesome for those who insist that constitutivism provide a refutation of all versions of skepticism. But it is hardly evidence of some serious difficulty with constitutivism as a general argumentative strategy. The issue is only whether constitutivism should be embraced by those philosophers whose primary aspiration is the refutation of the skeptic in all of his possible guises, including the skeptic-as-shmagent one.

5.5 In any event, if one is willing to settle for a less ambitious anti-skeptical strategy, constitutivism still offers a variety of anti-skeptical tools. In addition to the defusing of the skeptic-as-shmagent, constitutivism is not barred from attempting actual refutations of those skeptics who do not take the shmagency route but launch their challenges while standing *inside* agency. Likewise, constitutivism is not barred from engaging in rational conversations with (and, if necessary, refutations of) *defective* agents—including massively defective ones, at least as long as they have not yet stepped outside of agency.

In sum, although constitutivism might be unable to refute every kind of skeptic, it still offers a combination of anti-skeptical weapons—including the possibility of actual refutations—that

many should find reasonably satisfactory. Whatever limitations constitutivism might exhibit on this front, they hardly count as a devastating objection to it.

5.6 The indispensability of agency does not rule out the possibility of genuine skeptical challenges launched inside of agency. This is why constitutivism might be able to engage in actual rational argumentations with these ‘internal’ skeptics and attempt to refute them. At the same time, this shows that constitutivism might still be vulnerable to a *reductio ad absurdum*. This is what any skeptic who does not take the ill-fated shmagency route is going to attempt against constitutivism.²⁷ Nonetheless, the inability of constitutivism to rule out *a priori* the bare possibility of a *reductio* can hardly count as a criticism of it. In the absence of any specific suggestion of how the *reductio* is supposed to work, all that one might ask of constitutivism is a generic profession of intellectual humility, i.e., the acknowledgment that it is not in principle immune from a *reductio*. But the burden of proof still lies with the skeptic; he is the one who has to show that constitutivism fails on the face of inconsistent commitments. In addition, this skeptic cannot find any support in the discussion of shmagency. None of the characterizations of shmagency that we have encountered thus far suggests that constitutivism might suffer from any internal inconsistency. There is one last concern with the anti-skeptical implications of constitutivism. In adopting a kind of transcendental argument against the possibility of the shmagent (and the skeptic-as-shmagent), constitutivism might exhibit the same limitations as the transcendental arguments used against epistemic skepticism.²⁸ Barry Stroud famously argued that transcendental arguments fail at deducing substantive truths about the world from nothing more than the necessary conditions for the possibility of our thoughts and experiences.

²⁷ Here I am in agreement with Enoch’s suggestion that skeptical challenges are best interpreted as ‘highlighting tensions within our own commitments, as paradoxes arguing for an unacceptable conclusion from premises we endorse, employing rules of inference to which we are committed’ and that the philosophical task thus is ‘not to defeat a real person who advocates the skeptical view or occupies the skeptical position (what view or position?) but, rather to solve the paradox, to show how we can avoid the unacceptable conclusion at an acceptable price,’ Enoch (2006: 183-184).

The transcendental arguments are unable to establish non-psychological conclusions—truths about how things are—from mere psychological premises.²⁹ This failure leaves room for more modest arguments, which remain confined within the psychological realm but establish connections between different ways of thinking that are indispensable for us.³⁰ The weaker arguments show that some of our beliefs are invulnerable in the sense that ‘no one could consistently reach the conclusion that although we all believe that things are as that belief says that they are, the belief is false.’³¹ The limitation of these more modest arguments is that they cannot prove that the skeptical possibility is false. They offer *no refutation* of skepticism. For beings with radically different cognitive faculties or conceptual schemes, the skeptical possibility might be a live one. But the skeptical possibility is inaccessible to us as rational subjects because it is *inconsistent* with the correct operation of our own judgment-sensitive attitudes.³²

5.8 What are the implications for constitutivism of the modesty of the transcendental arguments? The problem seems to be that constitutivism leaves the logical possibility of normative skepticism open. However, I think we should be cautious about accepting this conclusion. The conclusion is based on an analogy with the transcendental arguments adopted against epistemic skepticism. Couldn’t it be that the kind of confinement or inaccessibility of the skeptical possibility might be specific to the epistemic domain and might not extend to the practical and normative one? There might be enough differences between the nature of these domains and the skepticisms that they invite to warrant a closer look at the specific structure of the transcendental arguments applied against normative skepticism before declaring them modest.

²⁸ See Wallace (2004: 458), cf. Gibbard (1999: 154).

²⁹ Stroud (1968/2000).

³⁰ See Stroud (1999: 165).

³¹ Stroud (1999: 166). See also Hookway (1999: 177) and Taylor (1995: 26, 33).

³² See Hookway (1999:177-178).

In any event, how troubled should we really be about the modest import of transcendental arguments? As modest as they are, they tell us that in the *correct* exercise of our *full rationality*, and while relying on our own *most basic* conceptual schemes, we cannot be persuaded by skepticism, given that it is *inconsistent* with the operation of our rational faculties and our conceptual commitments.³³ For a modest claim, this seems to be quite strong to me.³⁴ But this might just be a matter of philosophical temperament. However, isn't talk of clashing temperaments the place where many discussion of skepticism eventually lead?

