

Synthese (2007) 159:235–251 DOI 10.1007/s11229-007-9206-5

Inside loops: developmental premises of self-ascriptions

Radu J. Bogdan

Published online: 26 July 2007

© Springer Science+Business Media B.V. 2007

Abstract Self-ascriptions of thoughts and attitudes depend on a sense of the intentionality of one's own mental states, which develops later than, and independently of, the sense of the intentionality of the thoughts and attitudes of others. This sense of the self-intentionality of one's own mental states grows initially out of executive developments that enable one to simulate one's own actions and perceptions, as genuine off-line thoughts, and to regulate such simulations.

Keywords Self-ascriptions · Thoughts · Attitudes · Self-regulation · Simulation

1 Introduction

This paper develops an earlier argument (Bogdan 2005a) to the effect that self-ascriptions of thoughts and attitudes require a conscious sense of their intentionality. The background assumption is that one's grasp of the self-intentionality of one's body and mind in general depends on the sort of bodily and mental activities and states that one self-regulates and on the manner of such self-regulation. The conscious sense of self-intentionality required for self-ascriptions of thoughts and attitudes must be distinguished from other senses of self-intentionality, such as bodily or sensorimotor, and also from the resources required for the ascriptions of the thoughts and attitudes of others. In order to mark this latter distinction, Sect. 2 surveys the difference between the child's naive psychology before and after the age of 4 and suggests that, for self-ascriptions of thoughts and attitudes, the difference has less to do with developments in naive psychology and more with dramatic changes in the child's intramental self-regulation. This truth may be obscured by nonintentional self-ascriptions, which are based on or directed at one's own experiences. Experiential self-ascriptions (as I will call

them) need not presuppose a sense of self-intentionality. With these critical clarifications behind us, I will turn to the main constructive argument.

Section 3 introduces the notion of the self-regulation and surveys several of its forms in order to identify the specific form responsible for a conscious sense of the intentionality of one's own thoughts as the key developmental premise for self-ascriptions of attitudes, such as beliefs or intentions. According to Sect. 4, the executive developments after the age of 4 generate new capacities to simulate one's own actions and perceptions and, crucially for self-ascriptions, call for the self-regulation of these simulations. Such self-regulated simulations are the first ontogenetic instances of thoughts the child is aware of as her own world-directed thoughts. Their self-regulation is carried out by an executive metaseantics, construed as a capacity to monitor and control one's own simulative thoughts in terms of their representational relations. It is an intramental self-regulation that provides a sense of the intentionality of one's own thoughts. Section 5 explains why the development of intramental self-regulation is late, thus causing a correspondingly late development of self-ascriptions. Section 6 reviews the argument and points to some implications.

Before proceeding, a few words about terminology. For brevity, I call the ascriptions of one's own thoughts and attitudes 'self-ascriptions' and those aimed at the thoughts and attitudes of others 'other-ascriptions'. I will be concerned *only* with the self-ascriptions that register the intentionality (in the philosophical sense) or aboutness of the thoughts and attitudes involved. The ability to register the intentionality of one's own bodily and mental states and activities in general amounts to (what I will call) a 'sense of self-intentionality'—where the 'sense' may cover a variety of mechanisms whose operation can, but need not, be conscious.

Since I will be focusing mostly on thoughts, I take the terminological liberty of using the adjective 'thoughtful' to mean 'pertaining to or having to do with thoughts'. It is standard practice in recent philosophy to construe thoughts as representing the contents of attitudes, such as desires and beliefs. On this construal, one can self-ascribe attitudes only when one can self-ascribe thoughts. And one can self-ascribe thoughts only when one understands them as intentionally related to what they represent. The focus below is on what it takes to develop the capacity to self-ascribe thoughts as a prerequisite for the self-ascription of attitudes. The basic idea will be that only simulative thoughts, subject to self-regulation, afford a conscious sense of their intentionality, thus making self-ascriptions of thoughts and hence attitudes possible.

2 Blind to one's own intentional mind

My reading of the developmental data is that until around the age of 3–4 the child is mostly concerned with the outside world, for obvious reasons of social and cultural adaptation. As a result, the child's naive psychology is almost exclusively oriented toward *other* people, as illustrated by such dominant practices as gaze following, joint attention, and imitation (Hobson 1993; Perner 1991; Tomasello 1999). The spontaneity of these practices and their well paced development suggest adaptive responses to the pressures of social interaction, communication and cultural assimilation during early childhood (Bogdan 1997, 2000, 2003, 2005a).

There are good reasons to think that, although finely attuned to the intentionality of other minds, the young child is generally blind to the intentionality of her own mind. An indirect but convincing reason is the absence in early childhood of metacognition, introspection, and autobiographical memory—all mental activities that require an awareness of the intentionality of one's own mental states (Bjorklund 2000; Flavell et al. 1995; Nelson 1996). A more direct reason concerns the developmental and cognitive *asymmetries* between self-ascriptions and other-ascriptions that register the intentionality of the attitudes involved. The case of belief is symptomatic. The recognition of the false beliefs of others, around 3–4, seems to precede the recognition of one's own false beliefs, a year or two later (Astington and Gopnik 1988; Perner 1991). This asymmetry is further confirmed by data showing that children even younger than 3 seem to recognize, implicitly and preverbally, the false beliefs of others (but not their own) (Clements and Perner 1994). The belief asymmetry and other attitudinal asymmetries are not widely accepted in the developmental literature and new data force constant revisions of earlier accounts. Nevertheless, I think that if we widen the theoretical horizon, we will find good reasons to take these asymmetries seriously, as I argue below and also did so elsewhere (Bogdan 1997, 2003, 2005a).

