The rise of the metamind

Thomas Suddendorf

Cogito ergo sum – Descartes

Introduction

Every culture and epoch has had its ideas about the nature of mind and existence. We can reflect upon ourselves, upon others, and upon the world. Do animals do that too? Do they sit around and think that they are because they can think? Are they to be considered mindless if they do not reflect? How did our ability to think beyond the immediately present evolve?

The self-awareness implied by '*cogito ergo sum*', or 'I think, therefore I am', demands a reflective level of thinking that develops by about age four in children. Only then, recent research suggests, do children begin to reflect on their own mental states. It would be quite difficult, however, to convince people that younger children are mindless – mind can surely exist without being able to reflect upon its own existence. One can know, regardless of whether one knows that one knows. This means that the Cartesian assumption that the mind is necessarily transparent to the self is flawed (cf. Wimmer and Hartl 1991; Gopnik 1993). Instead, the reflective mind, or what I want to call the metamind, seems to depend on mental computations that gradually develop over the first four years of life and that have evolved over the last five million years of human evolution.

Rather than being given by God, as Descartes would have had it, metamind is the product of natural selection. It is not a basic starting block, but the product of a long process of cognitive evolution. The fundamental problem to be addressed is how, from a time when there was no consciousness on Earth, mind developed to the phenomenological experiences we have today. While physical features such as bipedal locomotion, the opposable thumb, increased cranial size, and stone-tool production mark important developments in human evolution, I maintain that it was the evolution of *mind* that takes prime responsibility for the extraordinary story of the human species.

In this chapter I will attempt to outline the natural history of the human mind, with special reference the emergence of representational skills. There are several evolutionary proposals that are in parts similar to the one that will be presented here (e.g. Lorenz 1973; Bischof 1985; Tulving 1985; Humphrey 1986, 1992; Whiten and Byrne 1991; Olson 1993; Dennett 1995). Perner (1991) proposed perhaps the most influential account of children's growing representational skills. I will extend and update Perner's developmental model and apply it to evolution. From this model one can derive various testable hypotheses about children's development and about other animals' capacities. I will show that there is a general fit to the data and that this model is very useful for understanding the basic progression in the evolution of the hominid mind.

Perner (1991) suggested that children's understanding of mind reflects their level of understanding representation in general. In brief, his theory suggests that children advance in conceptual capacity from the ability to form (1) primary representations of reality, to (2) entertaining secondary representations beyond current reality (e.g. representing past, future, or imaginary objects or events; or representing the representational content of other representational systems), to (3) understanding representational relations themselves-or in other words, metarepresenting representations *as* representations. This progression, which may seem confusing at first, can be illustrated with the example of understanding the representational nature of television. With primary representations one can merely perceive the reality of the TV image as a quickly colour-changing set of dots or as two-dimensional shapes that might be mistaken for the objects they resemble. Following these pictures in terms of the 3-D characters and events they represent means creating secondary representations (i.e., the content of the TV image is represented). With metarepresentation, one can further appreciate that this is just television (an American soap, say, with lots of bad acting). At this level, then, one can simultaneously follow the programme and evaluate whether the story is true or fictional, whether the editing is good or bad, and whether the actor always has this funny accent.

The mind can be regarded as a representational system (cf. Dretske 1995). Understanding mind might develop in the same fashion as understanding other representational systems such as televisions. For example, with primary representation one simply perceives other's actions. With secondary representations these actions can be interpreted in terms of what the person wants, intends, or pretends (i.e., in terms of not directly perceivable mental states). Only with metarepresentation, however, can one appreciate that these mental states are *just* representations. One can simultaneously entertain somebody else's beliefs and evaluate them as true or false, and wonder whether the other wants to deceive or how one could change his or her mind.

Children, animals, and human ancestors can, according to my model, be categorized in terms of their level of representational capacity (see Table 12.1). The representational level determines what an organism can mentally conceive of. With primary representations an animal can form only a *single updating model* of reality. With secondary representations the organism can entertain *multiple models.* That is to say, in addition to a model of current reality, such an organism can consider models representing past, future, or hypothetical situations. Furthermore, it can interpret the representational content of other representational systems (e.g. a picture or someone's pretend play). Because the different models can be compared I refer to this mentality as the *collating mind.* Finally, with metarepresentations an individual can form *metamodels.* The representational relations themselves can be represented (new ones can even be invented). With this capacity the individual can conceive concurrently of different ways the same object or event can be represented (e.g., by different people, by different media, by different plans, in different times, etc.). This capacity to form reflective metamodels is the cornerstone of *metamind,* which, I propose, further entails the ability to dissociate from primary perceptions and response tendencies to create a distinct level of mental executive control over actions. Metamind, I will argue, is uniquely human and at the root of humans' extraordinary position in the animal kingdom.

Table 12.1. A *model of the natural history of the representational mind*

	Primary mind	*Collating mind*	*Metamind*
Model of reality	single updating model	multiple models	metamodels
Representational level	primary representation	secondary representation, representing representations	metarepresentation, representing representations *as* representations.
Evolution	>150 million years ago birds, mammals	~ 15 million years ago great apes	~ 1.5 million years ago *H. erectus/ergaster*, humans
Development	fetus (~ 30 weeks)	end of infancy (~ 1.5 years)	pre-schooler (~ 3.5 years)
Characteristics	sensation, perception, emotion, schemes, instincts, reflexes, conditioning ...	plus secondary representation	plus metarepresentation, executive control
High expressions and applications	play, exploration, latent learning	pretence, planning, insight, self-awareness, other awareness, attribution of desire and intention	theory of mind, mental time travel, symbolic representation, generativity, creativity, teleology...
Resulting culture	no culture but nature	simple tool cultures, basic politics	mimesis, morality, religion, language, narrativity, history, science...

I acknowledge that any attempt at reconstructing the evolution of mind is inevitably an outrageous simplification. The story I will tell encompasses and connects recent findings from various research fields, to create a plausible and coherent – perhaps even true – account of the evolution of the representational mind. But a couple of hundred citations in the reference list merely scratch the surface of the complex debates in the disciplines. So even if I succeed at constructing a harmonious interdisciplinary picture, it is only one of several possible interpretations. Furthermore, to answer the question as to how we got to where we are now, I will need to employ two controversial strategies – I shall argue from currently living organisms to ancestral organisms, and from ontogeny (development) to phylogeny (evolution). Both of these strategies require some justification.

The main problem with arguing from living to ancestral species is one of deciding what features are based on common ancestry. Functionally or structurally similar phenotypes can depend on either homology or analogy. The decision between them is sometimes difficult to make, especially when it comes to behaviour and underlying cognition. For example, I observed a sub-adult male orang-utan maintaining in a horizontal position between two trees until a juvenile climbed down the trunk and used him as a living bridge. Ants can also be observed to build such bridges. Does the bridging behaviour indicate some level of self-awareness in orangs (cf., Povinelli and Cant 1995) but not in ants? The decision should be based on what I want to call *evolutionary parsimony*. Parsimony usually refers to Lloyd Mogan's canon which proclaims that we should use no higher level explanation for behaviour than is strictly necessary. Unfortunately, it is not at all clear what constitutes a higher and what a lower level of explanation. Evolutionary parsimony, however, can be defined more readily. We should favour the phylogenetic scenario that requires the least number of assumptions to explain the current phenotypes of species. For example, if all species of a particular family share a feature, it is more parsimonious to assume that they share that feature because a common ancestor had it, than to postulate that each species developed that feature independently. On the other hand, if two distantly related species (such as orangs and ants) share a feature that is not present in closer relatives, then it is more parsimonious to assume that the feature developed through convergent evolution, because otherwise one would have to assume that each of the closer related species had lost that feature during evolution. In this chapter, this reasoning will be applied chiefly in the context of similarities between humans and our closest relatives, the other great apes. Features shared between all these species are probably homologous and thus were present in our common ancestor some 12 to 15 million years ago. While analogous features can tell us something about selective pressures, only homologous features suggest that the same underlying mechanism is at work. Human characteristics not shared with our closest relatives, not even with the chimpanzees, most likely evolved after our ancestry split from the line that led to modern chimpanzees and have therefore emerged in the last five million years.