Finally, let's remember that the transcendental argument of constitutivism is successful against the shmagency objection, at least in its non-skeptical version. The issue raised by the shmagency objection is about the optionality of the engagement in agency, not about our dealings with all possible kinds of normative skepticism. With respect to the former issue, I maintain that the transcendental claims of constitutivism suffer from no troubling limitations. And this is all that we need to establish the viability of constitutivism.

6 *Constitutivism without Agency*

6.1 Despite the failure of the shmagency objection, the idea of shmagency might still be relevant to investigating the plausibility of constitutivism. In particular, concerns might be raised about the special role played by agency in constitutivism. Could we have constitutivism without agency?

One might accept the central claim of constitutivism—that categorical oughts are grounded on

³³ See Hookway (1999:178). The inaccessibility of the skeptical possibility is not a matter of some psychological *impediment*, as if we were unable to get rid of some obsessive thought or hang-up. It is rather a matter of the fully rational operation of our judgment-sensitive attitudes.

³⁴ If what the transcendental arguments prove is that the skeptical possibility is inaccessible to us because of the nature of rationality and the structure of our conceptual schemes, this limitation is not a fault of the transcendental *argument*, but a liability of *our nature* as rational beings. It seems to follow that even other anti-skeptical arguments, as long as they are launched inside of our conceptual schemes and while relying on our rational faculties, will be unable to refute skepticism. Likewise, if the transcendental arguments against normative skepticism prove to be similarly modest, then one cannot blame constitutivism for the weakness of its anti-skeptical import. Other meta-ethical views are supposed to suffer from exactly the same limitations, since they are *our own* limitations, not constitutivism's.

the constitutive standards of a special kind of enterprise—but reject the suggestion that agency qualifies as the special enterprise.

This proposal might take two forms. First, one might argue that the truly inescapable enterprise is some sort of uberagency, i.e., an enterprise that includes both agency and shmagency as optional subordinate enterprises.³⁵ Alternatively, one might argue that there is more than one inescapable enterprise. Shmagency might be as inescapable as agency.³⁶

6.2 These suggestions pose no serious threats to constitutivism if the notions of shmagency and uberagency are ultimately intended not to replace the *concept* of agency but to articulate a different *conception* of it. By a ‘conception of agency’ I mean a substantive articulation and specification of an otherwise uncontested concept of agency.³⁷ For instance, a discussion about whether agency is better understood in terms of self-understanding (as Velleman suggests) or self-constitution (as Korsgaard does) is a dispute among competing conceptions of agency. The undisputed concept of agency, instead, is meant to outline the basic structure of agency at a more general level. The concept is individuated by its role in relation to other equally general concepts such as—to mention a few—choice, intention, open alternatives, and autonomy. To illustrate, conflicting conceptions of agency would not disagree over statements like ‘agency is the capacity exerted when a subject acts intentionally as a result of her autonomous choice over alternatives she believes to be open to her.’ Statements of this sort are part of the articulation of the shared concept of agency.

³⁵ The regress argument against uberagency presented in § 2.5 above does not apply here, since the proposal under consideration accepts the constitutivist claim that the regress is stopped once we reach the level of the genuinely inescapable enterprise.

³⁶ Velleman (forthcoming) appears to read Enoch’s shmagency objection as suggesting something somewhat along these lines. I do not think that this is the best interpretation of Enoch’s argument, although this reading might be suggested by some remarks that Enoch makes in the original presentation of the shmagent, especially in his discussion of the similarity between shmactions and actions (see § 3.7 above). In any event, many aspects of Velleman’s response to Enoch can be persuasively applied to both readings of the shmagency objection (see §§ 6.6 and 7.5 below).

³⁷ For the distinction between ‘concept’ and ‘conception,’ see Rawls (1971: 5-6, 9).

Notice that, in spite of its generality, the concept of agency is sufficiently substantial to be the object of sustained philosophical scrutiny. The inescapability of agency, for instance, is a feature of agency that can be derived from the general features of the concept of agency. The defense of constitutivism presented in this paper is conducted at this level of generality. Nothing that I say here takes any stance about particular conceptions of agency.

6.3 The appeal to the possibility of shmagency or uberagency raises no concern about constitutivism if this appeal is interpreted as suggesting an alternative conception of agency, even if only in the guise of the more radical replacement of the concept of agency. Under this interpretation, the shmagent who says, ‘Classify my bodily movements and indeed me as you like,’ and ‘I am perfectly happy being a shmagent—a nonagent who is very similar to agents but who lacks the aim (constitutive of agency, but not shmagency) of self-constitution,’³⁸ is only targeting a *specific* account of the substantive constitutive standards of agency—the one formulated in terms of self-constitution. He is not really objecting to constitutivism. What he really means to say is something along these lines, ‘It is fine by me if you want to reserve the term ‘agency’ and its cognates to describe the enterprise aimed at self-constitution; the problem is that this enterprise is optional as evinced by my indifference to it and my ability to engage in a conduct that is very similar to agency in spite of my indifference to self-constitution. Therefore, *agency in the sense of the enterprise of self-constitution* cannot be the ground of the normativity.’ According to this interpretation, the shmagent is only making a linguistic concession to his opponent; he is promoting a different conception of agency under the label of shmagency. He is not really trying to replace the concept of agency as the genuine inescapable enterprise that plays the uncontested conceptual role articulated in terms of such notions as choice, intention, autonomy, etc. This kind of shmagent, therefore, is not really raising any

³⁸ See Enoch (2006: 179) and § 2.2 above.

problem for constitutivism about agency.

