

# **Metarepresentations**

**A Multidisciplinary Perspective**

edited by Dan Sperber

**OXFORD**  
UNIVERSITY PRESS

2000

## Chapter 2

### *Making Tools for Thinking*

Daniel C. Dennett

Representations are themselves objects in the world, and therefore potential objects of (second-order or meta-) representations. However, humans seem to be nearly unique in their ability to represent representations.

—Dan Sperber, 1996, in setting the topic for this conference

Just as you cannot do very much carpentry with your bare hands, there is not much thinking you can do with your bare brain.

—Bo Dahlbom and Lars-Erik Janlert (unpublished)

We use intelligence to structure our environment so that we can succeed with *less* intelligence. Our brains make the world smart so we can be dumb in peace!

—Andy Clark, 1997, p.180

If it be maintained that certain powers, such as self-consciousness, abstraction, etc. . . . are particular to man, it may well be that these are the incidental results of other highly advanced intellectual faculties; and these again are mainly the result of the continued use of a highly developed language.

—Charles Darwin, 1871, p. 101

Reflection on the issues raised by these observations has driven me to a tentative conclusion: in spite of what may well have been a surfeit of attention to definitional issues, we still do not know what we (all) mean when we talk about representation, and hence what we mean by metarepresentation. I am far from ready to offer a cure for this confusion, but I do have a few suggestions to offer about what the problems are and why they persist. My thinking has been much influenced recently by several new books that deal at least indirectly with these topics: Steven Mithen's *The Prehistory of the Mind* (1996), and Andy Clark's *Being There* (1997), but at this point I am still an asker, not answerer, of questions.

## 1. Florid and Pastel (Meta)representation

I must sneak up on the topic, not having figured out how to mount any frontal assault. The ethologists Robert Seyfarth and Dorothy Cheney once told me an unforgettable anecdote about some lions hunting in their neighborhood in Amboseli Park in Kenya. Robert and Dorothy had parked their Land Rover on high ground to show their guest, Don Griffin, the panorama, and below them nearby they spotted three lions crouching behind cover on some medium-high ground, overlooking a small herd of wildebeests. While the three ethologists watched in fascination, the lions put into execution a remarkable hunting ploy. One lion stepped forward into full view of the wildebeests, which all turned, with some nervousness, to eye her cautiously, ready to bolt the moment she made her move. While lion A held the herd's rapt attention in this fashion, lion B crept off to the left, circling around into position in a shallow ditch, unseen by the herd. Meanwhile lion C crept around to the right, more or less on the opposite side of the small herd from lion B. Once in place, lion C leapt out and gave chase. The herd bolted away, of course, and stampeded right over the ditch where lion B lay in wait. Lion B merely had to leap up to catch and bring down one animal, providing supper for all.

A delectable scenario; it leaves one wondering how much planning, how much coordination or even communication may have gone into the control and timing of that episode. It might be, of course, that appearances were deceptive – the lions' meal might have been more dumb luck than design. But probably not, and we can dimly imagine the sorts of cognitive mechanisms and practices that could account for their hunting success's not being an accident. Now let me add an embellishment that nobody would or should believe. Before the lions crept into their attack positions, lion A was observed to move a small bunch of twigs with her nose into a little group in the dust in front of the three lions; then she pushed three stones, one at a time, into another group, on a little bump in the dust. Then, checking for shared attention in lions B and C, she pushed two of the stones to left and right of the group of twigs, and then dashed the rightmost stone across the bunch of twigs, sending it over the stone on the left. And so forth.

*Much* too good to be true, of course. We have no grounds at all for believing that *that* sort of representing is within the talents or cognitive horizons of lions. Let us call this *florid representing*. What makes it florid, one sees dimly, is that it is deliberate representing, knowing representing, even self-conscious representing. It seems you cannot engage in florid representing without knowing and appreciating that you are engaging in florid representing. Or at least that will do for the moment as a mark of the florid. Now it is just a short and inviting step from this

observation to the suggestion that such florid representing involves *metarepresentation*, since knowing that you are representing involves having a mental state that is itself *about* the representing you are doing and, moreover, about that representing qua representing. And this invites a further step, or slogan: No florid representation without metarepresentation.

