# "Daring to Believe": Epistemic Agency and Reflective Knowledge in Virtue Epistemology

Fernando Broncano
Universidad Carlos III de Madrid
Fernando.Broncano@uc3m.es

## A feeling of paradox

"Knowledge" refers to a theoretical achievement while "agency" refers to a practical one. They are achievements of a subject able to hold beliefs, to have purposes and to make judgments and choices. In both cases, too, the subject eventually needs to reflect about her own position in order to make an ending judgment or decision. By taking such reflective stance the agent gets a reason to make up her own mind and hence she reaches a result in a fully intentional way. Suppose that one is asked: "Do you know that?", or, "Can you help me?" If accepting the challenge, the addressee will bring about a conscious answer, yes or no; and such choice will demand from her a careful evaluation of her own cognitive position, or ability to help the requester respectively.

Although the overall reflective process presents different facets, it is likely to suppose that some sort of self-knowledge should be available to such agent. However, this putative self-knowledge raises two controversial questions becoming the topic of a large amount of literature: the first is the supposedly transparency or luminosity of the states involved in the bringing about the asked answer (Williamson (2000). The second one, which precisely represents our main target in this paper, is a related question about what should be a necessary degree of integration for the faculties in charge of such reflective process. Since it is plausible to imagine that a non well-integrated mind could conclude in a possible right belief, and, from a normative point of view, such belief still could fail in attaining theoretical or practical success, the problem of integration arises as a question of epistemological force. Our aim is to argue that integration is a main requirement, in addition to other virtuous traits of the

agent's character, for a full agential response from the subject. "Subject" will mean here a very strongly normative level of the agent's constitution in the epistemic and practical domains.

Actually, one could state transparency and integration as a priori requirements, now then, as considering the issue from a naturalistic view, likely they should not be taken as given data, but as mere symptoms of a "healthy" degree of agential excellence. They are empirical signs of the righteousness of the own perspective before an epistemic or a practical task, or more precisely, signs of epistemic or agential "quality". Reflection will be, then, a stance the agent takes in order to estimate the quality of her position. The agent "reflects" about her own forces before to accept some belief, or before to act. Nevertheless, "reflection" can be a confusing term depending on how we consider the nature of the process in which the agent is involved. We could consider, for example, "to reflect" as an introspective process of inspecting the own mental states. This conception exploits a "visual" metaphor, as the very etymology of "introspection" suggests (looking within). That is the notion of reflection the legend of the Cartesian tradition seems to picture. "Reflecting" would be, then, to take some evidence from the own mental states about their epistemic or agential quality. But such "evidentialist" notion of reflection raises a new sequence of questions, since the "visual" metaphor seems to be so intuitively weird as metaphysically unjustifiable: Who looks?, where looks?, what sees as evidence?, etc. The point we want to consider is how to preserve a reflective stance without falling in a Cartesian image of it.

To begin with, we feel a certain sensation of paradox emerging from the notion of reflection in Cartesian lights. As an analogy, let us recall the well-known Vermeer's canvas *The Art of Painting*:

When one wonders what is represented in the picture, this famous canvas presents an intriguing paradoxical appearance: How it is possible to portray the own back while one is painting some scene? Certainly, one could regard the picture as representing the percept held by some indeterminate observer. But, how it is possible to depict perceptions of another mind? Clearly, between two impossibilities, one is led to interpret the canvas as representing an imaginative simulation of any one mind. That is, the canvas portrays a meta-representation. Now then, this answer leads us to inquire whose metarepresentation this is. Perhaps was Vermeer portraying himself?; or by contrast, was he picturing another person seeing him while painting?; or, finally, was he depicting a general scene of somebody looking at some other painting? The picture is the same, but the represented subjects are very different in the three cases.

Let us now consider the virtue epistemology, as exposed by Ernest Sosa in his recent twofold volume "A *virtue Epistemology*". Curiously, the jacket book of the first volume shows the Vermeer's painting. Not surprisingly, because a similar scenario is the one in which E. Sosa addresses the question of reflective knowledge, when escaping from the consequences of the skeptic's demands claiming that, in order to know that *p* one must also know that one is not dreaming (as well other undermining possibilities). The Sosa's answer is relevant because this requirement is very entrenched in the very folk notion of knowledge. Sosa distinguishes between "to know that one knows that *p*" and to have a reflective competence for valuing the risk one is taking when accept the deliverances of the own cognitive faculties. In order to attain a "virtuous" evaluation of the own epistemic status in a certain circumstance, the agents are gifted with meta-competences that eventually produce reflective knowledge. Such reflective knowledge is the response to philosophical skepticism[1] according to E. Sosa's virtue epistemology. It is relevant to note that such a faculty must be a kind of *meta-competence* that reaches further than the mere coherence of beliefs. For one thing, "coherence might conceivably be detached from the environing world of the thinker, so as to deprive him of reliable access to truth"[2]. Fleeing from the threats of circle or regress, Sosa stipulates this meta-competence with the function of examining the quality of the agent epistemic position. Reflective knowledge aims to exclude luck from the epistemically apt formation of true belief. This level provides a higher quality to the knowledge.