6.4 A true criticism of the focus of constitutivism on agency requires that the notions of uberagency and shmagency be meant as genuine replacements of the *concept* of agency. This is, however, no easy task. One cannot single out ‘agency’ as the only concept to be replaced. Doing without it requires finding a suitable replacement for the entire set of *agential* concepts, that is, the set of all those notions—such as action, choice, autonomy, reactive attitude, etc.—that are at least in part individuated in relation to the very idea of agency. The problem is that this constellation of agential concepts is one of the fundamental features of our conceptual scheme. It is only in terms of these agential notions that we can articulate some of the most fundamental distinctions that we make in our attempts at making sense of the world and of our relations to its denizens. The concept of ‘agency’ is essential to delineating the shape of a basic domain of reality—the domain of the ‘practical.’

6.5 For shmagency or uberagency to qualify as genuine conceptual alternatives to agency, they must be able to play a role in the shaping of conceptual schemes that is comparable to the one played by agency. This requires more than an abstract structural isomorphism with the distinctive features of agency. In principle, we might be able to make sense of phenomena that, at a suitably abstract level, might be said to be inescapable. That is, of phenomena that are closed under their distinctive operations—whatever those might be—and that have the largest possible ‘jurisdiction’ in their relevant domains—whatever those might be. But these structural similarities are not sufficient to show that we have a replacement for the concept of agency. The inescapability of agency is more than a formal property. The inescapability of agency is part of its distinctive *substantive* role, the role that makes the concept of ‘agency’ the linchpin of our understanding of the practical domain. Even the notions of ‘enterprise’ or ‘jurisdiction,’ which are used to characterize the structure of inescapability, seems ultimately to be concepts of a

practical/agential kind.

6.6 The notions of uberagency and shmagency could aspire to be adequate conceptual replacements of the concept of agency only if they could support a more *substantive* ontological and conceptual role comparable to the one of agency. They would have to do so, however, in their own *non-agential* terms. For instance, shmagency should bear comparable relations to the ‘shm-’ counterparts of such notions as action, autonomy, reactive attitude, and the like; that is, to *shmaction*, *shmautonomy*, and *shmreactive* attitude. In other words, as agency stands to the practical domain, shmagency should stand to a comparable domain, although one conceived in shmagency’s own terms—the *shmractical*, maybe?

What are we to make of a conceptual substitution of this kind? We could certainly make sense of the ‘shmractical’ notions as a mere notational restatement of the agential concepts but, of course, this would not be a genuine alternative to them. On the other hand, if the proposal purports to be more than a notational restatement, we need to get some grip on what this ‘shmractical’ domain is supposed to be like. But I think that, if we are pushed to this point, we have no conceptual resources in our repertoire that can help gain any insight into what this proposed replacement of the concept of agency is supposed to amount to. We have outstripped the limits of our conceptual imagination.³⁹ We cannot really do without the concept of agency. There is no problem with disputing about alternative conceptions of agency, but a replacement of the concept of agency in its role as a fundamental element of our conceptual scheme seems out of the question. This shows that there is a further sense in which agency is inescapable: agency is inescapable as an enterprise and indispensable as a concept.⁴⁰

³⁹ See Velleman (forthcoming: ch. 4) for a similar conclusion.

⁴⁰ Velleman (forthcoming: ch. 6) supports what he calls a ‘mild version’ of shmagency. He argues that, according to his conception of agency in terms of self-understanding, the principles of self-intelligibility might take a variety of forms. Some might consider a behavior intelligible in terms of a narrative understanding, others in terms of a causal-psychological explanation, others in terms of any combination of these two forms of explanation. The constitutive standards of agency remain silent on

7 *The Self-Validation of Agency*

7.1 In the previous sections, I have argued that the shmagency objection is unconvincing.

Because of the special status of agency—the inescapability of the enterprise of agency and the indispensability of the concept of agency—the question whether there is reason to be an agent cannot be raised and answered outside of agency. That is, it cannot be raised outside of the actual engagement in the agency enterprise and outside of a conceptual scheme in which the notion of agency plays an essential role. This does not mean, however, that constitutivism is exempt from the need to address the question, nor that it lacks the resources to do so.

Nevertheless, the question whether we have reason to be agents has to be taken up *within* agency. It has to be taken up by subjects who are not indifferent to the standards of agency and thus try to answer it in deference to these standards.