This may be the right way to think about florid representation, but I want to move more cautiously. If that is florid representing, what is pastel representing? Is just seeing one's prey pastel representing? Is it representing at all? One has a mental or perceptual state that is surely *about* the prey (qua prey, one might add), and if that perceptual state plays an apposite role in guiding one to the prey, this state should count as a representation *in use*, or a representation that exists *for* the predator in question. That is enough, according to a familiar theme of mutual agreement, for us to speak of representation. Alan Leslie, according to his abstract for this volume, begins at this starting point. Rodney Brooks (e.g., 1991), however, uses the term "representation" in a way that would permit him to deny that such straightforward perceptual states count as representations. For Brooks, representations are the sorts of data structures GOFAI trafficked in – the manipulable items of "physical symbol systems" – and he thinks (with many others) we can get intelligent behavior without *them*.

Suppose, for the time being at least, we follow Brooks in denying the label of "representation" to such perceptual states. What, then, of the various sorts of phenomena that seem to be intermediate between such a perceptual state and a token playing its role in an instance of florid representation? What, for instance, of somebody – a young child, say – yelling "snake!" when a snake slithers into view? Is the child's utterance not a representation in a stronger sense than her visual state that provoked and guided her utterance?

Note that the child's utterance may very well not pass our makeshift test of floridity. It seems right to say that young children speak before they realize they are speaking, use language without noticing that they are doing so, respond appositely to all manner of linguistic moves, their own and those of others, without any self-conscious or deliberate metarepresentational machinery invoked at all. What young children do with language is, it seems, more like the leopard cries of vervet monkeys than the imaginary token-manipulations of our lions. And yet it clearly goes way beyond what the vervets do as well, being famously amenable to combinatorial authorship and elaboration, for one thing.

In any case, children soon enough come to be self-conscious about their use of words, and when that happens, if not before, they are capable of florid representation. What are the sure signs of this advance? I do not have a settled view on this yet, but I can offer an example. When my son

was about five years old, he was captivated by what may have been the first joke he "got." "There's an old dead skunk lying in the middle of the road. I one it," I said. Now you say, "I two it." "I two it." "I three it" quoth I. "I four it!" he continued. "I five it." When it was his turn to say "I eight it," I looked at him with an expression of shock and disgust, and his eyes grew wide with delight. For several hours, we had to play this simple game, over and over, taking turns with the odd and even numbers. Such are the first faltering steps towards wit (and I have recently replicated this micro-experiment with my 5-year-old grandson). What shall we say about such moments? Then, if not before (and I am not at all sure about that), a child demonstrates a clear recognition of the existence of words as sounds, with roles to play, that might be combined in rule-bound ways – though of course the child would typically be incapable of putting it in such terms at the time.

But if this is an instance of unmistakable metarepresentational prowess, what are we to make, then, of a child's really quite adult and accomplished use of language preceding such an epiphany? (One is reminded, of course, of Molière's gentleman who was delighted to learn he'd been speaking prose all his life.) Was it somehow accompanied by, guided or informed by, metarepresentational states but just not *wittingly* so guided?<sup>1</sup>

## 2. Thinking about Thinking

I do not know how to answer that question yet, so I will turn to a related question which has recently been bothering me: Is there perhaps a big difference between *having beliefs about beliefs* and *thinking about thinking*?<sup>2</sup>

For some years now, I have been urging the importance of looking for evidence of "higher-order intentional states" – beliefs about beliefs, desires about beliefs, beliefs about desires about beliefs, and so forth – and all of this might very well be characterized as urging the importance of looking for evidence of metarepresentation. I do not at all mean to recant anything I have said on this score – not yet, at any rate – but I have begun to wonder whether it might not distract our attention from a neighboring source of insights, which might get us closer to an examination of the real powers of metarepresentation, properly so called.