Cognitively, there are major differences between the other- and self-ascriptions, insofar as the evidence for their intentionality is concerned. The intentionality of other-attitudes is rather easily detectible in the visible bodily features and behaviors of others, whereas the intentionality of self-attitudes is not visible, not in any obvious and conscious way, in one's own behavior or mental activity. The point here is not that children learn from visual and other forms of experience to recognize the intentionality of their own attitudes and that of others but rather that the recognitional abilities

involved have different tasks and maturation schedules, and recruit different resources.

The differences between intentionality-sensitive self-ascriptions and other-ascriptions derive from a more general gap between the child's naive psychology before and after the age of 4. It is not just the much-discussed turn of other-ascriptions to metarepresentation around 4, perhaps best illustrated by the false-belief test. There are also dramatic changes, mostly neuropsychological and executive in nature, which reorient the child's naive psychology toward complex mental states of others and of self, with imagined, counterfactual or abstract contents, and render the operation of naive psychology increasingly free from the perceptual here-and-now and from current motivation, and increasingly off-line and inferential. These changes include inhibition, a remarkable spurt in the growth and functionality of the prefrontal cortex, and a massive shift from a simpler and mostly other-oriented naive psychology, based largely on the left hemisphere, to a much more sophisticated successor that draws increasingly on the executive powers of the prefrontal cortex and the interconnectivity of the right hemisphere (Corcoran 2000; Diamond 2001). As we shall see, an important outcome of these neuropsychological developments is a conscious sense of the intentionality of one's own thoughtful mind.

The young child's blindness to the intentionality of her own thoughts should be carefully distinguished from the self-consciousness of her *experiential* mind and from her awareness of *lower-level* forms of self-intentionality. Consider the experiential mind first. The young child is surely aware of such experiences as feelings, emotions or desires, and this awareness yields a variety of explicit experiential self-ascriptions, thought about or even verbalized in terms of 'I feel pain' or 'I am sad' or 'I want an orange' or the like. One popular view in philosophy and developmental psychology construes the self-evidence of one's experiences as *sufficient* for intentionality-sensitive self-ascriptions of thoughts and attitudes. This experientialist view (if I may so label it) underpins the introspective version of simulation, which finds that self-ascriptions develop earlier than other-ascriptions (Goldman 1993; Harris 2000). I criticized this view elsewhere (Bogdan 1997, pp. 243–249, 2005a). Besides the fact that young children do not seem to introspect (Flavell et al. 1995), a careful reading of standard experiments reveals that children are frequently asked to report on their present or recent past *experiences* (perceptions, memories, feelings) without an *independent* determination of whether these experiences are actually recognized by children as intentionally directed at their contents (Astington and Gopnik 1988; Goldman 1993; Harris 2000). In such experiments a child

may be asked, for example, whether she saw a particular event or object X. Her positive answer may indicate visual familiarity with X or the presence of a memory image of X but need not entail an independent recognition of a past perceptual or attitudinal *relation* to X. This distinction is close to that between episodic memory, which requires experiential familiarity with a past event X, and autobiographical memory, which also requires awareness of self in *relation* to X. Not surprisingly, on my reading, autobiographical memory develops as late as do intentionality-sensitive self-ascriptions of attitudes and involves some of the same executive novelties (Bogdan 2005a).

Besides these empirical distinctions, there is a larger conceptual issue at stake here. An assumption of the experientalist view seems to be that if the young child knows the type of inner experience involved (say, desire) and its content (what is desired), and also knows that a particular sort of desire experience and a particular sort of content are always associated, then the child knows *that* the desire is directed at its target and therefore recognizes its intentionality. Thus, suppose one asks a young child what she desires, and she says an orange. And suppose further that, under further questioning, one also determines that the child's answer draws on her recognition of a past experience (say, taste) associated with oranges. Doesn't this show that the child recognizes her desire *as* a mental state *intentionally related* to oranges? If the child recognizes that she has a desire *for* oranges, isn't this *ipso facto* a recognition of the desire being *about* oranges (J. Perner, personal communication)? A related point can be made about memory (or any other inner experience associated with some intentional state): isn't the child's memory of an event X an experience the child recognizes as representing X (B. Malle, personal communication)? If the answer is yes, then the young child can be credited with an experience-based recognition of the intentionality of some self-attitude (e.g., desire or memory).

This is an answer I am determined to resist, for further reasons, in addition to those noted earlier. First, a conceptual point: having an experience of X, associated with an attitude about X, does not entail a recognition that the experienced attitude is about X (as in the orange example above). The experience may signal the attitude type (desire or memory) and its target (orange or some past event, respectively) but not the intentional relation between the two. The target of the attitude is transparent *in* its experience but not *in relation* to the experience. This is why experiencing an attitude and its target is not yet having the *concept* of the attitude, insofar as the concept factors in the intentional relation to the target. Young children have beliefs and memories, distinguish them experientially

from other mental types (such as desires or perceptions), know what their contents are (what they believe or remember) but do not, until a few years later, develop the concepts of such intentional types, qua attitudes, and hence the recognition of their intentionality.