7.2 As previously remarked, the distinctive operation of agency in its discursive mode is the practice of giving and asking for reasons for action. Because of the inescapability of agency, when an agent wonders whether she has reason to engage in agency, she can do so only by applying the distinctive operation of agency over her own agency. How does this self-directed operation affect the justification of one's engagement in agency? Does it guarantee that there is always reason to be an agent? That is, is agency necessarily self-validating? Or is this self-directed operation unacceptably circular?

which principle of self-intelligibility should be adopted. Velleman points out that this pluralism is not necessarily a problem for constitutivism about agency. The choice of principles might be arbitrary, but constitutivism about agency does not purport to claim that there can be no space for arbitrary choices. I think that Velleman is correct on this point. There might actually be other dimensions along which the concept of agency is underdetermined (for instance, the temporal extension and structure of the unit of agency, on which see Ferrero (forthcoming)). I take issue, however, with his suggestion that this is a concession to the shmagency objection. The underdetermination in the concept of agency suggests that there might be pluralism in the ways in which agents might fully specify the structure and boundaries of what they take to be the inescapable enterprise. This pluralism is nonetheless internal to a shared understanding of the basic structure of agency, including its role as the inescapable enterprise. The fact that certain features of particular specifications of the structure of agency might be arbitrary does not make agency optional in the sense required by the shmagency objection.

7.3 Let's consider circularity first. One concern is that, as an agent begins investigating whether she is justified in being an agent, she must already be taking this justification for granted. After all, because of the inescapability of agency, she has to conduct her investigation in deference to the standards of agency. True, but the agent needs only assume that her participation in agency is *provisionally* justified. As an agent, she cares about her participation in agency. As she embarks in the investigation, however, she is not yet assured that her care is eventually going to be proven justified (although she might hope that it is going to). Because of the inescapability of agency, however, there is no other place from whence she could launch this investigation. Given that the acceptance of the standards of agency is only provisional, the agent is not really begging the question when she begins pondering whether she should be an agent.

7.4 The provisional character of one's deference to the standards of agency, however, does not dispel all worries about circularity. There is also a concern about the validity of the criteria used in determining what counts as a correct answer to the practical question about agency, to the question whether one should be an agent. When one asks whether one should engage in an ordinary enterprise, one adopts the criteria of correctness set by the nature of intentional agency. This is because one is to answer the question whether one is to engage in that enterprise as an intentional agent; that is, one is to show her engagement to be supported by reasons for action. When one asks the practical question about agency itself, one is to defer to the same criteria of correctness since this is what asking a practical question consists in. But here lies the problem. If the validity of the criteria set by agency depends on our being justified in engaging in that enterprise in the first place, there seems to be an unacceptable circularity in justification: The criteria used in determining whether one is justified in being an agent are the same criteria whose validity depends on one's being justified in being an agent.

7.5 I will argue that, in spite of the appearances, there is really no unacceptable circularity here.

To show this I need to make some preliminary remarks about the kind of circularity that might affect ordinary enterprises. (In the following discussion, I take my cues from an argument recently made by Velleman in support of a similar conclusion.)⁴¹

When trying to validate a *move* made within a given enterprise, one has to appeal to the criteria of correctness set up by the constitutive standards of that enterprise. A move in the game of chess, for instance, is correct if and only if it abides by the rules of chess. These rules are given prior to and independently of that move. Questions about the validity of the game's criteria of correctness, rather than of its moves, cannot arise within that game. This is because the criteria determine what it is for a particular conduct to count as a correct move in that enterprise. It would be a misunderstanding of their role as correctness-setting if we were to ask whether these criteria are valid in their own terms. The rules of chess are not chess-valid, so to say, since they determine what chess-validity amounts to. To ask for the self-validation of these criteria would be meaningless rather than circular.

7.6 For some more complex enterprises, such as theoretical or practical reasoning, there is the possibility of a genuinely vicious circularity. One might reason theoretically about one's theoretical reasoning, or reason practically about one's practical reasoning. For instance, one might want to establish whether a rule of inference like *modus ponens* is theoretically valid. In doing so, one cannot rely on that very rule. This would be unacceptably circular. For there is an independently established criterion of theoretical correctness—conduciveness to the truth—that must be used to validate the rule of inference. To validate the rule of inference, one is to show, *without relying on that very rule*, that the rule meets the criterion of theoretical validity; that the rule is conducive to the truth. No similar question arises, however, for the criterion of theoretical validity, as opposed to a rule of inference. The criterion of theoretical validity cannot be

⁴¹ See Velleman (forthcoming: ch. 4).

theoretically validated since its role is to determine what counts as being correct in theoretical reasoning.