The second controversial strategy is to reason from development to evolution. Haeckel's notion that ontogeny recapitulates phylogeny has long been rejected. There appear to be as many cases violating this 'rule' as following it (e.g. Parker and Gibson 1978). However, this should not stop us from using developmental information to construct hypotheses about evolutionary history (see also Donald, this volume). From a state of no mind, a mind capable of reasoning *cogito ergo* sum evolved in phylogeny, and the same metamorphosis also occurs in the development of an individual human. For our discussion it is mainly the fact that one has to be able to form primary representations before one can form secondary representations, and secondary representations before one can form metarepresentations, that will be mapped from ontogeny onto phylogeny. My model predicts that various capacities and skills should co-emerge both in development or in evolution, because in both cases their emergence reflects the two representational transitions. Thus, recent empirical findings of developmental associations between markers of each representational level and other skills bear on the debate as to when these skills could have made their debut in evolutionary history.

Finally, the stage-like model I present does not imply strong claims about abrupt quantum leaps. The development and evolution of mental skills is a gradual process, but qualitative changes can emerge from gradual changes, just as water changes qualitatively from solid to liquid to gas as temperature gradually increases. Potentials and characteristics change in each qualitatively new level (cf. Lorenz 1973; Bischof 1985). I shall propose that, in children, the change from primary mind to the collating mind occurs at approximately age one and a half, and the change from collating mind to metamind occurs at approximately age three and a half. The corresponding changes in hominid evolution, I suggest, occurred approximately 15 million and approximately one and a half million years ago, respectively. However, the new skills might at first be crude; application and generalization might involve inconsistencies and might generally invoke a picture of gradual transition. Yet, if only to contribute to an easier understanding, I believe it to be useful to model this development as a simplified stage-like progression. Thus, I shall present the model in a three-stage manner-first discussing primary mind, then the collating mind, and finally the metamind.

I will place special emphasis on the emergence of metamind and its corollaries and present relevant findings from my own research. This change to metamind, I suggest, began in *Homo erectus/ergaster* about 1.5 million years ago. It marked the dawn of a new self-reflective force that challenged instincts and simple stimulus-response learning for the driver’s seat of behaviour. Once it emerged, the metamind was to change the face of the Earth. It is so powerful (creative, self-aware, and communicative) that it might help life itself to spread beyond the boundaries of the planet; so powerful, on the other hand, that it might destroy the whole enterprise of life on Earth.

The primary mind

Evolution of mind

Before discussing the evolution of complex representational abilities, I must briefly outline the basic assumptions about the origin of mind. During evolution more complex levels arose out of, and are ultimately dependent upon, less complex levels of organization while at the same time being profoundly different from those levels. Physical evolution gave rise to biological evolution, which gave rise to mental evolution, which gave rise to cultural evolution.

Various regulating mechanisms preceded mind, or mental functions, in biological evolution. Living organisms, by their very definition, reproduce, have metabolism, and respond to environmental change. These characteristics can be achieved by varied processes including regulating cycles and hard-wired response patterns. The entire plant world seems to run on such processes. Behaviour in early animals might also have been largely based on innate processes like today’s instincts, reflexes, and other “innate releasing mechanisms” (Lorenz 1973) that filter the environmental stimuli and select the appropriate response. Even modification of innate response patterns during an individual’s life-span can be achieved through classical conditioning, habituation, sensitization etc., and these are usually not regarded as affording anything mental (but see Lea, this volume).

Mental experience might have first emerged with the evolution of reinforceable plasticity. Dennett (1995) calls organisms with this capacity `Skinnerian Creatures’. The organism performs various actions and then selects the one `that works’ or that is `reinforced’. This is natural selection extended to the individual behavioural level. The selection process, however, is *not* within the environment but within the individual. Which environmental stimuli are reinforcing is ultimately a function of what the organism evaluates as ‘good’ or `bad’. The same stimulus may be good for one creature and bad for another; good today and bad tomorrow. But how does the individual `know’ what is good and what is bad, and why should it bother? These are surprisingly important questions that bear on the evolution of mind. Dennett (1995) suggests that some early candidates for this category `were no better off than their hard-wired cousins, since they had no way of favouring (selecting for an encore) the behavioural options they were equipped to `try out’, but others, we might suppose, were fortunate enough to have wired-in `reinforcers’ that happened to favour Smart Moves, actions that were better for their agents’ (p. 374).

But how are the ‘wired-in reinforcers’ of those lucky creatures supposed to work? In our experience something is positively reinforcing if it `feels good’. Perhaps the mind or mental experience evolved for this very reason-to represent certain stimuli as *feel-goods*, or, if they happen to be negative reinforcers, *feel-bads*. According to Humphrey (1992), evolution favoured organisms with affect-laden sensitivity (this is good versus this is bad) to events at their boundaries. Organisms evolved *interests*. Leahy (1994) traces this argument back to William James (1890) who claimed that consciousness evolved to make survival an imperative rather than a chance rule. `Minding’ created the adaptive striving for survival that marks the vertebrate world.

The legacy of Descartes' philosophy is a popular idea of mind that is far too intellectualized. As Milan Kundera (1991) put it: " 'I think therefore I am' is the statement of an intellectual who underrates toothache" (p. 200). Humphrey's (1992) alternative – `I feel therefore I am' – might come closer to the true essence of mind. That 'to feel' is the essence of consciousness is reflected in the everyday use of the word. Losing consciousness means more than being unable to think or reason, it means being unable to feel. Just as well, because being unconscious is supposed to mean that one cannot feel anything when one is cut open on an operating table.

Basic to Humphrey's (1992) analysis is the distinction between sensation (what is happening to me) and perception (what is happening out there), and the insight that sensations or feelings are the product of active neuronal processes separate from those producing perception. He cites phenomena such as blindsight and agnosia to make his case for a double dissociation between sensation and perception. The distinct activity of sensing, the logic goes, involves a reactivating loop with a particular duration in projection areas of the central nervous system, creating a subjective experience of sensation. The projection areas have a correspondence with the site of stimulation at the organism's boundary that determines the characteristic modality of the sensation. Sensation with an attached positive or negative evaluation is the fundamental non-intellectual assessment machinery which can be expected to be present in any organism with sufficiently developed sensory projection areas. Higher vertebrates such as birds and mammals are probably capable of this feat. They mind!

This assumption is substantiated by the resemblance of those animals' behaviour to the kinds of human behaviour that we ordinarily describe as based on feelings. Dogs, horses, and cats seem to *enjoy* being stroked and *dislike* being pinched, much as we do. This phenotypical correspondence is most likely of homologous rather than of analogous origin. That is, there is no apparent reason why convergent evolution, rather than ancestry, should have produced such strikingly similar phenotypes in all mammalian species. That in turn means that, just as we can assume that other people experience pain the way we do (e.g. when we see them scream while being hit), so too can we assume, with much the same certainty, that mammals (and birds and possibly some other species) experience pain when we see them `scream' while being hit. These creatures are sentient beings.

Nonetheless, the evolution and nature of feeling remain controversial. We do not need to resolve that controversy here. For my model of the evolution of representational skills it is the second kind of organismic information-processing that is important.

Perception is how the organism uses sensory information to gain knowledge about the outside world. This, of course, is of immense adaptive value. An organism that can respond appropriately to stimuli that indicate the proximity of, say, food or a predator, *before* that external object comes in direct contact with the body surface has an adaptive advantage. Organisms with highly fallible indication systems are soon starved or eaten. So, evolution favoured those with increasingly sophisticated and fine-grained categories for indicating stimuli of significant external events. Sensory organs became the means of inferring the state of the outside world. Stimulation of these specialized areas of the body surface (e.g. the retina) was used decreasingly for sensation and increasingly for perception. Some surface areas, such as the skin on our hands, maintained a double function of sensation (when being stroked) and perception (when touching objects). At specialized perception areas like the retina, however, the stimulation is no longer evaluated simply as `feels good' or `feels bad'. It is the configuration of the stimuli and what they indicate that is evaluated emotionally. Furthermore, perception does not happen to a *tabula rasa.* Animals form concepts, and concepts guide perception. These processes require an integrating central processor.