---

[1] Sosa expresses in this way the philosophical scepticism:
"A1. Any theory of knowledge must be internalist or externalist.
A2. A fully general internalist theory is impossible.
A3. A fully general externalist theory is impossible.
C. Therefore, *philosophical scepticism* is true"
Sosa, E. (2009) *Reflective Knowledge. Apt Belief and Reflective Knowledge, Volume II*. Oxford: Oxford University Press p. 154
[2] Sosa (2007) p 190

Sosa defines meta-knowledge in this way:

"Reflective knowledge goes beyond animal knowledge, and requires also an apt apprehension that the object-level perceptual belief is apt. What competence might a believer exercise in gaining such meta-apprehension? It would have to be a competence enabling him to size up the appropriateness of the conditions" (Sosa (2007) p 108)

Reflective knowledge then turns out to be the achievement of an epistemic meta-competence; that is, of a faculty or disposition to aptly evaluate the epistemic position and circumstances of knowing. Now then, the question is that, depending on how we conceive this competence, the evaluation could or could not discriminate among evaluative results as our possible interpretation of the canvas exemplifies. That is, the question is what is the process the meta-competence evaluates: it can be "the complete subject" from a kind of estrange third- person point of view, external to the first- person point of view or, by contrast, it can be a phenomenal calibration of the own situation. The point is that the subject is involved in very different ways in each case. The question we address will be then just about the degree to which the subject must be involved in reflective knowledge.

A first step is to note the particular contribution to the epistemic value that reflective knowledge confers to the overall process of knowing. For one thing, reflective knowledge adds justification to the first-order aptness as it strengthens the cognitive success in the particular circumstances by contributing to reduce luck in this achievement. The device that affords such justification is formed by two components, according to Sosa. The first one is the following:

Principle of epistemic ascent: "If one knows full well that p and considers whether one knows that p, then one must be justified in thinking that one does" (Sosa (2007) p.114)

The second one excludes possible undermining alternatives:

Principle of closure of epistemic justification: "If one is fully justified in believing that necessarily, unless it is so that q, it cannot be so that p, then one must also be justified in believing that q" ( Sosa (2007) p 115)

Stepping up the two principles, the subject is able to form the judgment that she justifiably knows that *p*. This judgement is enabled by the

Principle of criterion: "PC2. In order to know full well that p one must be justified in believing (at least implicitly or dispositionally, if not consciously) that

one's belief that p is formed in a way that is at least minimally reliable, that it has at least minimally reliable source (if the proposition that one's source is thus reliable is within one's grasp)" (Sosa (2007) p. 122) [3]

It is interesting to note that Sosa allows that the rational endorsement of reliability can be produced in unconscious or implicit ways. The reason is that reflective knowledge comes from a disposition to correctly evaluate the reliability of the faculties, and this disposition could work in some different ways. Reflective as well as unreflective knowledge both produce apt true beliefs, and this production, Sosa argues, is at some point independent of the degree to which subjects are voluntarily engaged. The sole condition is that the subject can be confident about her reflectively obtained belief, and it can be a result of an overwhelming disposition to confidently believe.

Answering a previous version of this paper (Sosa (2009b), Sosa distinguishes between the constraints of the assertions and the beliefs. Assertions, public as well as private, are voluntary judgments, and therefore submitted to the qualitative conditions of any intentional action. But confident beliefs do not necessarily have to be conditioned in such a way. It is sufficient that they come from a particular meta-apt faculty or disposition.

I concede that assertion is not a necessary condition for knowledge, although it could be for other aims, such as ulterior inferences from the belief, as well as the communicative or testimonial uses of the belief. My point is that the analogy between true belief and intentional action can be pursued beyond the voluntary and full intentional level (the level of full assertive and intentional acts) to a deeper level of a normatively constrained level of constitution. In this sense, I argue that aptness and control are normative conditions of knowledge and action respectively, and that both properties share the same requirement of good cognitive integration for the subject. Moreover, I will contend that a mere meta-apt well functioning cannot be sufficient to achieve knowledge and purposive action respectively.

---

[3] A kind of unreflective justification can be conferred by the first-order aptness. But this second-order meta-aptness provides a rational justification: *Reflective* rational justification, by contrast, is acquired at least in part through rational endorsement: *either* through endorsement of the *specific reliability* or one's basis (or at least the *safety* of one's basis, of the fact that it would not lead one astray in delivering the deliverance that p) *or* through endorsement of the *generic reliability* of one's basis" (Sosa (2009) p 239. The difference between the two kinds of justification is the lack of endorsement (in the case of unreflective justification) and the necessary rational endorsement (in the case of reflective knowledge.