It may also help to look at the matter from a wider comparative angle as well. If regular associations between one's own experienced attitudes and their targets were yielding a sense of the self-intentionality of the attitudes, then animals and very young children that can associate the two ought to be able to self-ascribe attitudes in some form. But they don't. Regular association is not enough. A version of this *reductio* applies to animals (say, dogs and parrots) that learn to associate sound or gestural experiences with targets in the world. If such learned associations yielded a sense of the intentionality of the experiences involved, these animals would understand word or gesture reference as a relation. But they don't. Indeed, the first words of human children, around 6 months, are learned by association, *without* a sense that they are intended to refer to their targets. That early understanding of words is radically different from, and significantly inferior to, the words learned 10 or so months later through attention shared with adults. Only the latter reveals to the child the intended referentiality of words (Bogdan 2008; Tomasello 2003). And, finally, if naive psychology itself were acquired by learning regular associations between experienced attitudes and their contents, we should find it throughout the animal world. But we don't.

Perhaps the most important argument—and certainly the most relevant here—against the notion that the experience of attitudes suffices to produce a conscious sense of their self-intentionality derives from the main conjecture of this paper, which is that such a sense depends essentially on what world-directed or intentional aspects of one's mind one self-regulates and how. Animals and young children do not self-regulate the intentional relations of their own thoughts. But they do self-regulate the intentional relations of their sensorimotor and perceptual experiences, which is why they have a sense, not necessarily conscious, of their sensorimotor or perceptual self-intentionality. This is also why, I think, the young child who recognizes her desire for an orange has at best a conscious sense of the gustatory-motor intentionality (so to speak) of her experienced desire (i.e., of what she is poised to do and feel, given the anticipated taste) rather than a conscious sense of her desire as an attitude related to its target. This last point deserves some elaboration.

Awareness of selfhood and self-intentionality operates at different levels. One may have a sense of a *thoughtful* self, of being the author of

one's own thoughts (compromised in schizophrenia, for example), which is distinct from a sense of a *sensorimotor* self, which is one's sense of ownership of and control over one's sensorimotor activities (normally not impaired in schizophrenia). Young children, and most animals, have a sense of sensorimotor but not of thoughtful self. I return below to this variety of selves.

A similar sensorimotor/perceptual/thoughtful distinction works for self-intentionality. Young children and most animals have a sense of their own sensorimotor or perceptual intentionality—that is, a sense that their sensorily or visually guided actions are directed toward external targets. For example, they can self-regulate their sensory approach to or perceptual perspective on a target (Russell 1996). But, I suggest, only older children develop a sense of their own *thoughtful* intentionality—that is, a recognition that their own thoughts represent something. This recognition is crucial for the self-ascriptions of attitudes, as is the recognition of thoughtful selfhood.

To sum up, so far, the young child's thoughtful mind is not yet conscious of itself *as intentional*, unlike her experiential mind, which is largely conscious of its own experiences. The young child may be aware of thoughts as experiences but not as intentional representations. The theoretical and experimental challenge is not to mistake the young child's experiential self-ascriptions for self-ascriptions that are intentionality-sensitive and therefore not to think that the young child's ability to self-ascribe attitudes is simpler and easier than it actually is, or that it emerges earlier than it actually does. The question, then, is what it takes to understand one's own thoughts and hence attitudes as intentional. This is the question explored next.

3 Self-regulated minds

The guiding idea is that a conscious sense of the intentionality of one's own thoughts, as a prerequisite of self-ascriptions of attitudes, depends on thoughts becoming targets of one's mental agency in general and of self-regulation in particular. This idea, initially far removed from the domain of naive psychology, takes us (briefly) to the basics of life—goal-directedness and self-regulation.

The background premise is simple and familiar. To be successful in life, organisms must monitor and control or, in a hyphenated word, self-regulate the means by which they reach their goals. Organisms have a variety of processes and activities that are self-regulated. In most species, these processes and activities are physiological, sensorimotor, perceptual,

and behavioral. Some species, mostly primate, also self-regulate their higher mental states and activities. The ability to regulate the world-relatedness of one's mental and behavioral states and activities is an intrinsic part of being intentionally directed at the world. This self-regulatory ability contains the seeds of (what we may call, quite liberally) an *executive metasemantics*—metasemantics, because the organism has some way of registering and checking the intentional relations (informational, referential, representational, meaning-like, as in language) it has to targets in the world; and executive, because of the regulatory nature of the ability itself. When such self-regulation meets some further conditions, the executive metasemantics also generates a *conscious* sense of self-intentionality in some dominant modality. When thoughts themselves become targets of conscious self-regulation, a developmental platform emerges for the self-ascriptions of attitudes. This is the idea developed in the remainder of this paper. I begin with some basic notions of self and regulation, which I need for my analysis.

To exercise self-regulation, an organism needs a capacity (as mechanism, procedure, scheme, or explicit representation) to register that it is related to some target and also that it is on target or not, or how close to or far from it, in order to guide and correct its representations and actions. In evolving such a capacity for self-regulation, an organism faces two evolutionary problems, whose solutions constitute the biological pillars for a sense of self and of self-intentionality, respectively.

3.1 Senses of self

The first problem is to distinguish between the information provided by internal resources and the information caused by the outside world. Only in this way can an organism distinguish what it does from what the world does and thus have a sense of authorship or ownership of its states and activities. This ability is not unlike that of the immune system that distinguishes between its body and intruders, based on some sense of the bodily self. Or consider vision, for a more detailed example. The movement of an object in the outside world results in a corresponding movement of the image of the object on the retina. The latter is perceived by the brain as actual movement of the object. But when the eyes are moved voluntarily, there is also self-initiated movement of the image on the retina, which is not perceived as movement of an external object. Why not? Because the brain makes an output (or efference or feedforward) copy of its instructions to the eye muscles, and the output copy is translated into the expected retinal

movement. If the real retinal movement is the same as the expected one, no outside movement is perceived. But any movement in the outside world will result in movement of the retinal image which has no counterpart in the output copy, and therefore will be perceived by the brain as the movement of an object in the outside world.