The brain processes the constantly changing incoming information to create a relatively stable picture of the environment. Size, shape, and colour constancy are formidable computational achievements that ensure that an object is perceived as *one continuing* object. Cross-modal integration of sensory information is an even greater achievement. Auditory, olfactory, and visual inputs are all taken as indicators of one and the same object in the environment. The brain creates a *single integrating* model of the outside world. All major primate groups have been shown to have cross-modal abilities (Ettlinger and Wilson 1990).

We can call this mental model, created by perception, a *representation,* since the inner and outer worlds are matched according to definable rules. Because it is an integrating model, it is not entirely stimulus-bound, as is often suggested. Indeed, a mental model might not only integrate information across modalities but also across time. If a cat were to stop hunting as soon as its prey disappeared behind a bush, it would have a very inefficient mental model indeed. This, of course, does not happen. The mental model maintains a representation of the prey as being behind the bush until there is new information displacing it (e.g. the prey re-emerges on the other side). Bischof (1985) called this capacity *diachronic identity,* as the mental model bridges past, present, and future, albeit to a very limited extent. The main limitation is that we are talking about a single model of reality that is constantly updated (cf. Perner 1991).

The ancestors of today's vertebrates were already armed with these devices of sensation, perception, and representation, and so exploited the potential design space to evolve in their diverse ways. Thus, what I present here as one basic category of mind is a very heterogenous class with very diverse degrees of sophistication, all of which are equally successful adaptations. They all share the same fundamental characteristics of mind, but lumping them into one category does their diversity an injustice. This rough treatment is inevitable, though, given the scope of this chapter and its primary focus on mental evolution in hominids.

There are two further refinements I wish to highlight because they lead up to the next level of representational capacity. One is curiosity. Some species developed active ways of improving their internal conception of the environment, and thus improving the information contained in the model. Exploratory behaviour is used to gather information about the environment, but the animal only engages in it in the absence of any competing serious motivation. A rat, for example, explores a maze and, when a stimulus elicits a flight response, uses the information gained to directly choose the most appropriate escape route. Such *latent learning* does not involve any operant conditioning but is a pre-emptive measure. The single updating mental model is enlarged to cover a wider spectrum of the current situation, including routes through the maze, by the information gathered through exploration. This is specifically adaptive for unstable or diverse environments, because it provides the information necessary for flexible appropriate responses in novel situations. It is not surprising therefore to find this trait especially in what Lorenz (1973) called the "specialists in non-specialization.... the rats from among the rodents, corvids from among the song birds, and man [sic] from among the primates" (p. 148).

Closely related to curiosity is play. Playing animals not only acquire information about the environment but also train their responses to it. Just as in exploratory behaviour, the organism usually stops the play activity when serious threats or motivations (e.g. hunger) emerge. This suggests that the action is implicitly tagged as a 'non-serious' drive only to be followed in the absence of `serious' motives. Learning activities such as play or exploration are inherently future-oriented and lead up to the next level of representational thought. But first we need to consider human development.

Development of mind

How did we develop from a non-conscious cell to a minding, perceiving, playing, and exploring infant? From the moment of birth, human infants seem to feel and to perceive the world, albeit in limited ways. Indeed, there is reason to believe that the child `minds' long before it is born. Prenatal voice recognition and other effects suggest such early developments (e.g. DeCasper and Spence 1986). While electrical activity above the brainstem starts to emerge from the 14th week after conception, it is not until the 30th week that cortical electrical activity is recorded. It is not clear when exactly the sensory projection areas of the fetal brain have matured enough to produce the reverberatory loops that, according to Humphrey (1992), produce sensation, and it will not become clear until we have identified the precise nature of these proposed neural loops in the mature brain. But it is clearly some time before birth, and in medical practice painkillers are now being used for the fetus in prenatal surgery, in apparent recognition that the sentient mind begins its development before birth (cf. Concar 1996).

There is early evidence for the existence of a single updating mental model (Perner 1991). From birth the infant seems capable of cross-modal integration, as is evident through imitation of facial expression (Meltzoff and Gopnik 1989). And by three months the time-bridging capacity becomes evident. Baillargeon has shown in several experiments that three- to four-month-old infants represent an object that is no longer perceptible (Baillargeon 1987; Baillargeon and DeVos 1991). The infants show significant signs of surprise when viewing a screen that appears to move through the space occupied by an obscured object without any signs of resistance. Apparently, the infants' model includes information about the no longer visible object, and expectations about its effect on other objects. The internal model therefore holds the information (e.g. there is this object) until it is erased through forgetting or until incompatible information updates the model (Perner 1991). The model's time-bridging capacity seems to be quite limited at first. Searching for an object hidden under one of two cloths deteriorates to chance level if the delay between witnessing the placement and starting the search is more than 8 s in ten-month-olds (Diamond 1985), and more than only 20 s in 16-month-olds (Daechler *et al.* 1976). This does not mean, however, that the infant does not store information and accumulate knowledge. Indeed, recent research has shown that young infants learn from single events and show deferred imitation (e.g. Bauer 1996; Rovee-Collier 1997).

The first two years witness an increasing sophistication in other areas of the infant's mental and physical capabilities. Piaget (e.g. 1954) aptly called this phase the sensorimotor period. Perceptions form mental concepts-or *schemas*, in Piagetian terminology. New experiences are incorporated or *assimilated* into existing schemas (applying an old schema to a new stimulus) and these schemas themselves are altered or *accommodated* to new experiential demands (adapting an old schema to a new stimulus). Piaget (e.g. 1951, 1952, 1954) described in detail the invariant sequences of sensorimotor development in regard to imitation, causality, means-ends, and object concept. Exploration and play soon make the formation and extension of the inner mental model an active preoccupation for the infant.

The development of the object concept, and other aspects of human sensorimotor development, are now being systematically studied in other primates. Indeed, a whole new research programme is devoted to this study. Parker (1990) called it *comparative developmental evolutionary psychology.* Monkeys and apes follow the same developmental sequence as humans, although they have fewer schemas, achieve the various stages at different rates, and reach different levels as their highest achievement. The general pattern of cognitive development appears to be an invariant sequence across the primate species (see, for example, Parker and Gibson 1990). Because this quasi-universal development in primates points to homology, it seems to be most plausible and parsimonious to assume that subjective phenomenological (mental) correlates of this development are homologous, also. In other words, the quality of mental experience in young infants and young monkeys might be very similar, and might have existed already in infants of our common ancestor.

However, in the second year the human infant acquires a new capacity to go beyond a single updating model to form multiple models. This has far-reaching consequences, and among the primate species this development might be shared by the other great apes only.

The collating mind

The next transition of representational capacity sets the stage for mental detachment from the immediate present. The mind goes beyond the single updating model of reality to entertain other hypothetical models. The creation of multiple models or secondary representations allows for a whole new set of skills. Throughout the discussion I will adhere to Perner's (1991) analysis of the nature of representations. In brief, a representation comprises a representational medium which represents something as having certain properties. For example, the medium might be a picture, the thing represented a house, and it might be represented as being blue (even if it's actually red).

A single updating model is based on primary mental representations. Its main function is to represent significant features of the outer world accurately. Secondary representations, on the other hand, are decoupled from the causal link to the real world. They can represent imaginary, past, or future situations. In this sense, play and curiosity might be precursors of the ability to consider multiple models. The collating mind can bring primary and secondary representations into propositional relation. That is, one can *think* of x (secondary representation) while looking at y (primary representation) and collate x and y (cf. Olson 1993). The mind can now not only feel, perceive and represent the world, but it can also *think* about things or events while feeling, perceiving, and representing the same or other things or events. This also enables the individual to interpret the representational content of other representations (e.g. of pictures, or others' thoughts). The primary perception of the pixel configuration on a two-dimensional TV screen is set aside, in favour of a secondary representation of the three-dimensional objects and events that the pictures represent. Note that these TV events are not simply mistaken for reality, although they might provoke emotional reactions and so forth appropriate to the realities they depict.

In the following section I will discuss some of the most important new skills a mind with multiple models possesses. I shall present and compare developmental and comparative data for each of these skills. My model predicts that they should all make their debut from around the same time in human development (by about one and a half years), and that our closest phylogenetic relatives, the great apes, are in principle capable of all these skills. This would support the claim that the ability to form secondary representations is a domain-general skill, and that this skill evolved before the line that led to modern apes and humans began to split, some 12 to 15 million years ago.