Let us imagine William Tell doubting about his ability to safely hit the apple on his son's head at the very moment of shooting. Let us consider now the content of the following propositions:

1. I am skilfully prepared to shoot

2. William Tell is skilfully prepared to shoot

3. I believe that I am skilfully prepared to shoot

4. William Tell believes that William Tell is skilfully prepared to shoot

5. William Tell believes of himself that he is skilfully prepared to shoot

The belief that one is prepared to shoot is the same, but obviously, the five possibilities are very different according to their consequences. (1) and (2) are propositions that can express the knowledge state of William Tell transparently. The other ones, (3), (4) and (5) can be involved in assertive judgements as well as in testimonial cases, but, according to Sosa, the trustworthy beliefs (1) or (2) are the only ones required for reflective knowledge. Bill Tell is confronted with exactly three options that are open to him:

> "(…) (a) " No, I don't know that," or (b) "Who knows whether I know it or not"; maybe I do, maybe I don't," or (c) "Yes that is something I do know" ( Sosa (2007) p 115)

Here is where the integration problem appears for the subject. According to Sosa,

> "Answer (a), and even answer (b), would reveal a certain lack of integration in that stretch of consciousness; only answer (c) of the three, entirely avoids disharmony within that consciousness at that time" (( Sosa (2007) p. 115)

Notice that Sosa is addressing here assertions of knowledge, and it is in this context where he detects a possible lack of integration. An ambiguity, in this sense, in this sense, lies here between the possible state of knowing and the positive assertion of that, that is, between (1) and (2) possible propositions ascribing knowledge to William Tell. How self-confidence and knowing are related in the subject is the question of integration that I wish to address.

Why does this question of integration turn out to be the apparently paradoxical scenario exemplified by the Vermeer painting? To answer the question we must notice

that the meta-competence that reflective knowledge shows us has to evaluate the belief in a particular way:

> "The object of evaluation is thus a particular item, but it is evaluated relative to its relevant wider context. And the wider context may include possibility space, as when an archer hits the bull's-eye with a shot that it is not only accurate but also "skilful", with its counterfactuals implications" (( Sosa (2007)  p 114)

Sosa considers apt belief as analogous to the apt shooting of an archer. Our example consciously blends the two cases by focusing on the self-confidence of William Tell when shooting. His belief is evaluated in this particular circumstance, but this evaluation is made in the framework of a possibility space within  which  a counterfactually supported relationship between skill and success is established: "this successful shooting is due to the archer's ability".  This relation derives its modal strength from the links between the agent's faculties and the particular causal circumstances of shooting. The same works in the case of knowledge, that is, a true and confident belief coming from cognitive faculties working in the evaluated circumstances (in this case, the belief is about one's own abilities to shoot).

The feeling of paradox would disappear if an external referee was the one who evaluated the archer's merits. The question arises because the referee and the evaluated subject are the same person. The knowing subject is self-ascribing a competence: Is the content of his evaluation the belief (1) or, by contrast, the belief (2)? Compare this question with the  case of the Vermeer painting: Truly, a representation of a shooting archer is not a shooting, to differentiate the painting of a painter painting, which it does. Some processes -- and, by the  way, some abilities— have self-referential properties: painting and knowing putatively have them, shooting does not. Imagining, language, simulating minds, metacognition are probably cognitive examples of self-referential abilities.

The point is that self-reference preserves the content when passing from a broad context to another narrower context.  This is the case when one focuses on the painter of *The Art of Painting*, or, by contrast, when one scales up to the whole canvas. The content of the primitive image is preserved, however the meaning is now very different, for it passes from a third-person view to a first-person view. Does the same thing happen when reflectively one knows something? Because the subject must pass from an objective description of her epistemic status to a first-person description (this is the difference the principles (1) and (2) exemplifies.

First of all, let us consider the difference between the question addressed here and the coherentist quarrels. Some epistemologists consider that the border between external and internal contexts also divides the space of causes from the space of reasons. But, the integration requirement we postulate refers, as Breyer and Greco also maintain, to the objective/subjective divide more than the external/internal one (Breyer, D.; J. Greco (2008)). Sosa seems to be conscious of the coherentist objection as for example when he considers it in this way:

> "No belief B is fully justified because it satisfies some condition F such that beliefs satisfying F are probably true. The believer must also be aware, at some level, that B satisfies the condition" (Sosa (2007) p 124)

But he considers that this requirement is properly met by the exposed principles:

> "Most interesting for us is the fact that Bonjour's Generalization (of Sellar's insight) is a member of our family of principles of the criterion" (Sosa (2007) p 127)

Reflective knowledge, analogously to Cartesian *scientia,* is functionally defined:

> "Scientia requires more. It is attained only through an adequate perspective on one's epistemic doings" (130)

And the increased quality reflective knowledge added to the animal knowledge is accounted by this functioning:

> "Knowing  full well thus requires some awareness of the status of one's belief, some ability to answer that one does know or that one is epistemically justified, and some ability to defend this through the reliability of one's relevant competence exercised in its appropriate conditions" ( Sosa (2007) p 132)

Notice that Sosa judges here as sufficient symptoms of possessing reflective knowledge to have "some awareness of the status of one's belief, or to have "some ability to answer that one knows", as well as to have "some ability to defend this"[4]. Now then, that which keeps still here of the archery analogy is that it exemplifies aptness as a way of luck avoiding.  That is, that successful shots are those that hit the target not by luck. I guess that the example of William Tell doubting about his shot could provide a suggestive model for that. In the following section I will discuss the possible analogy between beliefs and acts according their metacognitive statuses.