7.7 Although it makes no sense to ask for the self-validation of the correctness-setting criteria of any enterprise, these same criteria might be said to be ‘justified,’ ‘valid,’ or ‘correct’ if they meet the standards set by some other enterprise. For example, one might be concerned with the playability of a game (Is it too easy? Too tedious? Are the rules inconsistent?) and thus speak of the validation of the game’s rules in terms of playability. Alternatively, one might speak of the validation of the criteria in terms of the practical question, ‘Should I be playing this game?,’ ‘Should I be engaged in this enterprise?’ In this case, one might say that the criteria are valid if playing that game is justified, that is, if one is justified in *adopting* the game’s criteria. But this practical validation, as it might be called, does not make the game’s criteria valid *qua* correctness-setting criteria for that game. The rules of chess determine what counts as a legal chess move whether or not there is ever any reason to play chess. Finally, one might ask whether the criteria of a particular enterprise are constitutive of that enterprise. As part of the investigation in the metaphysics of chess, for instance, one might ask whether ‘castling’ is a valid rule of chess, by which one means to ask whether ‘castling’ *is* one of the rules of chess.⁴²

7.8 Similar considerations can be made about the criteria of agency. First, whether a particular move in the game of agency is a valid move in that game depends on the criteria of correctness set up by the constitutive standards of agency. Second, any attempt at giving a practical

⁴² A further issue with the criteria of both theoretical and practical reasoning is whether one is justified in adopting them as *explicit* guides in shaping one’s conduct. For instance, there might be cases in which the goal of reaching the truth is better achieved if one does not explicitly conceive of one’s conduct as aimed at that goal. Likewise for practical reasoning. For such cases, one might be justified in adopting some criteria as determining the *objective* correctness of her conduct but not as the standards that she is to follow as *subjective guides* in determining her conduct (see Railton 1997). There is no circularity, therefore, in wondering whether there is reason to use an objective criterion of correctness of an enterprise as one’s subjective guide in that enterprise. Nor is it a failure of practical justification of an enterprise, if the use of its objective criteria as subjective guides cannot be practically justified. This is true even about the objective criteria of agency.

justification for the use of a rule of practical inference that relies on that same rule is viciously circular. A rule of inference is to be practically validated in terms of the independently given criteria of correctness in practical reasoning (whatever they are supposed to be). Third, it is a misunderstanding of the role of the criteria of correctness of practical reasoning to try to validate them in their own terms, to show that they are *practically* valid. Fourth, one might validate these criteria in terms of the standards set by a distinct enterprise. In particular, one might embark on an investigation into the metaphysics of agency in order to discover the nature of the constitutive standards of agency and of the criteria of practical correctness.

In other words, the following questions are to be kept separate: The metaphysical question about agency: What are the true criteria of practical reasoning? The question of correctness in practical reasoning: how do I go about telling whether the answer to a practical question is correct? The practical question about agency: Do I have reason to be an agent?

7.9 We are now in a position to see why the self-application of the criteria of correctness of agency is not viciously circular. When one asks the practical question about agency—‘Is there reason to be an agent?’—one relies on the criteria set by the nature of agency to answer this question given that this is a *practical* question. What one is asking is whether one should *adopt* these criteria in shaping one’s conduct. One is not asking whether the criteria are valid in their status as setting what counts as correct in practical reasoning. This status is rather presupposed in raising the practical question. The status remains unaffected by the answer that one is going to give to the practical question. Like in the chess example above, the standards of agency continue to set what counts as practically correct *whether or not* one has any reason to adopt those standards. The worry of circularity arises only because of the mistaken impression that the practical question is supposed to establish that the criteria are valid as the criteria that sets what counts as practically correct.

7.10 It is worth remarking that, in addressing the practical question, there is no provisional assumption of the criteria that determine the nature of practical correctness. We accept them as independently established. What is provisional is rather our *adoption* of them in shaping our conduct. The adoption is provisional since we are still trying to figure out whether we are eventually justified in adopting them (see § 7.3 above). Once again, there is no objectionable circularity in the provisional adoption of them. It is nonetheless true that, because of the inescapability of agency, the practical question can only be addressed within agency. This implies that the criteria of practical correctness determine the ultimate fate of their own adoption. *If* the engagement in agency turns out to be justified, the criteria turn out to be *self-ratifying*: they justify in their own terms our actual use of them in the shaping of our conduct.⁴³ This is not a matter of circularity, but of closure.

7.11 Notice the conditional nature of the conclusion reached above. The fact that, when raising the practical question about agency, constitutivism is not threatened by vicious circularity offers no guarantee that we necessarily have reason to be agents. The constitutivist strategy offers no guarantee about the self-justification of agency and the self-ratification of its criteria. That agency operates over itself whenever we raise the practical question about agency leaves open what the answer is going to be (at least to a point, see § 8.2 below.⁴⁴ The only thing that appears to follow immediately from the inescapability of agency is that the investigation about

⁴³ Millgram (2005) suggests that there are two strategies that constitutivism might adopt to show that agency is non-optional. First, there is the metaphysical strategy directed at showing that we *cannot but be* agents. Second, there is the practical strategy directed at showing that we *cannot refuse the offer* to be agents. It is interesting to notice that, if the self-justification of agency proceeds according to the structure illustrated in this paper, the validation of agency is the result of both strategies. To begin with, the self-justification depends on the inescapability of agency, which is a matter of the metaphysics of agency. The self-application of the operation of agency is mandated by the fact that we cannot but be, in the relevant sense, agents. At the same time, the self-validation is a matter of practical necessity: a fully rational agent cannot refuse the offer to participate in the enterprise of agency: participating in the enterprise is what her reflective practical reasoning tells her *to do*.