Several developmental and evolutionary theorists have recognized the significance of the transition to multiple models. Many names have been used, depending on the authors' theoretical positions: these include *synchronous identities* (Bischof 1985), *imagining other possible worlds* (Byrne and Whiten 1992), *metarepresentation* (Leslie 1987), *representation* (Olson 1993), *secondary representation* (Perner 1991), and *symbolic function* (Piaget 1951). Dennett (1995) coined the term *Popperian Creature,* after Sir Karl Popper, who once observed that the new skill allows hypotheses to die instead of the individual. What he meant was that multiple models can be used as mental testing grounds for behaviour rather than having to try out the real world and suffer real consequences. This, of course, is what we call insight-the search for new means to goals by mental operations rather than physical trial and error.

Insight, planning, and object permanence

Insight involves the creation of a propositional relation between the reality model (primary) and the hypothetical model of a desired goal state (secondary). Simple planning, too, can be understood as requiring the perception of the current situation (primary) and keeping a goal state in mind (secondary). These primary and secondary models need to be collated in order to conceive how one can get from the current to the desired situation. Piaget's search tests for (stage 6) object permanence might also require secondary representation (Perner 1991), but here the secondary model represents a past, rather than a future, state of the world. A classic procedure is as follows: The child watches a desired object being put into a box. The box is then placed under a rug where the object is covertly hidden. Finally the empty box is revealed to the child. With a single updating model of the world the individual only has access to the current situation. An infant does not know where the object might be. With multiple models, the individual can consider, not only the primary representation (box is empty), but also a secondary representation (past-object inside box under the rug) to create a good guess as to where the object is (under the rug).

By one and a half years most children pass these hidden-displacement tests (Haake and Somerville 1985) and show signs of insight (Piaget 1952). Planning, too, becomes evident in that the children begin to show signs of monitoring, correcting, and controlling goal-directed actions, and showing positive affective responses to mastery (Kagan 1981; Bullock and Luetkenhaus 1988). The emotional pleasure appears to be the result of achievement; the primary representation of the present reality now matches the preconceived secondary representation of the goal state.

Great apes have been observed solving problems in a way that strongly suggests insight. Kohler's (1917/1927) classic experiments illustrate this. Faced with the problem of bananas hanging out of reach from the roof, the chimpanzee apparently contemplated the situation and suddenly enacted a solution (e.g. stacking boxes upon each other) without hesitation. The chimpanzee seemed to mentally manipulate components of the situation in its imagination until hitting upon a solution to the problem. In contrast with great apes, monkeys and other animals have not yet provided convincing evidence of insightful behaviour (cf. Visalberghi and Limongelli 1994).

The same applies to planning where, again in contrast with monkeys, great apes have shown considerable skill. Dohl (1970), for example, showed that the chimpanzee Julia was able to look up to five steps ahead in a sequential-planning task. The chimpanzee was confronted with a series of two keys in transparent boxes. She had to choose the right key at the first trial to get to the right key at the second, the right key at the second to get the right key at the third, and so on, in order to finally reach the food reward. Only by working mentally backwards from the goal could the initial right key be determined. In other words, in addition to the primary representation of the situation, Julia considered the goal state (secondary representation) and worked out the steps towards the goal state *before* acting (see Suddendorf and Corballis 1997 for further examples).

While the first stages in the developmental sequence towards object permanence are commonly passed by many species, hidden displacement has been proven in few, including the great apes (Chevalier-Skolnikoff 1983; Natale and Antinucci 1989; Miles 1990). There is also evidence that African grey parrots can pass such tests (Pepperberg and Kozak 1986), but it is debatable as to whether secondary representations are implied by the particular tests used in this work (Natale and Antinucci 1989). Although monkeys can extrapolate hidden movement (Filion et al. 1996), they fail object permanence tasks (deBlois and Novak 1994). Natale et al. (1986) showed by employing `catch' trials that a macaque used simple search rules in hidden displacement tasks, in contrast to a gorilla who, like a one-and-a-half-year-old human, showed systematic search implying secondary representation of the past path of the object.

Great apes and children from about one and a half years therefore appear to possess secondary representational skill as evidenced through planning, insight, and object permanence.

The beginning of symbolism – pretence, pictures, and language

Pretending that one object is another entails secondary representation because the object of perception (primary representation) is treated as if it were a different object (secondary representation). This requires decoupling from the primary representation (Leslie 1987). Further, in representing what someone else is pretending, one is interpreting their actions in terms of secondary representation; Leslie refers to this as metarepresentation, but in this chapter I reserve that term for the ability to represent representations as representations (cf. Pylyshyn 1978; Perner 1991; Astington 1994). As already noted, metarepresentation in this original sense does not emerge until about the fourth year of life-even in the realm of pretend play (Lillard 1993; Jarrold et al. 1994; Suddendorf et al., in press). But already in earlier pretence, the child has to create secondary representations (the as if situation) in addition to the primary representation (real situation). This behaviour begins by about one and a half years (e.g., Leslie 1987).

There is some controversial evidence suggesting that great apes can pretend. Sign-trained chimpanzees, orang-utans, and gorillas have been reported to engage in pretend play with dolls (Gardner and Gardner 1969; Patterson and Linden 1981; Miles 1990). Patterson and Linden claim, for instance, that the gorilla Koko frequently pretends that one of her plastic alligators has `real' properties, and uses it to `frighten' her human caregivers (Patterson and Linden 1981). There are also some anecdotes of imaginary toy play (e.g. Hayes 1951; Savage-Rumbaugh and McDonald 1988). Although these examples all involve zoo or home-reared animals, they are still evidence that apes can engage in such behaviours. Other species do not seem to pretend. A cat's prey-catching behaviour with a ball probably does not constitute pretence, because it is stimulus-elicited and inflexible (Whiten and Byrne 1991).

Pictures usually represent something, for example the characters and events in a story. Understanding what they represent, therefore, involves representing representational content. While infants in their first year treat pictures as just a piece of colourful paper, during the second year pictures develop a magical attraction. The child can now interpret pictures in the sense of forming a mental model of the depicted situation (Perner 1991). This requires multiple models, because the true current situation (e.g. mum holding the family photo album) and the one pictured (e.g., mum at a beach holiday) have to be differentiated. A mental model of a depicted situation (photo, drawing, or TV), like a model of a pretended situation, is a secondary representation. Great apes interpret pictorial information appropriately and show interest in videos and picture books (e.g. Premack and Woodruff 1978; Patterson 1991).

The interpretation of symbols such as words or signs also begins during the second year. This also might be based on the emerging ability to entertain secondary representations in addition to the primary representation of the sound or the sight. Attempts at teaching signs to animals have resulted in moderate successes in a few species. Chimpanzees, bonobos, gorillas, and orang-utans have been successfully taught production and comprehension of words (signs), often numbering in the hundreds (Gardner and Gardner 1969; Greenfield and Savage-Rumbaugh 1990; Miles 1990; Patterson 1991). Limited success has also been reported in dolphins (e.g. Herman et al. 1993), sea lions (e.g. Schusterman et al. 1993), and African grey parrots (e.g. Pepperberg 1990, 1993). The most linguistically competent animal so far appears to be the bonobo Kanzi, whose capacity for language has been assessed as approximating that of a two-year-old human (Savage-Rumbaugh et al. 1993). Other aspects of early symbolic understanding such as classification and negation (cf. Olson 1993) or the ability to deduce word reference from the speaker's focus of attention (cf. Baldwin 1993) might also be facilitated through secondary representational skills and great apes might be quite capable in these respects.

Self awareness: mirror self-recognition

Reflections are similar to pictures, and the development of an understanding of reflective surfaces might parallel that of understanding pictures. Reflective surfaces can be used to discover what one looks like. Researchers have developed a formal test to assess mirror self-recognition (e.g. Amsterdam 1972; Gallup 1970). Individuals pass the test by retrieving a mark, such as a sticker or rouge covertly placed by the experimenter, from their faces while viewing their mirror image. This test has become a standard tool for the non-verbal measurement of an early cognitive understanding of self.