Paying attention to *beliefs about beliefs* has certainly borne some good fruit. The theory-of-mind literature, the various paradigms of false-belief tests of children and animals, and all that surrounds this growth industry in investigation, has been more than routinely productive of insight in several fields, I think, but I have come to wonder if we are not missing some better questions in the immediate vicinity. A chimpanzee may be a natural psychologist, as Nicholas Humphrey has said, or may

have a theory of mind, as David Premack (Premack & Woodruff 1978) and Alan Leslie (this volume) and Simon Baron-Cohen (1995) and many others have said, but – as I have recently been urging (e.g., in *Kinds of Minds*, 1996, pp. 130 ff.) – chimpanzees never have opportunities to compare notes with other psychologists, never get to *ask for the reasons* that ground the particular attributions of mentality to others, never get to communicate to others *or to themselves* about these mental states of others. I am tempted to say that even if they do have beliefs about beliefs, they may well be incapable of *thinking about thinking*. They may, indeed, not really be capable of *thinking* at all (in some florid but important sense of “thinking”).

Our imaginary lions plotting their next moves were certainly thinking, using props as representations of the things in the world they were thinking about. But they were imaginary. When children playing touch football crouch down in the huddle and one says “Jimmy goes deep, Joey slants left, and I’ll block for Tom who looks for the open man” (with or without a diagram traced with a finger on the ground or on an extended palm), this is an unmistakable instance of florid representation, and communal thinking. It is not necessarily thinking about thinking, but we can readily turn it into thinking about thinking, by letting their plans hinge on deception of some particular sort (“Jimmy will pretend to slip and fall, and then, as soon as they ignore him . . .”)

What strikes me as clear enough is that there is *one* pathway to such florid thinking about thinking that moves from the outside in. It begins with the overt, public use of symbols and tokens of one sort or another (spoken words will do, perhaps, but only if they are used self-consciously), and creates practices that later can be internalized and rendered private. Since chimpanzees, for instance, lacking language, cannot play these external games any more than lions can, they cannot use *this* path as a route to private or covert thinking about thinking. This line of thought fits handsomely with the recent emphasis by Andy Clark, and indeed by me (in *Kinds of Minds*), on the claim that *minds are composed of tools for thinking* that we not only obtain from the wider (social) world, but largely leave in the world, instead of cluttering up our brains with them.

Might there, however, be other routes? Might there be pastel versions of metarepresentation that can arise by some process other than by the fading (or sophistication) of originally florid practices? It has commonly been supposed that there might indeed be such developments. The child’s acquisition of a theory of mind, for instance, might seem to be a developmental story that does not at any point *depend on* the child’s prior mastery of public, communal, florid practices of representation of the minds of others. After all, it has seemed quite clear that there is at least a strong genetic predisposition in our species – in normal, non-

autistic children – to develop this way of “thinking about” other people, and indeed about animals and almost anything that moves.<sup>3</sup>

But if so, and if this penchant for *unwittingly* adopting the intentional stance is something we share with at least some of our cousin species, how do we, and we alone, come to metarepresent in the more fluid, witting ways? How could such metarepresentational abilities arise in the first place? I do not think we philosophers have been sufficiently puzzled by these questions – indeed, we have hardly noticed them. We tend to take for granted the categories we find in folk psychology, which encourage us to impose a sorting in terms of these categories on all the intermediate cases. We ask: Is this one an instance of higher-order belief or not? Is this one an *echt* case of meta-representation, or should we put it in the “other” pigeonhole? These questions mask the developmental or evolutionary processes that must lead from the simple to the complex by a sequence of iffy or marginal candidates.