This analysis of visual self-regulation has been known since the 1950's and has later been extended to motor self-regulation as well (Jeannerod 1997; Wolpert et al. 1998). For our purposes, it is worth noting how the self-world distinction in vision is achieved implicitly, procedurally and unconsciously by way of how the self-regulatory machinery of the visual brain works. I think the same is likely to be true of other modalities of mental activity—perceptual, conative, memorial, thoughtful, and so on. In particular, we should not be surprised to learn that the ownership of the thoughts (and possibly memories) of the young child (before 3–4), which seem to be mostly perceptual or imagistic, is handled and registered more or less along the same lines as the child's visual images. Let us call these kinds of thoughts *me-thoughts*. Such thoughts (more or less) happen to the child, caused mostly by perceptions, memories, habits and drives, yet they are distinguished from what is going on in the outside world and also distinguished from their proximal causes.

After the 3–4 interval, the child can be said to graduate to *I-thoughts*. This new kind of thoughts reflects not only a fuller grasp of the personal pronoun but also (and relatedly) a multiplex mind that can initiate, operate with and control multiple thoughts, thanks to the development of inhibition, the significant growth of the prefrontal cortex and the increased interconnectivity of the right hemisphere (Diamond 2001). I-thoughts are self-initiated thoughts that are not only registered as one's own (ownership) but as initiated by oneself (authorship).

Yet the self need not be—and is likely not to be—separately, explicitly and consciously identified and registered in one's I-thoughts. That would be the development of *self-thoughts*, which occurs later, perhaps around the age of 6–7, as indicated by the emergence of introspection, narrative autobiographical memory and reflexive thinking (Flavell et al. 1995; Nelson 1996; also Bogdan 2000). The idea here is not that there is a self somewhere in the mind, which can be targeted for explicit representation (an idea vigorously and plausibly denied by Hume and many other philosophers), but rather that novel self-regulatory mechanisms generate a novel form of recognition of the authorship of one's own thoughts. With self-thoughts one is thinking explicitly about one's present, past or future self-initiated and self-owned I-thoughts. Self-ascriptions of

thoughts and hence attitudes become possible only with the development of I-thoughts and self-thoughts.

Yet having me-thoughts or I-thoughts or self-thoughts or indeed any other sensorimotor or mental state need not lead to a sense of their intentionality. The latter is the solution to the second major evolutionary problem faced by any active and autonomously goal-directed organism, which is the problem of registering its own intentionality in some world-related modality. There are several versions of the solution to this second problem. They are worth distinguishing, briefly, in order not only to identify the one involved in attitudinal self-ascriptions but also to show its executive credentials and pedigree.

3.2 Senses of self-intentionality

As with various selves, a sense of an organism's target-relatedness or intentionality originates in its machinery for self-regulation. In the sensorimotor modality, this machinery functions implicitly when, say, feedback from an action is compared with and adjusted to preset expectations or memories of prior actions. It is an implicit work because no explicit representations need be involved in or result from such a comparison and adjustment. As in the case of the self-world distinction, the idea is that the machinery for self-regulation would not have evolved, and would not operate as it does, unless its function is to register, check upon, and correct an organism's relatedness to its goals.

The standard forms of *physiological* and *sensorimotor* self-regulation are feedback and feedforward. They maintain the homeostasis of an organism against perturbations and guide its goal-directed actions. Relative to a preset value, feedback corrects perturbations or actions after the fact, whereas feedforward anticipates perturbations and actions. Primitive organisms, such as bacteria, monitor and control their movements to their goals by simple feedback mechanisms that regularly sample their surroundings. More complex organisms evolved internal models of their environment which allow them to anticipate by feedforward how the environment would look like in the immediate future. More sophisticated such models also anticipate how the environment would be affected by the organism's actions and even how, as a result, the organism itself would be affected. Thinking may have evolved out of these latter models, as I speculate later.

A key stage in this evolution, relevant to the present analysis, is metamotor imagery. Think of it this way. In order to track an organism's

bodily positions relative to the targets of its actions, motor images are fed into preset action schemas or monitored by feedback or feedforward routines that compare snapshot by discontinuous snapshot the sensory input with internal output copies of the targets of action. Suppose that the organism comes under selection pressures to track its actions continuously and modify them in terms of what the motor images represent at any given moment and of how these images compare with the intended targets of actions. The organism would do better if it can represent and monitor continuously the motor images themselves *in their relation to* their targets. This will be the job of higher-order and explicit *metamotor images*. The target-directedness or intentionality of motor images thus becomes an object of *metaintentional* self-regulation. This may be the earliest version of an executive metasemantics—a new representational capability selected for and involved in the executive work of self-regulation.

This executive-cum-representational innovation may have evolved when intricate motor behaviors, possibly connected, in primates, with complex arboreal behaviors or, in humans, with novel uses of hands, such as throwing or tool making and tool use. Whereas first-order motor images represent actions relative to bodily states and external targets, the metamotor images allow an explicit and continuous comparison between what first-order motor images represent and internal models from motor memory—the feedback part—and action predictions made by a planning center—the feedforward part (Damasio 1999; Jeannerod 1997).