---

[4] This nature  of the  ability to balance the epistemic status of one's belief also provides an answer to the alleged no luminosity of beliefs that Williamson maintains. For, although a belief lacks means to indicate its epistemic status, a meta-competence can do so.

**Making up the mind and self-calibration**

A well-known objection to the analogy between beliefs and acts, regarding their deliberative ascent, is that of their voluntariness.

> *Practical status of beliefs*: Beliefs are different from acts in that beliefs are not voluntarily held but actions are necessarily chosen to be intentional acts.

Beliefs, according this principle, are intentional because of their intentional content, and not because of their intentional production, as the actions are. A belief $p$ could be supported by a reason $r$ even though one believes $p$ independently of $r$. To be a reason $r$ for $p$, it suffices to have the proper logical (or epistemological) relationship with the content of $p$. By contrast, an act $a$ is supported by a reason $r$ if and only if r provides an explanation of the choice of $a$ instead of another alternative one. This different way of production would have consequences about the required consciousness for believing and acting. For knowledge can be achieved sufficiently by reaching apt beliefs. Moreover, a meta-apt belief suffices to reach reflective knowledge, , no matter if this belief was explicitly and consciously produced or not. All that it is required is a well-functioning metacompetence.

> "(...) It does lead me to think of the difference between the two sorts of knowledge, the animal and the reflective, as difference of degree. The higher brutes may be credited, along with small children, with some minimal degree of perspectival, reflective knowledge, of the implicit, subconscious sort, which largely resides in hosted inference patterns" (Sosa (2003) p 129).

Obviously, this primitive knowledge may have a very low epistemic quality. Sosa accepts it:

> " In richness, explicitness, and explanatory power, that  falls short of the reflective knowledge to which a human can aspire, especially someone philosophically inclined" (Sosa (2003) p 129).

Nevertheless, the significant point is that reflective knowledge can be attributed to creatures lacking higher degrees of deliberative consciousness. Surely, things turn out different when someone asserts belief in a public context, for example, when  giving forensic testimony in court. Then, the reflectively held belief amounts to being part of the act of the assertion. But, according to Sosa, from the point of view of conditions to reach the status or meta-apt belief, the full consciousness is not required for belief, as the traditional coherentist epistemologist surely claims.

Firstly, we will examine this view by considering an empirical candidate to accomplish a function similar to that expounded by Sosa; secondly, we will argue that a malfunctioning of this cognitive mechanism could sheds some light on the integration problem; thirdly, we will conclude from this example that Sosa needs a kind of involvement of the first-person point of view that is not necessarily equivalent to assertion. Therefore, we will conclude that a first-person perspective reliably means that an integrated epistemic agent is in charge of the task of knowing.

The cognitive function we are referring to is one which psychologists and neurologists have dubbed as *metacognition* (Koriat, A. (2000), Proust, J. (2007), Metcalfe, J. (ed.) (1994),  Necka,E.; J. Orzechowski (2005)).  It is often  described as a cognitive device that aims to "know(ing) about knowing, that is, a cognitive function to distinguish what one knows about one's own cognitive abilities, states of knowledge, and actual performance form the cognitive abilities, states of knowledge and performance per se" (Koren, D.; L.J.Seidman; M. Goldsmith; P.H. Harvey (2006)  p. 313). The case of our example of William Tell's hesitation about his shooting is relevantly a case of metacognition, as "metacognitive processes are required for decision making, troubleshooting, strategy selection and performance of non-routine actions" (Fernández-Duque, D.; J.A.Baird.; M.A. Posner (2000) p. 289).

Metacognition is probably the best candidate to exemplify one of the functions of reflective knowledge in cognitive systems. Although reflective knowledge can be characterized as a higher intellectual and conceptual process, surely it is also based on more basic cognitive mechanisms than metacognition is. And the working of these devices can help us to clarify our question of integration requirement. As a functional system or ability, it probably is already present in some animals other than human beings. Some studies with simians, dolphins, and even rats have shown that many animals refrain  from acting when the cognitive conditions of a formerly known task become harder(Smith, D. (2005), ;  Foote, A.L.; J.D. Crystal (2007)).  The alleged experiments do not allow us to clearly conclude that certain animals are gifted with self-consciousness, but they are sufficiently expressive to ascribe some degree of metacognition to them. These are not news for Sosa's concept of reflective knowledge, for, as for him, "a minimal degree of perspectival, reflective knowledge" can be possessed by animals (Sosa (2009) p 239). Notice however that behaviour of refraining seems to indicate the existence of strong links between information evaluation and control of acts.  The links do not amount to confusing acts and beliefs in animals (among other things because they lack intentions both in actions and in beliefs), but, however, the issue is that the control is based here in an appreciative feeling of

ignorance. It does not matter here if alternative possibilities of action are considered by the animal.