Sometimes I am struck by the weird complacency with which we are inclined to accept these categories at face value. How on earth *could* any agent, any organism, come to develop the *knowing* competence to *use representations*? It might be “only” a matter of practical – not logical – necessity that one acquire this special ability via the external, social route, but that fact – if it is one – might still be the key to understanding how it is possible at all. It might, for instance, be the key to understanding the general distinction I have drawn (1969) between the personal and subpersonal levels. A reminder may be in order: parsing sentences is not a personal-level activity for many of us; it is something few of us ever do except on those rare occasions when we discuss niceties and perplexities of grammar in an academic setting; at the subpersonal level, in contrast, parsing sentences is a phenomenon that occurs *within* us whenever we follow a conversation or read with comprehension. We must be careful not to confuse the two phenomena, and in particular not to assume that a good model of the former, rare phenomenon is also a good model of the latter, ubiquitous phenomenon. In the present context we may ask when, if ever, animals engage in personal-level activities at all, let alone personal projects of metarepresentation.

Let me propose a case in which we can see the philosophical sleight-of-hand happening right in front of our eyes: a puzzle is masked by our accepting an invitation to treat a curious phenomenon as unproblematically falling into our standard, human, folk-psychological categories. Many years ago, Bertrand Russell made a wry observation: “Animals studied by Americans rush about frantically, with an incredible display of hustle and pep, and at last achieve the desired result by chance. Animals observed by Germans sit still and think, and at last evolve the situation out of their inner consciousness” (Russell, 1927, pp. 32–33). Wolfgang Köhler’s (1925) early experiments with chimpanzees were the

inspiration for Russell's witticism, which helps to perpetuate a common misunderstanding. Köhler's apes did not just sit and think up the solutions. They had to have many hours of exposure to the relevant props – the boxes and sticks, for instance – and they engaged in much manipulation of these items. Those apes that discovered the solutions – some never did – accomplished it with the aid of many hours of trial-and-error manipulating.

Now were they *thinking* when they were fussing about in their cages? What were they manipulating? Boxes and sticks. It is all too tempting to suppose that their external, visible manipulations were accompanied by, and driven by, internal, covert manipulations – of thoughts about or representations of these objects, but succumbing to this temptation is losing the main chance. What they were attending to, manipulating and turning over and rearranging, were boxes and sticks, not thoughts.

They were *familiarizing themselves* with objects in their environments. What does that mean? It means that they were building up some sort of perceptuo-locomotor structures tuned to the specific objects, discovering the affordances of those objects, getting used to them, making them salient, and so forth. So their behavior was not all that different from the incessant trial-and-error scrambling of the behaviorists' cats, rats and pigeons. They were acting in the world, rearranging things in the world – without any apparent plan or insight or goal, at least at the outset.

Animals at all levels are designed to tidy up their immediate environments, which are initially messy, confusing, intractable, dangerous, inscrutable, hard to move around in. They build nests, caches, escape tunnels, ambush sites, scent trails, territorial boundaries. They familiarize themselves with landmarks. They do all this to help them keep better track of the things that matter – predators and prey, mates, and so forth. These are done by "instinct": automatized routines for improving the environment of action, making a better fit between agent and world.

This wise husbandry of one's own behavioral territory is economically focussed on the most important and ubiquitous features. Pre-eminent among these portions of the environment is the agent's own body, of course, always present, its condition always intensely relevant. Animals instinctively groom themselves and engage in behaviors that are apparently designed (although they need not realize this) to repair flaws and maintain and improve their coordination, muscle tone, and, in effect, to familiarize themselves with their own bodies. A part of the body that must not be overlooked in this maintenance and improvement schedule is the brain. It, too, can become messy, confusing, inscrutable, an overwhelmingly complex arena of action. So we should expect animals to be instinctually equipped to engage in mindless, automatic routines that tidy up their own brains.



We should especially expect it in *Homo sapiens*, whose huge brains are so plastic, so inundatable, so at the mercy of invading memes and memories. Resource management for a young human brain is, I think, a major task, and we should expect it to be accomplished by activities that are rooted, at the outset, in our biology, in our "instincts," but which also get enhanced in major ways by techniques that are themselves part of the influx of new resources.