An understanding of the properties of a mirror is implied by the realization of the identity between directly and indirectly (mirrored) perceived objects (object match) and between directly and indirectly observed actions (event match or contingency testing). Self-recognition is a special case of object and event matching. Parts of the visible body (e.g. hands) match their mirrored counterparts, and proprioceptive and visual information about action also correspond with their mirrored equivalents. Such identification, or what Bischof (1985) called *synchronous identity,* seems to require secondary representation because one event or object has to be held in mind (secondary representation) while the perceptual system engages with the other (primary representation), in order that the relation between the two can be recognized.

The mirror mark test provides stronger evidence for the presence of secondary representations than the mere use of mirrored information, because the latter can be based on associative learning. For mirror self-recognition the individual has to hold in mind a picture of what the reflection `ought' to look like, based on past experience (secondary), in order to realize that the mirror image it currently perceives (primary) is different. The reflection might, for example, show an unexpected red dot on the forehead. Only if this discrepancy is noted does it make sense for the individual to investigate its own forehead in search of the dot. Thus, passing the test is evidence for secondary representation and the existence of some kind of mental image of self (as seen from the outside). In human development the mark test is passed at about one and a half years and co-emerges with self-recognition in photos (Amsterdam 1972; M. Lewis and Brooks-Gunn 1979). Some argue that it marks the onset of autobiographic memory (Howe and Courage 1993, 1997), but I argue against this in the metamind section below.

Over the last twenty years the mirror self-recognition test has been used to study a wide variety of animals ranging from birds to dolphins (Parker *et al.* 1994). Macaques (Anderson 1986), elephants (Povinelli 1989), and parrots (Pepperberg *et al.* 1995) have all been shown to use mirrors appropriately, but only great apes have been shown to pass the classic mark test. This includes at least some chimpanzees, bonobos, orang-utans, and gorillas (Gallup 1970; Suarez and Gallup 1981; Patterson 1991; Hyatt and Hopkins 1994). Event-matching is clearly present in these species. Menzel *et al.* (1985) produced additional evidence that chimpanzees, but not macaques, could relate indirect perception and proprioception, because they could use a video image to guide their hand movements. Interestingly, dolphins, while not testable in the classic way, might also engage in contingency testing (Marten and Psarakos 1994, 1995).

Other-awareness – synchronic imitation, empathic behaviour, and mental attribution

To be aware of the perspective of others, or other-awareness, might entail secondary representations. Another individual's perspective is not part of the perceptual field but is an inference, a propositional relationship between a secondary representation and a primary representation. Other-awareness and self-awareness emerge together, presumably because they both depend on the ability to form secondary representation (cf. Asendorpf *et al.* 1996).

During the second year other-awareness in children becomes evident in various ways, including self-consciousness while the centre of another's attention (M. Lewis *et al.* 1989), cooperation with peers (Brownell and Carriger 1990), prosocial behaviour towards victims of distress (Zahn-Waxler *et al.* 1979), and communication through synchronic imitation (Asendorpf and Baudonniere 1993). The last two have been found to be strongly associated with self-recognition (Bischof-Kohler 1989; Asendorpf and Baudonniere 1993; Asendorpf *et* al. 1996). At the same age, children also begin to attribute mental states to others (e.g. Dunn 1991). For example, they might re-enact what an adult (but not an inanimate object) seemed to *intend* to do rather than what the adult actually did (Meltzoff 1995). Early mental attribution is about desires and intentions (Wellman 1990). Because mental states are not observable, they cannot be derived from direct perception. Rather, they result from cognitive processes relating conception (secondary) and perception (primary) in the observing individual. Secondary representations are therefore critically important for the attribution of intentionality.

Of course, most animal species must have some awareness of the presence of others. But evidence for other-awareness based on secondary representation is again strongest for the great apes. An example in chimpanzees was reported by de Waal (1982)-a subordinate male who displayed sexual interest to a female immediately covered his erect penis with his hand when he noticed that a dominant male was approaching. Macaques, on the other hand, failed to show an understanding of another's perspective even after training (e.g. Kummer *et* al. 1996). A collection of anecdotes of primate tactical deception (Whiten and Byrne 1988; Byrne and Whiten 1990, 1992) contain many examples of self-conscious behaviour shown by great apes when they were the centre of others' attention.[1] The record for lesser apes and monkeys is much scantier and more controversial.

In deception, other-awareness is used to take advantage of another individual. It might also be used for mutual advantage. Cooperation is of course common throughout the animal (and even the plant) kingdom. In most cases, however, it takes the form of long-term symbiotic behaviour. It is short-term cooperation in innovative problem solving that is more likely to involve some kind of perspective taking. There are many records of such cooperative innovation in chimpanzees (e.g. Kohler 1917/1927; Menzel 1974; de Waal 1989), but without a clear behavioural definition of when cooperation entails taking the perspective of the other, such anecdotes cannot be cited with confidence as evidence of secondary representation skills. Experimental work by Povinelli and colleagues (Povinelli *et* al. 1992a, b) showed that chimpanzees, but not rhesus monkeys, were capable of role reversal in a cooperative task, and the authors argued that only the chimpanzees showed `empathy', that is, understood their partner's role – although Heyes (1993) has provided an alternative account of these data, based on associative learning.

There are very few records of spontaneous empathic behaviour in animals. There are the classic tales of dolphins helping humans in distress. Some credible accounts relating to great

---

[1] At face value, the anecdotes of tactical deception in great apes appear to suggest metarepresentational capacities (cf. Whiten and Byrne 1991). I argue, however, that the difference between monkey and ape deception can be explained on the basis of the difference between a primary and a collating mind. Since writing this chapter my attention has been drawn to a paper by Whiten (1996) in which he re-evaluates the deception data. Independently, considering imitation, pretence and mindreading, he comes to the same conclusion that great apes can entertain secondary representations, but not metarepresentations.

apes have been published. Washoe, the Gardner's home-reared chimpanzee, saved a young chimp, who had fallen into a moat, from drowning (Fouts and Fouts 1993). Boesch (1992) has reported evidence for chimpanzee compassion and empathy from the field. Most recently a gorilla saved a three-year-old boy, who fell 15 feet into the gorilla enclosure at Chicago's Brookfield Zoo, by carrying him to a door and alerting zoo-keepers. Although controversy still surrounds the issue, O'Connell (1995) reviewed the evidence and concluded that chimpanzees do have empathy.

Opinions about imitation in animals have changed dramatically in the last decades. Behaviours that were formerly considered as clear examples of imitation have recently been explained in terms of social learning processes such as stimulus enhancement and social facilitation (Meador *et al.* 1987). The most frequently cited case is the learning of song by birds, but this behaviour seems to be based on an innate program enabling direct comparison of the individual's own sound production with the memory trace (`tape recording') of the model sound (e.g. Lorenz 1973; Byrne 1994). Imitation of visually perceived behaviour might be significantly more difficult, because motor tasks might look very different from different perspectives. Indeed, it might be necessary to take the perspective of the model if one is to compare one's own behaviour with that of the model, particularly if imitation is to occur *syn-chronously* with the model behaviour. Contrary to common belief, there is *no* convincing evidence for imitation in monkeys (Cheney and Seyfarth 1990; Whiten and Ham 1992). The evidence for `aping' in great apes is also scant, but more convincing (Goodall 1986; Meador et al. 1987; Byrne 1994). Dolphins also seem to be capable of imitation (Herman et al. 1993).

Great apes have also provided evidence for mental attribution of intention and desire. Indeed, the whole enterprise of studying the development of `theory of mind' stemmed from a seminal article presenting evidence for the attribution of intention in a chimpanzee (Premack and Woodruff 1978). The conclusions remain controversial (cf. Heyes, 1998, and commentaries), but experimental (e.g. Povinelli et al. 1998a) and observational (e.g. Whiten and Byrne 1988) data suggest at least limited understanding of mind in great apes-perhaps comparable with that of a two-year-old child.

Summary and conclusion

In accordance with my model, comparative and developmental data seem to converge across skills. Great apes show evidence for a collating mind through their apparent knowledge about self and others, and skills such as pretence, planning, insight, sign-language learning, and mental attribution. In human development the child begins to display these capacities during the second year. Thus, it was argued that children of this age and great apes have collating minds.