3.3 A thoughtful mind

We reach now a critical juncture in the explanation of the formation of a sense of thoughtful self-intentionality as a premise for self-ascriptions. Once brought under one's voluntary control, the metamotor images, particularly of the feedforward sort, seem to be plausible evolutionary forerunners of *thoughts* as mental representations that anticipate and rehearse or (in short) *simulate actions*, and hence how the world and the organism's place in it would or could be changed by its actions, and also *simulate perception*, and hence how the world would or could look like. (These are simulations of one's own actions and perceptions as part of voluntary self-regulation, and not simulations of others in a naive-psychological sense, as in other-ascriptions or imitation.) The idea, developed next, is that thinking as simulation of action and perception and its self-regulation constitute the early premises of a sense of thoughtful self-intentionality and therefore of the self-ascriptions of thoughts and attitudes.

Before proceeding, I note that the notion of simulative thought is stronger than the standard notion that portrays thoughts as mere mental representations. On the latter notion, practically any organism may be said to have thoughts, usually generated by perceptual or memory inputs or schemes of action. Without arguing the point here, I prefer the notion of an active and potentially voluntary and thus self-regulated thinking, which is why I regard the simulations of actions and perceptions as the first genuine signs of thoughtfulness. In other words, I take a thoughtful mind to be one that regulates its actions and perceptions by simulating them, and also regulates its own simulations.

4 Thinking as simulation

It has been often said that simulative thinking is covert or suppressed behavior. It may be more accurate to say that simulative thinking evolved—and perhaps could only have evolved—when self-regulation became disconnected from action and aimed instead at the preparation and planning of action. Both in evolution and ontogeny it is motor behavior that is first simulated, prior to its execution. But action cannot be simulated alone, without a perceptual surround. It is the primary job of mental imagery to simulate perception, particularly visual (Hesslow 2002). Mental imagery may thus be the second form of self-regulatory simulation to emerge in evolution as well as human ontogeny.

Construed as simulations of actions and perceptions, thoughts are structured representations of which the thinker is aware as having a mind-initiated existence and simulated or projected (as opposed to world-caused) contents that anticipate or rehearse some future state of affairs. A thinker, in this sense, is aware of her own simulations, does not confuse their simulated contents with world-caused contents (perceptual, for example), and can voluntarily initiate and use the simulations to some anticipated end. Such simulation amounts to thinking, to the extent that it is conscious and deliberate initiation and manipulation of action images and mental images. A closer look at how simulative thinking works reveals how it underpins a sense of the self-intentionality of thoughts.

4.1 Thought as simulation

I propose to call the mental simulation of action *simulaction*. My reading of the neuropsychological literature suggests the possibility—admittedly speculative but rather plausible—that simulaction may have evolved out of

two prior self-regulatory mechanisms for feedforward control and metamotor imagery. The idea is that capacities for simulation evolved to upgrade and employ metamotor images with the aim of anticipating and rehearsing off-line the actions to be undertaken. I first expand this idea and then document it empirically.

We recall that the function of metamotor images is to monitor the *relation* between motor images, which represent the condition and position of the body, and the targets of its action. It is the feedforward or anticipative employment of the metamotor images that is likely to morph them into simulactive or motor-oriented thoughts. So employed, simulations inherit the self-regulatory function of metamotor images, which consists in tracking the intentionality of one's own body and actions directed at external targets.

The difference is that simulations exercise this function off-line, in advance of action. Unlike metamotor images, which are processed on-line, automatically and unconsciously, simulations are deliberately and consciously intended, and constructed and manipulated off-line. Simulations first emerge as new representational modalities of self-regulating future actions directed at future targets. To be effective in fulfilling their self-regulatory intentions, the simulations themselves must be monitored, evaluated and controlled—that is, self-regulated, at a higher level. Only such higher-level self-regulation can distinguish deliberate simulations from (say) motor hallucinations, form the right simulactive intentions or improve simulations by training and feedback. Since simulations rehearse off-line how the body and its actions relate to targets in the world, the intentionality of the simulations themselves must now be factored into their higher-level self-regulation. The result is no longer just a sense of metamotor self-intentionality, bearing on the relation of motor imagery to motor targets, but rather a distinct and higher-level sense of a simulactive-thought-to-motor-action-to-target set of relations—or, in short, a sense of the *ideomotor* intentionality of one's own simulations as off-line metamotor thoughts. So construed, a sense of ideomotor self-intentionality is necessarily a sense of one's own simulations being directed, off-line, at one's intended actions in relation to their targets. This is the idea I want now to document empirically.

Two neuropsychological facts about simulation support the analysis just sketched. One fact is that the simulation of movement seems to involve some of the *same* brain areas as does the preparation of motor behavior. The same cortical overlap holds between consciously representing an action and intentionally executing it. This means that simulation draws on many of the same brain resources as does the formation of motor intentions. The only

major cerebral difference between simulation and the intended execution of an action is the inhibition of the execution (Jeannerod 1997, 1999, 2003a,b). The other supporting fact is that the execution of an action takes the *same time* as its mental simulation (Decety et al. 1989).

These results are open to divergent interpretations. One major issue is whether simulation is simulated movement rather than simulated perception of movement (Currie and Ravenscroft 2002, chapter 4). Leaving this interesting but difficult question aside, the implication that matters here is this. Since metamotor imagery handles both the simulation and the execution of motor actions, and does so with many of the same brain resources and in the same time period, it appears that simulation does employ the metamotor-imagery machinery in an anticipative or feedforward mode. Marc Jeannerod is explicit on this point: motor representations resemble feedforward models in that they estimate the outcome of an action without waiting for the sensory reafference or even without performing the action. This is to be expected from a mechanism that plans and controls the execution of an action by simulating it (Jeannerod 1997, p. 175). As suggested earlier, simulations are therefore likely to emerge out of the self-regulatory work of metamotor imagery and to inherit from that work a sense of ideomotor self-intentionality.