Another interesting characteristic of metacognition apparently supporting Sosa's views is that it does not necessarily involve metarepresentation (Proust (2007). Metarepresentation is a higher-order process which supposes to have reached previously advanced steps in cognitive development (children reach it at the age of three and half) for metarepresentation requires open consciousness and deliberative stances. If metacognition were equivalent to metarepresentation, then Sosa could legitimately argue that he is postulating a more basic metacompetence or skill that is not in such a high layer of cognitive processing (although a full reflective stage of knowledge is possible). Nevertheless metacognition is perhaps a good candidate to exemplify a skill to evaluate the own cognitive perspective without supposing the status of assertion or deliberation to act. Furthermore, metacognition draws in structurally important functions as they are "theory of mind" (or simulation ability), memory retrieval, transmission of learning, executive control, etc. It amounts then to a structural component of any cognitive task in which the epistemic quality was involved. In fact, the lack of metacognitive skills is a symptom of mental disorder. This metacognitive deficit is called *Anosognosia*, and it is observed in patients suffering schizophrenia, Alzheimer's and others kinds of dementia (Cosentino, J.; Y.Stern (2005)). Anosognosia is a sort of unawareness of one's own cognitive and functional impairment. Patients with this symptom are invited to discover their deficit by indirect means, because they show denials of the deficit and a very poor insight of their cognitive status.

According to the hypothesis of most authors on the subject, metacognition is a faculty composed of two more basic functions: *monitoring* the cognitive status, and *controlling* the ulterior processing of the given information. For example, refraining from action is one of the possible outputs of the control mechanism (but all the same, it could also be refraining from believing). Monitoring and control do not constitute two separate functions aimed at belief and action respectively, but two necessary aspects of any meta-competence, even though they operate in conscious levels that are not too explicitly represented. Now then, this double aspect could be considered as establishing some symmetry between belief and action at the basic level of their production: reflective knowledge would require both monitoring and control, the same as reflective action, even though an explicit deliberation would not take part in the process.

Maybe the case that one would argue against the strategy of resorting to empirical findings when the issue is in fact of an exclusive conceptual nature. I will not quarrel about the required degree of naturalistic considerations in epistemology. My point is that in spite of the fact that most basic levels of cognitive faculties were working, a problem of lack of integration could be addressed (and consequently, much more in higher epistemological levels). In this way, metacognition furnishes us with a framework where the integration problem can be formulated without reaching so higher levels, sorting out beliefs from assertions.

## Integration and first-person perspective

The integration problem is one which arises when we the personal level becomes the prime focus. By personal level I do not mean here necessarily highest levels of deliberative stances, but a more constitutive stage where the overall equilibrium of the system is the issue. The idea is that a person is healthy constituted when her mental faculties are reasonably functioning and acceptably coordinated (obviously, disabilities, local malfunctioning, etc., cannot be discarded). In this sense, metacognition is a mechanism that can accomplish its task only in a systemic way; that is, that it works insofar as other cognitive mechanisms coherently work as well. The coherence required for this integration is unlike mere logical or informational coherence; actually we would say that the system demands a sort of metacoherence. Furthermore, in the task of knowing, working with a higher degree of harmoniously systemic coordination is required in order to engage the overall cognitive system in such a demanding task. This requirement derives from the very nature of knowledge: for knowledge in virtue epistemology is a kind of achievement from the agent's character. The conceptual point here is that this requirement claims a first-person perspective. We contend that systemic coherence, as described from a third-person perspective, will not be sufficient. Something in the working system is required to ensure that it is the agent, and not a mere part of him, who is engaged in the task. In this regard, metacognition, when involved in reflective knowledge, always entails first-person perspective on the own cognitive processing. Hence, metacoherence, integration and first-person perspective keep going or fall together.

Let's now examine a plausible argument against the engagement of personal levels:

(1) A personal level is relevant insofar as human autonomy is in question

(2) Autonomy requires deliberation between alternative possibilities

(3) Intentional action supposes alternative possibilities

(4) Belief formation does not necessarily suppose alternative possibilities

(5) Therefore intentional action claims a personal level but belief formation does not.

This argument assumes that reflective knowledge is part of the belief formation process (at least of the apt belief formation process). Consequently, reflective knowledge, as implicated merely in belief formation to difference of more demanding levels (as for instance assertions), would not require being located at a personal level.