To many developmental and comparative psychologists, this kind of generalization is a challenge. They are eager to show that much younger children can pass this or that test, if it is appropriately simplified, or that some bird or insect can solve the problem, if the test uses species-relevant stimuli. Despite this tendency, and the accumulation of some evidence for gradual improvement, I maintain that these changes, in the second year of human life, and between the monkey and ape lines of descent, can best be described as a qualitative shift in representational capacity. The capacity for secondary representation shows itself in a range of domains, but these domains themselves might have their roots in earlier developments. Thus, precursors of the later skills might suggest more gradual development.

The proposal does not, of course, exclude the possibility that some other lines of descent have also evolved collating minds, and strong arguments might be made for secondary representation in large-brained birds, such as parrots, and in aquatic mammals, especially dolphins. Any ape-like abilities in these species, however, must be the result of convergent evolution. But among the primates, the fact that all our closest sister species, the great apes, show evidence for a secondary representation in all of the respects discussed above is strong support for Darwinian evolution of mind. Evolutionary parsimony suggests that our common ancestors had already evolved the basis of a collating mind some 15 million years ago.

Metamind

During the fourth year children seem to change quite dramatically in the way they see the world, others, and themselves. Parents observe that their children begin to make their own plans, have their own long-term goals (such as what they will do when they grow up), recall what one told them last week (especially when it contradicts today's explanation), consider other people's minds, deceive and lie, restrain themselves, start to read and to follow complex story lines, invent their own stories (generating entirely novel scenarios), their own symbols, and perhaps an imaginary friend, draw moral conclusions, and actively ask *why* and *what for* in their attempts to make sense of the world. While these pre-schoolers are clearly just children, prone to all sorts of mistakes and silliness, we adults notice, often with pride, their first precocious attempts at reason and reflection. A fascinating new realm has opened up to the child. Almost out of the blue, the young person's mind is completely and without a doubt far beyond the reaches of any animal's mind. What happened?

I propose that the child has developed a *metamirzd.* A key aspect of metamind is the ability to metarepresent. Metarepresentations, according to Pylyshyn (1978), are representations of representational relations. Representations themselves can now be represented as representations (Perner 1991). This has far-reaching implications. Perner emphasized that to understand (i.e. metarepresent) something (e.g. a picture) as a representation one has to master the distinction between what something represents (e.g. my house) and how it represents it as being (e.g. the house is blue in the picture although it has since been painted red).

Through making this distinction one can form higher-order predicates such as `is true' or `is false' about the representation (cf. Olson 1993*).* Many different representations (e.g. an old picture, my memory, your memory, or an `artist's impression' from a real-estate advertisement) might refer to the same object or event (e.g. my house) and represent it in very different ways (e.g. as new, beautiful, and blue, while it is truly old, shabby, and red). With metarepresentation, then, representational relations can be tagged with predicates (e.g. your view; my memory), and the individual can now simultaneously entertain several distinct representations of the same object or event in reality without running into paradoxical conflicts. It is not clear, however, whether the relationship is not the other way round. It might be that the ability to entertain various conflicting representations simultaneously is what enables the individual to create representations of representational relations (i.e. metarepresentations). Here, I will treat both abilities as the two sides of the same coin. With metamind the child can appreciate representations as representations and can entertain various conflicting representations of the same object or event. It follows that the child can now understand that people's minds, as well as pictures, words, and other representational systems, might represent the same world in different (e.g., true, false, exaggerated, imprecise, outdated, wishful ...) ways.

This is also important for reflections on one's own mind. One can be wrong. Things are not always what they seem. A distinction between appearance and reality requires one to represent the same object or event in two conflicting ways simultaneously (e.g. what it looks like and what it really is). With a metamind, children can grasp that their own primary representations might be false or distorted, whereas to the collating mind the primary representation defines reality, even though the secondary representation might deviate from it.

Metamind enables the individual to entertain various ways of looking at the same thing-representing what it is, looks like, was, could be, should be, and so on. Consideration of these various perspectives allows for a greater, and more informed, choice of behavioural options. On a social level, they help the individual to understand that other views are taken as the truth by other representers. The child can now understand that people will search for a desired object where they think it is, not necessarily where the child knows it is. And, as will become evident in later discussion, these examples only scratch the surface of the vast number of skills that are dependent on the representational advance to metamind.

The term *metamind* comprises all those kinds of thought and reasoning that are based on metarepresentational understanding. At times, metamind can be somewhat disengaged from immediate primary representations. It can `wander off', as it were, and entertain a variety of propositions. It is the stage for complex reasoning, considering `what if?', theorizing, reconstructing the past, and planning the future. Metamind enables the cognitive apparatus to function off-line (cf. Bickerton 1995). I am referring to that part of our mind that is dreaming or reasoning while the on-line processes are, say, driving the car. We can be so far removed from the on-line processes that we sometimes wonder how we drove to where we are now, with no recollection of the journey.

The second key aspect of metamind is an increase in executive control. It is important to note that metamind is only adaptive because the results of this disengaged thinking *can* be related to current perceptions. The conclusions drawn from these mental exercises can positively affect survival and procreation. Instead of acting simply on the basis of the current situation (i.e., what seems to be true-according to primary representation), the individual can now cognitively base behavioural decisions on what was, what could be, what should be, and what might be true, as well as on what seems to be true for others. In order to benefit from these considerations the individual has to be able to (at times) suspend 'lower-level' impulsive response patterns.

Metamind requires an ability to disengage or dissociate from the immediate response to perception. Mental computation increasingly demands priority in determining behaviour. Thus, executive control is crucially important to any adaptive advantage of the metamind. Of course, even we adults are not entirely rational and are often guided by immediate impulse rather than reason. Metamind emerges as a new addition to a cognitive apparatus that is already sophisticated. Instincts or 'gut-feelings' might sometimes be better than our best reasoning processes at assessing what is good or bad for us. However, large parts of our lives and our culture are based on our capacity to override immediate behavioural predispositions and base our decision on our reflective metamind.

In sum, metamind comprises advances in representational and executive abilities. Metarepresentation enables the individual to form higher-order predicates such as `is true' and `is false' or `your belief and `my belief. Various conflicting representations of the same object or event can now be entertained concurrently. These representations can be brought into complex relations, and reasoning with and about things well beyond the immediately present becomes possible. For this ability to unfold fully, the individual must learn to disengage from the preoccupation with immediate perceptions. Metamind enables a more accurate and present-transcending modelling of reality. If it is to be of any adaptive advantage, metamind must be able to override impulsive response patterns and replace them with responses based on reflective reasoning. In other words, the mind increasingly exerts executive control.

While there is ample evidence that our closest relatives, the great apes, have also evolved a collating mind, there is a marked lack of evidence that they possess the key elements of metamind. Although they have the rudiments of many human capacities, it seems that they lack precisely those aspects that catapult these capacities to new heights, beginning in humans at about age four. However, it is of course impossible to prove that a species is *not* capable of X or Y. After all, it takes only one individual at any one point in time to disprove the claim. Nevertheless, in the light of current evidence it is most parsimonious to assume that none of the great ape species, or any other animal for that matter, has developed a metamind (cf. Heyes 1993, 1998). Indeed, naturalistic observations confirm that despite the great apes' remarkable skills, they did not invent morality, history, music, justice, art, religion, language, or any other human cultural universal that, as will be seen, implies metamind.

The achievements of the growing cognitive apparatus between ages three and four have been studied intensely over the last decade, but authors have emphasized different aspects. Some have referred to the `inner eye' (Humphrey 1986), others talk of `theory of mind' (Premack and Woodruff 1978; Wimmer and Perner 1983), `mental time-travel' (Suddendorf and Corballis 1997), 'second-order intentionality' (Dennett 1978), or `executive function' (Russell et al. 1994; Zelazo et al. 1996). The achievements in all these domains, I suggest, are based on metarepresentation and increased executive control. That is, the concept of `metamind' embraces the reflective self-reference of `inner eye' (an introspection organ), the social and abstract component of `theory of mind' (a mind-reading organ), the temporally extended sense of reality implied in `mental time-travel' (autobiographic memory and remote future plans), the active association of 'second-order intentionality' (action plans and strategies), as well as the regulatory aspect of `executive function' (the mental government of behaviour). I shall provide logical and empirical grounds for bringing together all these changes that occur at the transition between three and four years of age under the single heading metamind.