Yet this is only half of the story. Unlike metamotor images, simulations are deliberately and consciously initiated and handled. Therefore, the simulator must not only be aware of simulations not being actual actions. She must also be aware of her simulations being about her own future actions as intended to be directed at some target of interest. Jeannerod (2003b) notes that the dorsolateral prefrontal cortex is activated during the preparatory stage of simulation but not during the execution of an action. The prefrontal cortex is the conscious initiator and manager of simulations.

Jeannerod also notes that (despite the inhibition of motor execution) the motor cortex is also active during simulation, and opines that this activity may provide information for consciously monitoring the simulations themselves. This monitoring information, coming from the motor cortex, about the simulated action, joins the information about the goal and target of action held in working memory and available to prefrontal self-regulation (Jeannerod 2003b). This assembly of metamotor and prefrontal information seems to enable the simulator to self-regulate her simulations by tracking their intentional relations to intended actions and their targets.

To sum up, the self-regulation of simulation as mental activity is distributed, probably hierarchically, among several brain centers having distinct functions but ends up providing one with a conscious sense of the intentionality of one's own motor-oriented thoughts. This is the key result at this point in the analysis. Given the primacy of action and sensorimotor conation and cognition in human ontogeny, a sense of the ideomotor intentionality of one's own simulations is very likely to be the child's first major step toward self-ascriptions of I-thoughts. It is a necessary step. But is it also sufficient, even with language, consciousness, and some early naive-psychological concepts thrown into the mix, all of which are available to and used by the young child even before the age of 4? I think not, for at least two reasons.

First, simulations would not work unless they project an action in a spatial context populated by sundry objects and events in relation to the body and actions of the agent. Such projections rely on distal perceptual information—mostly visual. These off-line visual projections are handled by mental imagery. Second, human thinking involves a great deal of simulation that has little if anything to do with one's body and actions and a good deal to do with figuring out or imagining—often abstractly and impersonally—future, hypothetical or counterfactual situations. Such imaginative simulation is also likely to take off from mental imagery. It turns out, ontogenetically, that conscious and deliberately intended and manipulated mental imagery is a rather late development.

4.2 Thought as simulimage

Despite debates over its encoding format (pictorial or language-like), there is growing neurological evidence that mental imagery uses many of the same mechanisms as visual perception (Kosslyn et al. 2001). Imagining a visual scene or performing a task that involves visualization activate the same primary visual cortex (Le Bihan et al. 1993). Damage to the visual cortex most often impairs the ability to form mental images (Farah 1988). We have a strong parallel here with simulations, which were shown also to exploit many of the same brain centers as actual intentional execution of actions. And, just as simulations were said to inherit a sense of motor self-intentionality from the machinery of motor imagery, it makes sense to expect mental imagery or simulimagery to inherit a sense of perceptual self-intentionality from how the visual system works (Jacob and Jeannerod 2003).

Yet, again as in the case of simulations, such a basic but low-level sense of perceptual self-intentionality may be necessary but not sufficient for

an internal sense of the thoughtful intentionality of one's own simulimages. It is not enough to know that a mental image I form of X is about the X I remember seeing some time ago. The similarity may be accidental. Or I might hallucinate or fantasize a mental image that happens to be close enough to or even match a former visual image. The fact that mental images represent what they do, by exploiting the visual system and its outputs, is part of how they acquire their own intentionality, but that fact does not yet explain *my sense and awareness*, as active simulimager, of the intentionality of my mental images.

As in the case of my simulations, I must also consciously *intend* my mental images to represent the X in question (or whatever). And intending my mental images to represent what they do entails my being able to *self-regulate* my mental imagery as a purposeful cognitive activity. Solving a visual problem (e.g., how to fit the bulky suit-case in the puny trunk of a small car) or figuring out visually how parts may fit into a whole (e.g., organizing the furniture in a room) requires an awareness that one's own visually simulative thoughts represent what one intends them to represent. It also requires an ability to check on and improve the fit between the intention to represent and the actual outcome. For some tasks, such as making and using tools, some people (or even some primate species) may be better mental imagers than others, and for many tasks training can refine mental imagery. Both mastery of mental imaging and improving it by training may be a matter of working on the self-regulation of mental imagery.

How the self-regulation of mental imagery works and therefore how it delivers a sense of the self-intentionality of one's simulimages are difficult topics that are beyond the scope and competence of this paper. Some evidence suggests that the ontogenetically simpler and earlier forms of mental imagery are likely to take off from the visuomotor machinery of simulation and the regulatory machinery of perception itself, even when dealing with abstract tasks that do not involve actions or specific percepts, respectively. Thus, it has been shown that transformations of mental images—as, for example, in action-free mental rotation—are largely guided by (what was called earlier) metamotor imagery (Wexler et al. 1998). Such a result comports with the behavioral immediacy and concreteness (so to speak) of the young child's simulative thinking. And if it turns out that language and language-like processes are also involved in generating and regulating mental imagery—as they are likely to, since we often literally talk ourselves into simulimaging and revising simulimages—then the sense of self-intentionality of the young child's mental images is likely to be shaped in part by the semantics of her early language. As I read developmentally the

message of cognitive linguistics, the child's early semantics is likely to be embodied (in the sense of embodied cognition) and tied to action and perception patterns and other forms of interacting with concrete situations (Tomasello 2003).

Having made a case for the tight link between the child's sense of thoughtful self-intentionality and the self-regulation of her simulative thinking (itself a lower-level form of self-regulating future actions and perceptions), I want to conclude with the suggestion that the late ontogenesis of off-line simulations and simulimages and of their intramental self-regulation confirms the fact that intentionality-sensitive self-ascriptions of thoughts and attitudes develop differently from other-ascriptions and largely for different reasons.