I don't want to follow the course of the controversy over the voluntariness of belief, even though some objections could be addressed against the steps (2) and (4) from a Frankfurtian standpoint, that is to say against alternative possibilities as a condition for autonomy. And still the example of metacognition opens up another different line of rejoinder. The point is yet about the cognitive rightness of premise (2). Seemingly, this premise considers as autonomous only those systems capable of explicit, discursive, and conscious deliberation between alternative possibilities. Certainly it is not difficult to grant that conscious deliberation is sufficient for autonomy. However the question is if there can be autonomous systems in lower steps of cognitive development. Regarding the symmetry between belief and action insofar as an integrated system is required, you need not consider alternative possibilities. Take into account for instance cases of perception or spontaneous and yet intentional reactions (let's assume that they are reactive attitudes): they are not result from deliberative processes, but still they are intentionally higher forms of behaviour.

Accordingly since, although metacognition is a part of an assembled bundle of faculties that compounds an agent, nevertheless when metacognition works it indicates that such assembly and the subject are the same agent. In cases of knowledge, the knower is self-ascribing a competence. Notwithstanding such self-ascription needn´t be processed in higher deliberative layers. As formerly said, metacognition has the functions of monitoring and controlling the cognitive status while the organism confronts a particular task, but such function can be performed in a quite automatically mode. It is out of the question if this double function was performed by a single mechanism or, by contrast, required two different ones. The relevant thing is that

metacognition performs the twofold function as part of a singular but compound cognitive task. Now then, notice that in spite of the fact that monitoring can be considered as a more passive engagement of the organism, nevertheless the case becomes different for control, as it spontaneously and actively engages the overall organism in the production of a right outcome. The function of metacognition would be then to evaluate the ability of the organism to deal with a difficult cognitive task. It does not matter if deliberation precedes this function or not. The system works properly as far as it is able to detect a state of dangerous ignorance, that is, when it detects that the available knowledge is not sufficient.

Let's compare now this function with the alleged case of William Tell at the very moment of shooting the apple on his son's head (instead of shooting his son). In the case (a) Tell thoroughly deliberates the action and consciously decides that the shooting will be safe. In the case (b) Tell does not consider the question and simply shoots. He is very confident in his skill to hit the target. In both cases reflective knowledge (at least metacognition) is involved, but in the case (b) awareness of the situation does not necessarily imply an explicit discursive deliberation. Tell will be trustworthy insofar as his control system works and will be able to refrain from shooting when the accuracy is jeopardized.

Therefore, the premise (2) can be reformulated in broader terms:

(2) Autonomy involves enough control function to refrain from following the process.

Rejecting the premise(2), and substituting it by a less demanding and more general capacity of behaviour refraining, means that symmetry between actions and beliefs there can be re-established, something that is stronger than an analogy, regarding the whole system engagement in performing.

My second point is that only a first-person perspective guarantees this required agential engagement in cognitive tasks. Since, as for instance blindness disorder shows, it is possible that a perceptual system can work properly at some level, and yet the suffering patient is unable to say that he sees the object (an object that actually he is able to correctly manipulate). Analogously, although a patient suffering from Anosognosia could be described from a third- person perspective as properly working in a first level of knowledge (take, for example, remembering something), I do not understand how metacognition functioning could be described in a third-person perspective.

It happens that in the case of metacognition normal subjects report having a "feeling of knowing" when the system is working well. Consider, for example, the tip of the tongue cases when, after some efforts to recall a name, one is able to retrieve it. Then a feeling of knowing overwhelms us, and this feeling signals that that the task is accomplished. However the philosophical significance of this feeling is not easy to assess. On the one hand, this feeling of knowing can be considered as a symptom of a proper working of metacognitive skills. One who feels he knows seemingly reaches a better stage that a mere belief occurring him. Even though I do not claim that in epistemology necessarily requires this feeling to account for reflective knowledge (perhaps it is a mere psychological indicator), yet it tell us that something is happening, namely, that the agent experiences the knowledge ownership. Possibly some epistemologists worried only by the objective functioning of a faculty would argue that the option between first or third person perspectives makes no difference. Nevertheless the symmetry between belief and action is relevant here. Compare for instance the case of an aeroplane guided by the pilot with, by contrast, another operated by some automatic system in a remote control tower. In spite of that fact that both ways could be described as functionally equivalent, there is nevertheless a noteworthy difference between the two. Intuitively, the personal perspective matters here. Now then, does it matter from an epistemological point of view?