It is the theory-of-mind aspect that has received most attention over the last decade. `Theory of mind' refers to the explanation and prediction of behaviour based on the attribution of mental states such as intention, knowledge, or belief. There is a fast-paced on-going debate about such questions as whether the underlying mechanisms are innate or socially constructed (e.g. Carruthers and Smith 1996), whether imitation or pretence is the precursor (e.g. Moore 1996), or whether theory of mind develops through simulation or through theory construction, but the debates are really about the relative importance of these factors (Astington 1994). In respect of whether theory of mind is based on simulation or on theory construction, for example, it is clear that adults can use both strategies-they can reason abstractly about, say, the probable mental state of someone they have betrayed, but they can also gain further insight into what the other might feel by mentally putting themselves into his or her position. Thus, instead of entering these debates (see Carruthers and Smith 1996), I want to look at the overall interconnected change that occurs at age four. It is now time to put empirical `flesh' to the theoretical bone structure I have presented so far.

'Theory-of-mind' tests of metamind

Ever since Wimmer and Perner's (1983) pioneering research on mental attribution, understanding false beliefs has been the crucial test for metarepresentation and `theory of mind'. People act according to how they represent the world, rather than according to how the world actually is. In the case of true beliefs, representation and reality are identical, but in the case of false beliefs they differ. Only the attribution of false beliefs can therefore unequivocally reveal an understanding of the representational nature of mind (Dennett 1978; Wimmer and Perner 1983). False-belief attribution also implies the other metamind skills, namely, the ability to dissociate or disengage from the immediate perception (e.g. regardless of where a desired object truly is, the person will look where she thinks it is) and to simultaneously entertain two conflicting representations of the same object or event (e.g. I know the object is here but she thinks the object is there).

At about the same age that children become able to ascribe false beliefs to others (by about three and a half to four years) they become able to attribute false beliefs to themselves. Gopnik and Astington (1988) showed that younger children fail on tests of representational change-when their belief was changed, they reported having held the current belief all along. When asked what they thought was inside a candy box before they were shown that there were pencils (not candy) inside, they stated that they originally thought there were pencils inside. Again, in order to entertain the past false belief (that there was candy in the box), the child has to be able to metarepresent the beliefs as true and false (or current and past) and disengage from the current knowledge about the true state of affairs.

Flavell and colleagues (e.g. Flavell *et al*. 1986) showed that children begin to distinguish between appearance and reality at around the same time. Younger children have problems understanding the difference. Pouring milk into a blue glass does not change the milk's colour but only our perception of it. To differentiate between the two, the child has to establish a relationship between the propositional relation held in mind and the perceived situation (Olson 1993*)*. The two representations (e.g. `is blue' and `is white') have to be metarepresented or tagged as `looks like' and `is truly' and the child has to disengage from the current perception (blue) to answer the question about the reality (white).

One might be concerned about whether young children simply fail these tasks because they do not know what the researchers mean or want (e.g. Siegal 1995). Flavell *et al*. (1987) showed in various ingenious ways that children's conceptual problems are genuine. For example, a white card was held under a blue filter and the experimenter detached a pre-cut piece from the card and presented the child with this white piece and a corresponding blue piece. Asked which piece was taken from the card three-year-olds tended to point to the blue piece, thereby failing to recognize the distinction between appearance and reality. Passing false-belief, representational change, and appearance-reality tasks have been found to be correlated (Gopnik and Astington 1988; Moore et al. 1990), substantiating the notion of a common underlying representational mechanism (e.g. Flavell 1993).

Another aspect of how the representational mind works is the basic relationship between informational access (i.e. perception) and knowledge. Children younger than three and a half do not fully comprehend that, for example, seeing leads to knowing (Wimmer et al. 1988; Perrier and Ruffinan 1995; but see Pratt and Bryant 1990). Again, children need to metarepresent the other's perception (or ensuing representation) and to disengage from their own perception in order to reason about what is and what is not available to the other's mind. Further, this seems to be true also for the reconstruction of how information entered one's own system. Asked how they know what they know (e.g. being told, having seen, or having inferred the location of an object) three-year-olds fail to respond above chance (Gopnik and Graf 1988; O'Neill and Gopnik 1991; Woolley and Bruell 1996). Indeed, even four-year-olds tend to claim that they have always known what in fact they have just learned today (Taylor et al. 1994). Source memory also requires that the current mental state is set aside and a different past state is entertained.

Like conflicting knowledge and belief, conflicting desires seem to be understood only by about age four (Moore *et al*. 1995). Younger children ascribe desires and intentions to others. But, when children younger than five hold a strong desire they seem unable to recognize that another person might desire something different. Their own desire might be too overwhelming to dissociate from. If there is no strong desire on the part of the ascriber, the attribution does not pose a problem. Older children, through metarepresentation, can tag the desires as `mine' and `yours' and the conflict is resolved since both desires can be simultaneously entertained.

Once again, a task that is similar except that it involves the self rather than others produces comparable results. Zelazo and colleagues (e.g. Frye et al. 1995; Zelazo *et* al. 1996) found that changing the rules in a card-sorting task poses a problem for three but not for five-year-olds. For example, children had to sort cards first according to shape and then according to colour. The younger children continued to sort according to the first rule. They needed to disengage from the first rule (or desire to order by shape) in order to sort according to the second rule. Understanding embedded rules (i.e. use rule 1 in game A and rule 2 in game B) might also require representing representational relations (cf. Zelazo and Jacques 1997). The younger children showed evidence of understanding both rules, but they could not use this knowledge to guide their action. Similarly, there is evidence that even three-year-olds can remember their earlier false beliefs (Freeman *et* al. 1995 used cued recall to elicit such memory), yet without support this knowledge is not accessible for action control.

Some recent research suggests that in particular circumstances children's performance on theory-of-mind tasks can be enhanced (e.g. Chandler and Hala 1994; Saltmarsh *et* al. 1995). Mitchell (1997) argues that younger children's problems with theory-of-mind tasks are due to of an attentional magnetism to a tangible reality. Thus, theory-of-mind tasks that involve the child and provide tangible objects corresponding to, say, false beliefs, can be passed even by three-year-olds. In effect, these easier theory-of-mind tasks impose less of a demand on the child's ability to disengage. Recent research by Clements and Perner (1997) suggests that children might have some implicit understanding of false beliefs before developing it explicitly. In any case, the point of the research was *not* to show that three-year-olds lack a theory of mind; rather, it was to find a way of *showing* that a child might have a representational theory of mind (Dennett 1978). When children reliably pass false-belief tasks we know that they understand that people act according to their representations of the world.

Evidence from animals has proven elusive, because few theory-of-mind tests have been successfully adapted for comparative research. Povinelli *et* al. (1990) argued that chimpanzees recognized the distinction between a `guesser' and a `knower', and this result seemed to have promising implications for chimpanzees' capacity to recognize the relationship between perception and knowledge. Further studies, however, proved that the design did not yield evidence for mental attribution performance even in humans (Gagliardi *et* al. 1995), and that young chimpanzees are ignorant of the fact that seeing leads to knowing (Povinelli and Eddy 1996). Premack (1988) provided the only attempt to test false-belief attribution in a chimpanzee, and the single subject in that study failed to provide evidence. Premack concluded that chimpanzees do not have a fully-fledged theory of mind (Premack and Dasser 1991; see also Heyes 1998). What they appear to lack is a metamind.

Metamind and social knowledge

Being able to understand that others have minds is, of course, of tremendous value for all kinds of social interaction. By age four, children have developed an impressive basic understanding of the workings of mind. Contrary to the impression gained from the literature, the particular theories of how mind and world interact vary profoundly between cultures. Lillard (1998) identifies four basic types of variation between different folk psychologies: "magic, unmandated conceptual distinctions, denial of the negative, and varying values" (p. *23).* Vast differences can also be observed within western culture. Some groups, for example, believe in mental powers such as telekinesis or telepathy. Others believe in spirits and in `god' having access to one's private thoughts. Children acquire their `theory of mind' within these socio-cultural contexts. While there is great variation between these theories, the basic capacity for generating such concepts about the mental (i.e., having a metamind), I argue, is a human cultural universal.