⁴⁴ Rosati (2003: 522) might thus be too hasty in suggesting that the operation of the motives and capacities constitutive of agency is self-vindicating just in virtue of their self-application.

the justification of agency cannot be conducted by someone who is indifferent to agency. But nothing immediately follows about the outcome of this investigation.

7.12 Whether we are justified in being agents is a substantive question whose answer depends on the particular criteria of correctness suggested by individual versions of constitutivism as spelled out in their specific *conceptions* of agency. I see no reason to deny that some of the criteria that might be advanced by specific versions of constitutivism might fail to prove that agency is self-justifying. However, to establish whether particular versions succeed at this task is a question that falls outside of the scope of this paper.⁴⁵ For present purposes, all that matters is the proof that constitutivism does not produce a self-justification of agency that is either trivial or viciously circular.

8 A Paradox of Self-Validation?

8.1 I have just argued that the inescapability of agency allows, although it does not guarantee, that there might be an unproblematic self-validation of the engagement in agency. In this final section, I will argue that the inescapability of agency seems also to imply that, at least under one particular respect, agency might be *beyond* validation. This is what might follow from the pragmatic paradox that would be faced by any agent if one were to discover that agency does not self-validate and that one has no reason to be an agent.

8.2 What happens if the criteria of practical correctness fail to ratify their own adoption? What is a rational agent to do if she were to discover that she has reason *not* to be an agent? If, while playing chess, a rational agent discovers she has a conclusive reason never to play chess, she is to immediately stop playing. The same should also be true of agency. Giving up one's participation is what reason demands in response to the discovery of unjustified engagements.

⁴⁵ Velleman (forthcoming) explicitly argues that his own constitutivist view in terms of self-understanding is self-ratifying.

Hence, if a rational agent discovers to have reason not to be an agent, she is supposed to give up immediately her participation in agency, that is, to stop being responsive to her acknowledgement of reasons for action. However, part of what she is no longer supposed to respond to includes the very reason that she has just discovered to hold of her, viz., the reason not to be an agent. The conclusion of her practical reasoning requires her to give up *hic and nunc* her rational agency *tout court*.

But this is not something that she can intentionally do as a *rational* agent. Any step that she would take in order to intentionally exit from agency would be *in compliance* with the demands of rationality, since in doing so she would be appropriately responding to the conclusion of her reasoning. At the same time, any step she would take to exit agency would also be *in violation* of the same demands of rationality. For, in taking that step, she would still be behaving, even if for the last time, *as a rational agent*, which is exactly what she has discovered she *is not to do*. Hence, the pragmatic paradox induced by the discovery that one's own agency fails to self-validate, that it fails to justify practically our own engagement in it.

8.3 The situation faced by the agent who discovers to have reason not to be an agent is different from the *non-paradoxical* cases of 'rational irrationality,' cases where one has reason to suspend temporarily one's rational responsiveness in order to increase her long-terms success as a rational agent (for instance, those situations in which one might induce one's temporary irrationality in order to avoid being the successful target of a coercive threat).⁴⁶ Rational irrationality is not problematic. It has the same structure as rational suicide, which—as previously discussed in §3.5—is not paradoxical. In both situations, when the agent is implementing her decision, she does exactly what she is supposed to do in response to the dictates of practical reason. Throughout this implementation, she continues to see herself as

⁴⁶ See Parfit (1984: 13).

justifiably bound by the constitutive standards of agency, since even if she is eventually going to opt out of agency forever, she is not giving up on agency *tout court*.

By contrast, if there is no reason whatsoever to be an agent, one cannot be required to *intentionally* opt out of agency. Rather, the demand should be that the agent immediately *disappear* but not as a result of any exercise of her intentional agency, of her responsiveness to reasons. What practical reason seems to demand, in this case, is not that the agent *do* something but that something *happen* to her. However, happenings are exactly the kind of things whose occurrence cannot be directly *demand*ed.⁴⁷ The paradox is that the rational agent is required to respond to reasons by immediately and permanently ceasing to be responsive to them. But at that point any intentional action would be one intentional action too many.⁴⁸

8.4 How does this paradox bear on constitutivism? What generates this paradoxical possibility is the inescapability of agency. This does not mean that constitutivism is weakened by it. The paradox does not show that constitutivism would *fail* at the self-validation of agency. The paradox does not show that the conceptions of agency championed by individual versions of constitutivism are unable to provide a self-justification of agency. Rather, the paradox raises some doubts about the *legitimacy* or the *intelligibility* of asking that agents validate the exercise of their own agency. This is because there is something peculiar to raising a practical question whose negative answer would throw the agent into the pragmatic paradox illustrated above.

⁴⁷ There can only be an *indirect* rational demand for happenings in that an agent might be required to create the conditions for something to happen at a later time. But this indirect demand is not what reason requires in the case discussed in the main text.