5 The late development of inside loops

Let us first recall from earlier sections three major features of the naive psychology of early childhood: its predominant orientation toward other minds, its situatedness in the here-and-now of perception and motivation, and its on-line operation. I think these three features conspire to keep the mind of the young child blind to itself, not experientially but executively and hence metaintentionally or metasemantically, due to the late development of intramental self-regulation and anticipative off-line simulation. If the argument of this paper (and of Bogdan 2005a) is right, this late development prevents the young child's naive psychology to turn to its thoughtful self and attitudes.

This conclusion is fortified, I think, by a revealing connection between the left hemisphere and the early dominance of situated and on-line other-ascriptions. The left hemisphere is the seat of basic, specialized and mostly evolution-installed faculties (Corcoran 2000; Deacon 1997). In intensely sociopolitical species like primates and humans, one would expect natural selection to pressure for first figuring out the minds of conspecifics from their observable behaviors and relations to a shared world, and one would expect the same in the helpless and adult-dependent primate and human offspring (Tomasello 1999; also Bogdan 1997, 2000, 2003).

Yet nothing said so far should suggest that young children cannot simulate their own actions and perceptions for self-regulatory purposes. The suggestion, rather, is that their simulations and simulimages are likely to be *situated* and *on-line*, and hence anchored in and mostly guided by current desires and perceptions. Young children can imagine partly possible situations as limited variations on currently perceived or desired themes.

They imagine situations that resemble a currently perceived or desired one but may differ through some permutation, addition or omission of some elements. In contrast, the older child's simulations of actions and perceptions are different in that they can be unsituated, off-line, and often totally divorced from current perception and motivation.

I think that pretending, mental imagery and mental rehearsal exhibit this contrast. The pretend play of early childhood is situated and guided by current desires and perceptions, whereas the pretense of later childhood can be unsituated and thoroughly imaginative and often speculative. Or so I read the data from the psychological literature (Harris 2000; Perner 1991), as I argued elsewhere (Bogdan 2005b). Also, according to Piaget and other psychologists, young children are poor mental imagers, apparently unable to form mental images required in mental rotation tests or to use images to solve visual-spatial problems (Foulkes 1999, pp. 65–66; Piaget and Inhelder 1971). Piaget distinguished between reproductive images that reenact previous experiences and anticipatory images that represent novel and unexperienced transformations of images. He thought that until around 7 children cannot entertain images of the latter sort. This pessimism is not shared by more recent research (Bjorklund 2000, chapter 7). Even if 7 is too late, Piaget's distinction may still work for early childhood and, for that period, it seems to fit other limitations mentioned here.

Mental rehearsal is a concept meant to cover a family of forms of mental simulation and anticipative imagination. Mental rehearsal seems to follow the same gradual and laborious development as mental imagery, autobiographical memory and narration. Preschoolers rarely plan and rehearse a simple arbitrary sequence, when required to retain it over a brief period (Bjorklund 2000). They fail to realize the value of mental rehearsal even after demonstration, and cannot reflect on the successive parts of their behavior or relate their mental rehearsal to subsequent task performance (Flavell et al. 1995).

Imaginative pretense, mental imagery, mental rehearsal, sequential reasoning, and in general simulative thinking—all require inhibition, before anything else, as a step to unsituatedness and off-line operation. Or, as already noted, inhibition becomes operative only after the 3–4 interval. These same simulative activities also require an ability to operate simultaneously with and interconnect many and often conflicting representations. A multiplex mind owes a great deal to the interconnectivity of the right hemisphere, which also develops slowly, particularly after the age of 4. Inhibition and the multiplex ability to open simultaneously several mental screens and switch from one screen to another as well as connect

them bring about unsituated, off-line simulations about possible situations or indeed about other simulations (thoughts about thoughts). The self-regulation of all these kinds of off-line simulations, targeting their intentional directedness and adequacy, is managed, by way of inside loops, by an executive metasemantics deployed by the prefrontal cortex. The latter, too, was noted to undergo a major burst of development only after the 3–4 interval (Diamond 2001).

It is in this neurodevelopmental context that the child acquires a sense of the intentionality of her own thoughts, first as simulations of actions and perceptions, and then as unsituated, anticipative and fully imaginative off-line simulations, including simulations of simulations—all critical conditions for the self-ascriptions of thoughts and hence attitudes.

6 Envoi

Little has been said here about the naive-psychological *concepts* that govern the self-ascriptions of thoughts and attitudes, except that the concepts incorporate the recognition of thoughtful self-intentionality. If the argument of this paper is on the mark, the child's concepts of self-attitudes cannot be *formed* in quite the same as the concepts of other-attitudes, as is most often assumed in the developmental literature (Bogdan 2005a). A full account of self-ascriptive concepts of thoughts and attitudes must explain when and how the executive metasemantics aimed at one's own intentional mind is aligned to the public and other-directed ascriptions of attitudes and in general to the public epistemic and naive-psychological discourse and ascriptional practices. That is a story for another time. But one implication can be discerned even at this stage.

A self-directed naive psychology is likely to be a *hybrid* construction that combines other-directed ascriptional concepts, self-regulatory mechanisms, and norms of public discourse and ascriptional practices. I think (but have not argued here) that even the other-directed naive psychology of early childhood is a hybrid of specialized mechanisms that join a variety of resources, with different functional agendas, not all initially implicated in naive psychology (Bogdan 1997, 2000). The upshot is that, contrary to prevailing opinions, no unitary and domain-specific account is likely to explain naive psychology, other- as well as self-directed. But this need not be a surprise: the human mind itself is a complicated hybrid. Why should its naive psychology be any different?