### 3. Machines Made from Found Objects

Noam Chomsky has often said that birds do not have to learn their feathers and babies do not have to learn their language. I think there is a better parallel between birds and language: a child acquiring language is like a bird building a nest; it is a matter of "instinctual" or "automatic" resource-enhancement, taking found objects and constructing something of great biological value—part of what Dawkins (1982) calls the extended phenotype—which blurs the boundary between an organism (or agent) and the environment in which it must act.

The primary found objects, of course, are words, which, like the blacksmith's legendary tools, are tools for making more tools. As Andy Clark puts it, "The linguistic constructions, thus viewed, are a new class of objects which invite us to develop new (non-language-based) skills of use, recognition, and manipulation. Sentential and nonsentential modes of thought thus coevolve so as to complement, but not replicate, each other's special cognitive virtues" (1997, p. 211). It is because the lions cannot talk to each other that they also cannot come to use tokens of other sorts to represent, non-linguistically. And it is by that indirect route, I think, that we come to construct our minds. Here is a little fantasy about the Planet of the Baboons. The baboons that inhabit this planet are approached one day by an alien technocracy and offered an amazing prosthesis: an artifact that will permit a sort of time travel, far into the past and even into the future, and a sort of time compression: tasks that normally took days or weeks could be accomplished in a split second. What a boon these baboons were being offered!—but of course they could not even understand the offer, being baboons. We human beings can understand the offer, however, because we have already accepted it. We have already received the prosthetic extension of our brains that permits us to play such glorious tricks with time: we call it language.

First consider time "travel": Unlike the baboons, who cannot even conceive of *10 years ago*, and who may not even be capable of recollecting the specific events of their own past experience that have shaped their current dispositions,<sup>4</sup> we can conceive of the distant past, and use accurate information about events that happened centuries ago to guide our

current projects. We can also see accurately into the future, predicting not just regular patterns such as nightfall and daybreak, and the coming winter (a trick that is in some regards within the competence of many animals and even plants), but such one-off events as the much anticipated millennial celebrations around the world, and, somewhat farther in the future, thank goodness, the extinguishing of the sun some billions of years hence. No non-human creature has a clue about any of that. Their time horizons may stretch for some limited purposes from seconds to hours to a year or two (think of the faithful biennial return of the albatross to its mate), but aside from that they have no conception. There is some remarkable evidence to suggest that elephants have some inkling of their own mortality, but it is equivocal. Our time horizons, unequivocally, extend to the point of evaporation; we can even think about – if not yet think about very well – the question of whether time itself has a beginning or end.

Second, consider time compression: We can teach a child, in a few seconds, lessons that animal parents have to drill into their offspring. Even in the highly specialized cases of learning what foods to eat and not to eat, where the Garcia effect and similar special-purpose varieties of one-shot learning have been found, animal parents have scant control over *which* lesson they can teach.

Consider, as a thought-experiment, the problem of a mother dog needing to warn her puppies not to attempt to eat the bright orange toads, which are poisonous. Not being able to raise the subject in conversation, she must wait till such a toad is in the offing, or perhaps seek one out for purposes of instruction, and then whenever the puppies approach it, she must bark or otherwise ward them off. It would be a form of operant conditioning, evolutionarily tuned, no doubt, so that the behavior shaping might be much swifter than what can be accomplished by “ecologically invalid” scientists in laboratories, but still vulnerable to the same ambiguities and indeterminacies.<sup>5</sup>

To dramatize the problem we might imagine the puppies *asking themselves* “Is it just *this* orange toad, or orange toads that hop like this one, or small brightly colored moving things, or orange things that smell like this, or . . . ?” There are only two ways of answering these questions for the puppies: letting evolution answer them (by relying on whatever generalization-proclivities are built in), or varying the circumstances in further laborious training to clarify the boundaries of the classes of positive and negative stimuli. Can a mother dog learn, and teach to her young, that the food shaped and colored like hearts and spades, but not clubs and diamonds, is to be avoided? I do not know what empirical studies may have shown about the limits on such instruction – I think it is an interesting question worth pursuing, if it has not already been answered.