⁴⁸ The problem is not one of the timing of the response. The problem is not that the action would be one action too late, as it might happen if one discovers that one *has no reason* to do something that one *has just done*. If agency fails to self-validate, it fails to self-validate atemporally. If an agent discovers that there is no reason to be an agent, she would thereby find out that she was never justified in her past actions. With respect to those actions, there is nothing that she can do, now or ever. But this is not paradoxical. It is simply the consequence of the impossibility of changing the past. The paradox arises, however, about the present exercise of agency in its necessary projection into the immediate future, which is the time where one is supposed to begin discharging the rational demand (Don't be an agent!) that one has just discovered to apply to oneself.

But even if we do not go as far as claiming that the question ‘Should I be an agent?’ could not be legitimately raised, the pragmatic paradox that would be generated by a negative answer to this question reinforces one of the basic claims of constitutivism about agency: the special status of agency vis-à-vis the issue of its practical justification. Agency is an enterprise of a very different kind and nature from ordinary enterprises. Drawing analogies between agency and other enterprises is a very risky strategy, since the purported analogies might be more misleading than illuminating, as the arguments in this paper has shows. It is even questionable whether it is really appropriate to speak of agency as a ‘game’ or as an ‘enterprise,’ as it is often done in the literature on constitutivism, this work included.

8.5 In closing, let me briefly mention some ways in which the special structure of agency as the inescapable ‘enterprise’ affects questions about its practical justification.

For any ordinary enterprise, the basic form of the question whether to engage in that enterprise looks the same whether or not the agent is already engaged in that enterprise.⁴⁹ Not so for agency. To begin with, it is unclear whether there could really be a reason to participate in agency for any subject who is not yet an agent. It might not make sense to ask for reasons to opt into agency, given that there is no intentional action of *opting into* agency. *Transitions into* agency cannot be imputed to the agent: the agent is the end product, not the initiator, of these transitions.⁵⁰ In a similar fashion, the pragmatic paradox discussed above might suggest that it makes no sense to ask about the practical justification for the agent’s *current* exercise of her agency. At the same time, questions about the justification for intentionally opting out of agency

⁴⁹ This is a claim about the ‘basic form’ of the question. I am not denying the obvious fact that answers to specific practical questions are often path-dependent: In particular cases, whether one is already engaged in a given enterprise often makes a difference to the practical question whether she is to engage in that enterprise.

⁵⁰ Cf. the ‘paradox of self-constitution’ in Korsgaard (2002: §1.3.2-1.3.3).

in the future are perfectly in order, as already shown in the discussion of suicide and rational irrationality.⁵¹

8.6 Another important issue concerns the ‘guise’ of agency under which the question of the justification of agency is raised. In this paper, the discussion has been conducted not at the level of specific conceptions of agency, but at the level of the undisputed but general *concept* of agency (see § 6.2). At the latter level two significant negative results have been achieved: the rejection of the threat of circularity and the discovery of a paradox in the self-validation of agency. But the concept of agency might turn out to be too generic for a constructive argument in support of the actual self-validation of agency. It seems that a positive self-validation can only be found by looking at the more concrete characterizations of agency spelled out in the *conceptions* of agency championed by specific versions of constitutivism.

8.7 The difference between concept and conception is also relevant to the derivation of substantive normative claims from the constitutive standards of agency. The concept of agency might offer too thin a ground for this job. The richer characterization of agency offered by specific conceptions appears to be a more plausible starting point for the derivation of substantive norms and requirements. Whether this is sufficient to assuage the worries about the normative fertility of constitutivism, however, is not a question that can be addressed in this paper. My goal has only been to prove the general viability of constitutivism against the shmagency objection and the worries about circular self-validation. Putting these concerns to rest only completes half of the job in defense of constitutivism. We still need to consider whether constitutivism can be true to its other ambitious aspiration, the derivation of *substantive* norms

⁵¹ David Wiggins raises the possibility of a similar asymmetry in the case of the related issue of the existence of reasons to support the temporal continuity of individual lives: “[O]ne may muster the courage to ask the question what is so good, either absolutely or for me, about my own mental life’s flowing on from now into the future. Surely this depends on what kind of person I am or think I am, and what sort of mental life it is. Well, not quite. There is something instinctive here and as irreducible as the rational

and principles from the constitutive standards of agency. But this is a question for another occasion.⁵²

commitment to make prudent provision for the future. These are things that we need reasons to *opt out of* rather than things that we have to look for reasons to *opt into*,” Wiggins (1979: 307, my emphasis).

⁵² An earlier version of this paper was presented at the Metaethics Workshop at the University of Wisconsin at Madison in 2007. I am grateful to the audience for their comments and criticisms, especially to David Copp, Connie Rosati, Mark Schroeder, Jacob Ross, and Peter Vranas. Many thanks to Russ Shafer-Landau for the wonderful job he did in organizing the workshop and editing this volume. My interest in constitutivism was first sparked by an invitation to comment on Peter Railton’s work at a symposium in his honor at the University of Rome. I thank Tito Magri, Peter Railton, and Barry Stroud for the comments and encouragements I received on that occasion. I thank Jennifer Morton and Assaf Sharon for illuminating conversations at the early stages of this project. I am very grateful for the extensive written comments I have received from David Enoch, Elijah Millgram, and two anonymous reviewers.