Acknowledgements I borrowed the title ‘Inside Loops’ from Keith Lehrer’s excellent book *Self-Trust* (Oxford University Press, 1997). This paper went through several versions, read in several places, where it occasioned many useful and insightful comments, criticisms and suggestions. I thank Keith Lehrer very warmly for his thoughtful input and for being so supportive in so many ways during the work on this paper. I also thank audiences at the University of Arizona, Tucson, and the University of Bucharest, as well as select group of participants at the 3rd International Colloquium in Analytic Philosophy in Bucharest, May 2005. Warm thanks to Dan Dennett for an extensive discussion of the notion of self-intentionality while driving through and admiring the bucolic landscapes of southern Transylvania.

References

- Astington, J. W., & Gopnik, A. (1988). Knowing y’ve changed your mind. In J. Astington, P. Harris, & D. Olson (Eds.), *Developing theories of mind*. Cambridge: Cambridge University Press.
- Bjorklund, D. F. (2000). *Children’s thinking*. Belmont: Wadworth.
- Bogdan, R. J. (1997). *Interpreting minds*. Cambridge: MIT Press.
- Bogdan, R. J. (2000). *Minding minds*. Cambridge, MIT Press.
- Bogdan, R. J. (2003). Watch your metastep: The first-order limits of early intentional attributions. In C. Kuznietz et al. (Eds.), *Persons*. Vienna: obv&hpt.
- Bogdan, R. J. (2005a). Why self-ascriptions are difficult and develop late. In B. Malle & S. Hodges (Eds.), *Other minds*. New York: Guilford Press.
- Bogdan, R. J. (2005b). Pretending as imaginative rehearsal for cultural conformity. *Journal of Cognition and Culture*, 5, 191–213.
- Bogdan, R. J. (2008). *Predicative minds*. Cambridge: MIT Press.
- Clements, W., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9, 377–397.
- Corcoran, R. (2000). Theory of mind in other clinical conditions. In S. Baron-Cohen et al. (Eds.), *Understanding other minds* (2nd ed.). Oxford: Oxford University Press.
- Currie, G., & Ravenscroft, I. (2002). *Recreative minds*. Oxford: Oxford University Press.
- Damasio, A. (1999). *The feeling of what happens*. San Diego: Harcourt.
- Deacon, T. (1997). *The Symbolic Species*. New York: Norton.
- Decety, J., Jeannerod, M., & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioural Brain Research*, 34, 35–42.

- Diamond, A. (2001). Normal developments of prefrontal cortex from birth to young adulthood. In D. T. Stuss, & R. T. Knight (Eds.), *The frontal lobes*. Oxford: Oxford University Press.
- Farah, M. J. (1988). Is visual imagery really visual? *Psychological Review*, *95*, 307–331.
- Flavell, J. H., Green, F. L., & Flavell, E. R. (1995) Young children's knowledge about thinking. *Monographs of the Society for Research in Child Development*, *60*, 1–96.
- Foulkes, D. (1999). *Children's dreaming and the development of consciousness*. Cambridge: Harvard University Press.
- Goldman, A. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences*, *16*, 15–28.
- Gopnik, A. (1993). How we know our minds. *Behavioral and Brain Sciences*, *16*, 1–14.
- Harris, P. (2000). *The work of imagination*. Oxford: Blackwell.
- Hesslow, G. (2002). Conscious thought as simulation of behavior and perception. *Trends in Cognitive Science*, *6*, 242–247.
- Hobson, R. P. (1993). *Autism and the development of mind*. Hillsdale: Erlbaum.
- Jacob, P., & Jeannerod, M. (2003). *Ways of seeing*. Oxford: Oxford University Press.
- Jeannerod, M. (1994). The representing brain. *Behavioral and Brain Sciences*, *17*, 187–245.
- Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford: Blackwell.
- Jeannerod, M. (1999). To act or not to act. *Quarterly Journal of Experimental Psychology*, 1–29.
- Jeannerod, M. (2003a). Consciousness of action and self-consciousness. In J. Roessler & N. Eilan (Eds.), *Agency and self-awareness*. Oxford: Oxford University Press.
- Jeannerod, M. (2003b). Simulation of action a unifying concept for motor cognition. In S. Johnson-Frey (Ed.), *Taking action*. Cambridge: MIT Press. 123
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*, 635–642.
- Le Bihan, D., Turner, R., Zeffiro, T. A., Cue'nod, C. A., Jezzard, P., & Bonnert, V. (1993). Activation of human primary visual cortex during visual recall. *Proceedings of the National Academy of Sciences*, *90*, 11802–11805.

- Nelson, K. (1996). *Language in cognitive development*. Cambridge: Cambridge University Press.
- Perner, J. (1991). *Understanding the representational mind*. Cambridge: MIT Press.
- Piaget, J., & Inhelder, B. (1971). *Mental imagery in the child*. New York: Basic Books.
- Russell, J. (1996). *Agency*. Hillsdale: Erlbaum.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge: Harvard University Press.
- Tomasello, M. (2003). *Constructing a language*. Cambridge: Harvard University Press.
- Wexler, A., Kosslyn, S., & Berthoz, A. (1998). Motor processes in mental rotation. *Cognition*, 68, 77–94.
- Wolpert, D., Miall, R., & Kawato, M. (1998). Internal models of the cerebellum. *Trends in Cognitive Science*, 2, 338–347.