But we must not lose sight of the obvious fact that the puppies cannot literally (floridly) ask themselves these questions, lacking language. It is far from clear, in fact, that the mother, also lacking language, can even frame for herself the project of teaching anything so specific to her young. Her project, we may suppose, is to replicate in her young the disposition vis-à-vis orange toads that she acquired from her mother, but we must not impose our language-borne distinctions on the characterization of that disposition. For instance, in all likelihood, she has no way – and no need for a way – of distinguishing between passing along the generalization *in sensu composito*:

Believe, my young, the following universal quantification: For all  $x$ , if  $x$  is an orange toad, ingesting  $x$  is to be avoided.

and passing along the general lesson *in sensu diviso*:

If you ever encounter an orange toad, my young, believe at that moment: ingesting *this* is to be avoided.

The latter is the general disposition to have particular, perception-anchored beliefs about particular encountered toads; the former is the more intellectual state of mind of deeming true a universally quantified proposition about orange toads in general. In the normal course of canine affairs, the difference is not apt to loom large enough to make a difference, but in human affairs, it can be readily discerned in special circumstances. Small-town Tom, knowing everybody in town, believes all the town's Republicans are mean-spirited *in sensu diviso* (thus would he sort them, *seriatim*, if asked to list the mean-spirited). However, not knowing the political affiliations of all his fellow citizens – or not knowing whether or not he knew all the people in town – he would sincerely express agnosticism when given the generalization *in sensu composito* for assent or dissent. The latter expresses a proposition that is news to him even though his sorting behavior would reliably conform to the generalization. There are things you can do with the proposition – such as pass it along, verbally, or use it as a premise in a formal argument – that you cannot do with the sorting disposition alone. You can reflect upon it, for instance, and ask yourself what are the true boundaries of the class of the universally quantified predicate (the sort of questions the puppies cannot ask, even if their experience partially answers them *de facto*). It has often seemed innocuous to attribute general beliefs to animals on the basis of their “sorting” behavior without noting the possibility that a giant step may be concealed in this usage, the step from implicit to explicit generalization. That step may require the special thinking tools that only a natural language can provide.

What sort of mind do you need in order to acquire language in the first place? (Why will not a baboon's or chimp's mind be enough?) That is one question to which we still do not have a good answer, though the ground is littered with clues. A different question: What are the special talents of the sort of mind you acquire once you have language installed? Ray Jackendoff has recently expressed to me a striking way of answering it: we human beings can transform ourselves on a moment's notice into a somewhat different "machine," taking on new projects, following new rules, adopting new policies (personal communication, 1997). When psychologists devise a new experimental setup or paradigm in which to test such non-human subjects as rats or cats or monkeys or dolphins, they often have to devote dozens or even hundreds of hours to training each subject on the new tasks. Human subjects, however, can usually just be told what is desired of them. After a brief question-and-answer session and a few minutes of practice, we human subjects will typically be as competent in the new environment as any agent ever could be. Of course, we do have to *understand* the representations presented to us in these briefings

We are transformers – that is what a mind is, as contrasted with a mere brain. A chameleonic transformer. A virtual machine for making more virtual machines. And where are they, these virtual machines? Centered on a brain, to be sure, but not explicable without looking outside the brain into the world. Obviously some brains might be more amenable to being turned into powerful minds than others – differences in the operating system, you might say. The role that these differences might play is highly malleable, however, and may be largely eliminable. We let the myopic wear glasses; we may let the forgetful take their books along with them – why not? If we allow paper and pencil, why not a slide rule, why not a calculator, why not a colleague or a brains trust, when you get right down to it? Give me a choice between a not-too-swift Senator who has the wit and good taste to surround himself with high-powered advisors, and a brilliant do-it-myselfer who probably does not know his own limits, and I may well vote for the former. It is this distribution of the tasks of intelligence that makes our minds so much more powerful than all other animal minds.