To take stock of our steps, observe this increasingly significant list: "metacognition", "feel of knowing", "first-person perspective", and "reflective stance". Each stage means a higher degree of integration of the system regarding its agential status. Likely, all of them can occur at diverse points of cognitive processes. However, the point is that a first-person perspective characterizes a certain form of integrating the cognitive and executive systems, i.e., an agential shaping of behaviour in such a way that the subject is the owner of her outcomes. Even though the task does not reach a full deliberative and conscious status, still the agency is supported by the self-confidence agents have in their capacity to attain the goal. The personal level is sufficiently expressed in the open intentional action, as assertive discourse is, but this level is also required in several tasks which can be performed in more implicit ways. In the case of knowledge, the personal level is required for a knowing subject when self-ascribing an epistemic competence. Observe now that the implicit/explicit divide does not equals to the personal/sub-personal divide. The personal level is required when meta-coherence is at stake. William Tell, before his dreadful experience demanding maximum self-trust, exemplifies this requirement of full integration for a system. Although a full explicit reflective stance can be a further stage for an agent who

deliberates in an open as well as in internal forum, nevertheless we should distinguish between the coherentist claim that full consciousness is required for a full justification and our view about the first person engagement as necessary to show an integrate agent. Meta-coherence, in our sense, is a structural property of the personal level, and it does not imply a more Kantian approach. As is well-known, this approach demands in addition to obeying a rule, also obeying because of the concept of the rule. Less-than-Kantian approaches, such as mine, simply require a good integration for a sufficient agency, an epistemic agency, in this case.

**Taking epistemic risks and epistemic responsibility: the agential model of knowing**

The symmetry between belief and action, which becomes apparent at the personal level requirement for integration, has to do with the status of being an agent. We contend that a subject enjoys an integrated status nature when she behaves as an agent both in the course of knowledge as well in action (observe that the inverse could fail: the integration is only a necessary condition for agency). Knowledge and action denote the presence of agent in the scene of theoretical or practical processes. The reason is that both share the common nature of agency as the distinctive trait expressing the human autonomy. It means a capacity to self- determine a particular state (in the world, as well in the own subject) as resulting from competencies the subject possesses as being a subject.

Hence knowledge signifies an expression of a subject able to determine a peculiar mental state of her own: the state of believing as the product of the own competencies. It might be that this judgment takes the form of an internal assertion as well as an open avowal. In such cases, the outcome has been preceded by an explicit judgment, but it needs not to be so for representing an agent's self-expression. It suffices with the well-integrated personal making up the mind even in an implicit way. Certainly, very often beliefs are not products of such self-determination, for they come from perceptual, emotional or automatic cognitive mechanisms. Possibly, these beliefs deliver information indeed, and even they can reach the status of knowledge. But in these cases, the merit is attributed to those component faculties and not to the subject itself. Like in knowledge, in the case of action there are forms of behaviour that do not attain a full intentional status: they can be forms of behaviour resulting from skills or ways of knowing-how that do not call for a full agential engagement. Of course this mechanical behaviour constitutes a very common way of acting in daily life. The agent

needs not always to express itself as an agent unless relevant circumstances demanded such so higher manifestation. The personal engagement is needed when the agent's position it at stake. Sometimes it claims an explicit judgment in the form of an assertion or decision, but the point is that the need is activated when the agent must calibrate her possibility to attain the goal. When such degree of control occurs she becomes responsible and praiseworthy.

From a normative perspective, it is essential to consider such state-determinations as attaining a certain normative level of success. To be a success is a property dependent of a broader context in the life span of a creature. Dewey explained in *Art and Nature* that "ends", in this normative sense, are characteristic points of special relevance to life. For example, births or deaths are indifferent points in the course of physical chains of causes and effects, but it is not difficult to understand that, in the framework of the life of living beings, these points are of an exceptional relevance. In the same sense, self- determinations of mental as well as physical states in the agential way are of normative relevance when they are also relevant points that can be qualified as successes[5].

To resume the discursive line on reflective knowledge, we must notice knowledge as an end is some more valuable than true belief. And that this difference matters for agency. To reach knowledge is some more than to get a true belief: to become knowledge, a true belief must reach the status of an achievement due to the competence of the agent. Such degree of accomplishment claims a subject worthy of it. The significant thing is that such achievement does not means merely to arrive at the proposed end, but to attain these ends because the active engaging of a well-integrated subject in the task of knowing.

It is telling that this way of considering knowledge gives us one of the best justifications for Virtue Theory. Since Virtue Theory Subject provides a causal background to the agential concept of subject. In other traditions, as for instance in a more the intellectualist tradition, the subject becomes a very weird metaphysical notion: although is considered is a substance able to have spontaneity, still this spontaneity seems to come from nowhere. In contrast with this, Virtue Theory proposes reflective competencies as the way in which an agent becomes responsible for

---

[5] Dewey profoundly observes: "It is not easy to distinguish between ends, as *de facto* endings, and ends as fulfilments, and at the same time to bear in mind the connection of the latter with the former. We respond so directly to some objects in experience with intent to preserve and perpetuate them that it is difficult to keep the conception of a thing as terminus free from the element of deliberate choice and endeavour; when we think of it or discourse about if, we introduce connection" Dewey (1929) p 111.

the quality of her epistemic view; and hence how she becomes a sort of causal singularity in the universe. This status is attained because she controls the risk of believing, that is to say, because she dares to believe given her own epistemic position in the world. This capacity to take epistemic risks converts the agent in an autonomous believer, namely, in an epistemic agent.