References

- Cohen, G. A. (1996) 'Reason, humanity, and the moral law,' in C. Korsgaard, *Sources of Normativity* (Cambridge: Cambridge University Press).
- Coleman, Mary C. (unpublished) 'Toward Modest Constructivism.'
- Darwall, Stephen (1992) 'Internalism and Agency,' in J. E. Tomberlin (ed.), *Philosophical Perspectives* (Atascadero, Ca: Ridgeview), 155-174.
- Enoch, David (2006) 'Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action,' *Philosophical Review*, 115: 169-198.
- Ferrero, Luca (forthcoming) 'What Good is a Diachronic Will' *Philosophical Studies*
- FitzPatrick, William J. (2005) 'The Practical Turn in Ethical Theory: Korsgaard's Constructivism, Realism, and the Nature of Normativity,' *Ethics*, 115: 651-691.
- Gibbard, Allan (1999) 'Morality and Consistency in Living: Korsgaard's Kantian Lectures,' *Ethics*, 110: 140-164.
- Hookway, Christopher (1999) 'Modest Transcendental Arguments and Sceptical Doubts: A Reply to Stroud,' in R. Stern (ed.), *Transcendental Arguments: Problems and Prospects* (Oxford: Oxford University Press), 173-187.
- Kolodny, Niko (2005) 'Why Be Rational?' *Mind*, 114: 509-563.
- Korsgaard, Christine (1996) *The Sources of Normativity* (Cambridge: Cambridge University Press).
- ____ (1997) 'The Normativity of Instrumental Reason,' in G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 215-254.
- ____ (1999) 'Self-Constitution in the Ethics of Plato and Kant,' *The Journal of Ethics*, 3: 1-29.
- ____ (2002) 'Self-Constitution: Action, Identity and Integrity. *The John Locke Lectures*.'
- Lavin, Douglas (2004) 'Practical Reason and the Possibility of Error,' *Ethics*, 114: 424-457.
- Marmor, Andrei (2001) *Positive Law and Objective Values* (Oxford: Oxford University Press).
- Millgram, Elijah (1997) *Practical Induction* (Cambridge (Mass.): Harvard University Press).
- ____ (2005) 'Practical Reason and the Structure of Actions,' in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
- Parfit, Derek (1984) *Reasons and Persons* (Oxford: Oxford University Press).
- Railton, Peter (1997) 'On the Hypothetical and Non-Hypothetical in Reasoning about Belief and Action,' in G. Cullity and B. Gaut (eds.), *Ethics and Practical Reason* (Oxford: Clarendon Press), 53-79.
- Rawls, J. (1971) *A Theory of Justice* (Cambridge (Mass.): Harvard University Press).

- Rosati, Connie (2003) 'Agency and the Open Question Argument,' *Ethics*, 113: 490-527.
- Rysiew, Peter. (2002) 'Reid and Epistemic Naturalism,' *The Philosophical Quarterly*, 52: 437-456.
- Schapiro, Tamar (1999) 'What is a Child?' *Ethics*, 109: 715-738.
- Setiya, Kieran (2003) 'Explaining Action,' *Philosophical Review*, 112: 339-393.
- ____ (2007) *Reasons without Rationalism* (Princeton: Princeton University Press).
- Safer-Landau, Russ (2003) *Moral Realism: A Defence* (Oxford: Oxford University Press).
- Sharon, Assaf (unpublished) 'Unreflected Glory: A Challenge to Metanormative Constructivism.'
- Strawson, Galen (1994) *Mental Reality* (Cambridge (Mass.): MIT Press).
- Stroud, Barry (1999) 'The Goal of Transcendental Arguments,' in R. Stern (ed.), *Transcendental Arguments: Problems and Perspectives* (Oxford: Oxford University Press), 155-172.
- ____ (2000) 'Transcendental Arguments,' in B. Stroud, *Understanding Human Knowledge* (Oxford: Oxford University Press), 9-25.
- Taylor, Charles (1985) 'What is Human Agency?' in C. Taylor, *Human Agency and Language: Philosophical Papers* (Cambridge: Cambridge University Press), 15-44.
- ____ (1995) 'The Validity of Transcendental Arguments,' in C. Taylor, *Philosophical Arguments* (Cambridge (Mass.): Harvard University Press), 20-33.
- Velleman, J. D. (2000) *The Possibility of Practical Reason* (Oxford: Oxford University Press).
- ____ (2004) 'Replies to Discussion on *The Possibility of Practical Reason*,' *Philosophical Studies*, 277-298.
- ____ (forthcoming) 'How We Get Along?' (Cambridge: Cambridge University Press).
- Wallace, R. J. (2004) 'Constructing Normativity,' *Philosophical Topics*, 32: 451.
- Wiggins, David (1998) 'The Concern to Survive,' in D. Wiggins, *Needs, Values, Truth* (New York: Blackwell).