## Notes

- 1 Andy Clark proposed (at a Santa Fe Institute workshop in May, 1996) some more stringent requirements for "robust" representation, requiring "surrogates" that are specifically "manipulated" (for instance, in "forward models" of the sort one can find in both engineering and animal motor-control systems). This can be seen to be a proposal to split the difference, in effect, requiring

- manipulanda *of sorts*, but not requiring the manipulation to be a personal-level, accessible, activity. This is a very tempting suggestion, I think, but I do not yet see how to follow it up.
- 2 Clark discusses "thinking about thinking" as a "good candidate for a distinctively human capacity" (1997, pp. 208–209).
  - 3 The scare-quotes are to remind us that maybe this is not robust, or florid, thinking at all, the sort that involves manipulation of surrogates. Some participants at the Vancouver conference took me to be claiming that there was no evidence for metarepresentation *of any sort* without the prior establishment of natural language, but my point was rather to warn against the assumption that this variety of apparently higher-order cognition is all there is to meta-representation.
  - 4 I have been challenging ethologists for several years to provide clear and persuasive evidence of episodic memory in non-human animals. The ethological and psychological literature, inspired by Tulving's (1983) work, treats episodic memory as a well-recognized category, but so far as I can see, there is a crucial equivocation in the discussions. True episodic memory is a tricky and sophisticated phenomenon. The dog that buries the bone and returns to it the next day does *not* show episodic memory; it merely shows that it has an expectation, engendered by a single episode of burying, that there will be something good to be found there. It may or may not be able to accompany its digging with an actual episode of recollection. Learning from a single episode – one-shot learning – is importantly unlike learning via lengthy conditioning (what I call ABC learning in *Kinds of Minds* and "Labeling and Learning") but it is not the same as episodic memory. We must not endow animals with a capacity for *recollection* just because they exhibit one-shot learning.
  - 5 Dan Sperber has suggested a natural ramp up which learning of this sort could proceed, beginning with what might be call impromptu observation of another (by an animal that has rich metapsychological powers to begin with), and leading by gradual steps to directed teaching, all without benefit of language. I wonder if this can be successfully modeled in a computer simulation – a nice new toy problem for artificial life.

## References

- Baron-Cohen, Simon (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.
- Brooks, Rodney (1991). Intelligence without representation. *Artificial Intelligence Journal* 47, 139–159.
- Clark, Andy (1997). *Being there*. Cambridge, MA: MIT Press.
- Dahlbom, Bo, & Janlert, Lars-Erik (1996). *Computer future*. Unpublished book manuscript.
- Darwin, Charles (1871). *The descent of man*. New York: Appleton.
- Dawkins, Richard (1982). *The extended phenotype*. Oxford and San Francisco: Freeman.
- Dennett, Daniel (1969). *Content and consciousness*. London: Routledge & Kegan Paul.

- Dennett, Daniel (1993). Labeling and learning. *Mind and Language* 8, 540–548.
- Dennett, Daniel (1996). *Kinds of minds*. New York: Basic Books.
- Humphrey, Nicholas (1986). *The inner eye*. London: Faber and Faber.
- Köhler, Wolfgang (1925). *The mentality of apes*. New York: Harcourt Brace and World.
- Mithen, Steven (1996). *The prehistory of the mind*. London: Thames and Hudson.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1: 515–526.
- Russell, Bertrand (1927). *An outline of philosophy*. London: Allen & Unwin.
- Sperber, Dan (1996). Personal communication, in setting the topic for the 10th conference of Vancouver Studies in Cognitive Science, at which the papers in this volume were presented.
- Tulving, Endel (1983). *Elements of episodic memory*. Oxford: Clarendon Press.