But, how is that to assume epistemic risks provides a criterion of epistemic agency? The answer is that the decisive aim of agency would be to grant that a particular move makes sense for the subject. David Velleman has convincingly argued that the human way of acting is ordered to self-understand the course of action one is choosing. Consider this way of explaining how an agent get a reason for deciding:

> "I believe that the process of improvisational self-enactment constitutes practical reasoning, the process of choosing an action on the basis of reasons. Why do I think that the self-enactor chooses his action? Because it his idea, which he puts into action in preference to other ideas that he might have enacted, if this one hadn't made more sense. Why do I think that he chooses for reasons? Because he chooses his action in light of a *rationale* for it, which consist in consideration in light of which the action makes sense" (Velleman (2009), p.18)

We could imagine in contrast with this that taking risks is, from the agential point of view, like mechanically trespassing some threshold of probability. But presumably any mechanical engine, lacking still the status of agent, could take this chance. However a reflective agent is one for whom make sense to dare to certain belief or decision. The reflective stance is not, then, a sort of faculty added sequentially to former springs of beliefs or decision. The previous discussion on meta-cognition showed us that this faculty can be possessed by a long class of animal species. If agential way is a human feature, then it must consist in something more: our contention has been that it consists in being the expression of the subject in the scene of knowledge and action. And such appearance only occurs when a well-integrated agent takes a course of action that makes sense for him. What is the risk she takes? Such risk is for her the chance of not being able and failing to achieve the aim, given her own cognitive resources and competences, and given the circumstances of the undertaken task. The risk of failure is a risk the agent must calibrate given her self-confidence, but also, her objective capacity to undertake the task.

As William Tell before his tragic performance, an epistemic agent must decide to accept a belief that can be of an indeterminate relevance for other aims, theoretical or practical, but that is of a constitutively centrality to the task of knowing. Consider

dramatic scenarios like doctors answering the fearful patients: "Do you know that it is cancer?" The degree of engagement and attention the epistemic agent devote to the quality of his epistemic perspective can depend on the demands of the involved question, but in any case, the agent must calibrate the own powers to make an avowal. At what extent the answer constitutes or not an epistemic achievement it is something that would not depend on the ethical or practical relevance of the question, but the virtuous character of an agent which dares to believe.

## References

Breyer, D.; J. Greco (2008) "Cognitive Integration and the Ownership of Belief: Response to Berneker". *Philosophy and Phenomenological Research LXXVI/1* 173-184.

Cosentino, J.; Y.Stern (2005) "Metacognitive Theory and Assessment in Dementia. Do we Recognize our Areas of Weakness?" *Journal of the International Neuropsychological Society 11*: 910-919

Dewey, J. (1929) "Nature, ends an Histories", en *Experience and Nature*. New York: Dover 1959

Fernández-Duque, D.; J.A.Baird.; M.A. Posner (2000) "Executive attention and Metacognitive Regulation" *Consciousness and Cognition 9*: 288-307, p. 289.

Foote, A.L.; J.D. Crystal (2007) "Metacognition in Rats" *Current Biology 17/6*: 551-55

Koren, D.; L.J.Seidman; M. Goldsmith; P.H. Harvey (2006) " Real—World Cognitive and Metacognitive in Schizophrenia a New Approach for Measuring (and Remediating) More "Right –Stuff" *Schizophrenia Bulletin 32* 310-326.

Koriat, A. (2000) "The Feeling of Knowing: some Metatheoretical Implications for Consciousness and Control" *Consciousness and Cognition 9*:149-171.

Metcalfe, J.(ed.) (1994) *Metacognition: Knowing about Knowing*. Cambridge MA: MIT Press.

Necka,E.; J. Orzechowski (2005) "Higher-Order Cognition and Intelligence", in Stenberg, R.; J. Pretz (eds) (2005) *Cognition and Intelligence*. Cambridge: Cambridge University Press.

Proust, J. (2007) "Metacognition and Metarepresentation: is a Self-directed Theory of Mind a Precondition of Metacognition" *Synthese 159*: 271-295.

Smith, D. (2005) "Studies on Uncertainty Monitoring and Metacognition in Animals and Human" in H.J. Terrace, J. Metcalfe (eds) *The Missing Link in Cognition. Origins of Self-Reflective Consciousness*. Oxford: Oxford University Press.

Sosa, E. (2003) "Knowledge, Animal and Reflective: a Reply to Michael Williams" *Proceedings of the Aristotelian Society. Supplementary Volume*. 113-30.

Sosa, E. (2007) *A Virtue Epistemology. Apt Belief and Reflective Knowledge*. Oxford: Oxford University Press.

Sosa, E. (2009) Reflective Knowledge. Apt Belief and Reflective Knowledge, Volume II. Oxford: Oxford University Press.

Sosa (2009 b) "Respuestas a mis comentadores", *Teorema* XXVIII/1, 2009, pp. 112-124.

Velleman, D.(2009) *How We Get Along*. Cambridge: Cambridge University Press.

Williamson, Th. (2000) *Knowledge and Its Limits*. Cambridge: Cambridge University Press.