In general, others' beliefs, intentions, and knowledge are great predictors of their behaviour, and cooperation as well as deception become far more effective if one knows what is on the other's mind. Indeed, changing others' behaviour is usually most effective when one changes their minds. Research (mainly in western cultures) has shown that when children pass theory-of-mind tasks, they improve in their social understanding. Lalonde and Chandler (1995), for example, found that teachers rated children who pass false-belief tasks to be higher in social-emotional maturity than those who fail. The new level of social understanding also shows in particular areas such as deception (Peskin 1992; Ruffinan et al. 1993). With metamind children begin to appreciate the subtler and more sophisticated aspects of the social world. The vast majority of our stories, histories, and fairy tales, for example, revolve around mental states like intentions, knowledge, false beliefs, surprise, betrayal, selflessness, goals, deceits, and morals. Only with metamind can such narratives be truly understood and the lessons inherent in the stories be learned (cf. Peskin 1996). Indeed, cultural learning in the sense of instructed learning, verbal self-regulation, and collaboration depends on metamind skills (cf. Tomasello et al. 1993). It is not surprising then that many researchers emphasize the social effect of metamind skills (e.g. Wimmer and Perner 1983; Dunn 1991; Byrne and Whiten 1992; Astington and Jenkins 1995; Baron-Cohen 1995). For a thorough discussion of the development of social understanding, see Barresi and Moore (1996).

The hypothesis of Machiavellian intelligence, which postulates that human intelligence was born out of increasing social intelligence in primates (Jolly 1966; Humphrey 1976; Byrne and Whiten 1988, 1992), fits in nicely with these effects. However, even if its origin was social, metamind must have had significant non-social side effects. These have been largely overlooked. To balance the score I want to emphasis these non-social effects here.

Metamind and self knowledge

Understanding other minds is bound to go hand in hand with understanding one's own mind. If one assumes that other minds are understood through simulation or putting oneself into the other's shoes (e.g. Harris 1991), then understanding one's own mind must come first.

Self-awareness can be viewed from a functional or structural perspective (Gibson 1995). The origin of self might lie in functional developments (e.g. experience of agency). Perceived control is of paramount importance, and primates prefer tasks in which they experience control over tasks where rewards are received independently of own action, even if the actual reward is greater in the latter tasks (e.g. Rumbaugh *et al.* 1994). Interactions with the environment, especially the experience of agency, give rise to knowledge about the agent (the self), its features, and competencies. This amounts to what from a structural perspective is `the' concept of self. This idea of self, me, I, or soul, is universally generated (cf. Brown 1991). I am not arguing that there is such a core entity, far from it, but it is important to explain why, how, and when children form the idea we might structurally call a personal identity. In the following I will present several ways in which metamind functionally influences aspects of self which add to the new structural conceptualization of self.

As we have seen already, the attribution of mental states to self appears to co-emerge with the attribution of mental states to others (Gopnik and Astington 1988). By age four children begin to know that and what they know, believe, want, etc. With this reflective thought we have the transparency necessary for Descartes' cogito. The child can now reflect on her own mental states and can potentially come to the conclusion: I think therefore I am. But who or what is the structure of this I?

The ultimate question of where we come from, what we are, and where we are going (Humphrey 1986) can now, at least in principle, be addressed by the child. I have argued elsewhere (Suddendorf 1994; Suddendorf and Corballis 1997) that episodic memory is an active reconstructive process that depends on metamind skills. To travel mentally into one's past and inspect one's history requires mental disengagement from the present, tagging of tense (i.e. `pastness') to the representation, and active metarepresentational reconstruction of episodes. Past mental states need to be appreciated and the source of current knowledge needs consideration in this process. Perner and Ruffman (1995) found evidence for an association between free recall and understanding the relationship between knowing and perceiving. Following Tulving (1985), they reasoned that episodic memory is based on autocueing through episodic traces which is reflected in free recall but not cued recognition. They found a robust correlation between free recall and perception-knowing tasks that was independent of variations in age and verbal intelligence. This substantiates the notion that episodic memory emerges with metamind (cf. Suddendorf and Corballis 1997). Recent neuropsychological work links this with the prefrontal cortex (Wheeler et al. 1997).

This new function, episodic memory, has an impact on self conceptualization in that it is the origin of autobiographical memory: `that episode is what happened to me' (but see Howe and Courage 1997). Personal identity depends on a personal history. The development of autobiographical memory is often described as a social interactive process (e.g. Nelson 1992) suggesting gradual improvements. Without denying that social effects are important (Welch-Ross 1995), the metamind model predicts a sharper transition. Events that occurred before metamind are not part of autobiography, whereas events after this transition can be. As there is an age-correlate, the model predicts that four but not three year-olds would create an episodic memory of a salient event that is retrievable years later. Pillemer et al. (1994) recently provided empirical support for this assertion. Two groups of pre-schoolers (mean age three and a half and four and a half) were interviewed two weeks after a school evacuation in response to a fire. All children had some memory of the event. Seven years later, only children of the older group produced an accurate narrative memory of the event. Only 18 percent of the younger children were able to produce even a fragmentary memory while 57 percent of the older children did. With the rise of metamind around age four, childhood amnesia ceases and one's personal history (i.e. autobiographical memory) begins (Suddendorf 1994; Perner and Ruffman 1995; Suddendorf and Corballis 1997).

But the platform of metamind is not only used to reflect on what is and what *was*, but also on what will or could *be.* That is, the same mechanism is used to extrapolate from the past so as to conjure up scenarios of the future (Suddendorf 1994; Suddendorf and Corballis 1997).[2] Indeed, the mechanism's selective advantage must lie in its benefits for future survival, not in reminiscence per se. Secondary representation makes limited planning possible, but the individual is tied to the present, as primary representations alone represent reality. This might be especially limiting with respect to one's own drives and needs. Without anticipation of future needs and drives there is no point in imagining a future more remote than the satisfaction of current needs. With such anticipations, on the other hand, it seems imperative to secure not only the fulfillment of present but also of future needs. The idea that only humans have developed the capacity to anticipate future needs and drives is called the Bischof-Kohler hypothesis, and might be illustrated by the, admittedly over-simplified, claim that "while a full-bellied lion is no threat to nearby zebras, a full-bellied human may well be" (Suddendorf 1994,

[2] Since writing this chapter I have been made aware of a manuscript by Bischof-Kohler (in press) in which she makes a similar argument and reports having found supporting correlations between the development of theory of mind and future-oriented behaviour (see also Moore *et al. in* press).

p. 45). Humans, presumably from about age four onwards, can consider possible or likely future states of need and alter their current behaviour in accordance with these anticipations.

The agent (the self) travels mentally in time (Suddendorf and Corballis 1997). A personal identity through time is the natural consequence by which the past, present, and future can be united under one umbrella. This is necessary to make the mental representations of past and future relevant to the present acting self. While we are changing dramatically in physical and mental make-up over time, all these stages are considered aspects of the same *me.* An extraordinary new structure appears to emerge: a sense of self that is not bound to time and body.[3]

With understanding and reflecting upon own mental states also comes the ability to form attitudes about oneself. Self-esteem is an important aspect of the new personhood, and early developments might have long-lasting effects on a person's life. More specifically, the child can now form attitudes about particular mental states such as: `I don't like myself bossing others around' or `I don't want to know'. Metamind includes self-assessment and self-control. Metavolition (e.g. `my desire to play will not interfere with my concentration on work') makes the child capable of governing its own motives (cf. Frankfurt 1988). Perner (1991) submitted that the observation that children around age four become `reasonable' might not so much be a result of them becoming `logical' (in concrete operational thinking) but of them becoming rational about their own desires.

Mischel and colleagues (e.g. Mischel *et al.* 1989) demonstrated that by age four children can successfully delay gratification in order to receive a greater reward later. Choice between an immediate small and a delayed but greater reward clearly depends on the values of the rewards and the expected waiting time. However, other factors have been identified which are more relevant to the discussion of the impact of metamind. Exposure to the reward decreased average delay time for the children, and so did thinking about the reward versus thinking about something else. Disengaging from the primary representation of the reward and from one's desire for it increased self-control. Focusing on abstract qualities of the reward or exposure to the image of rather than the actual rewards also increased delay. In fact, merely imagining that the actually exposed reward was an image increased delay time. These findings support the idea that mental disengagement is crucial for executive control of the metamind and that abstractions (e.g., through symbolic representation) improve this ability.

Only the great apes have proven that they understand what they look like in the mirror self-recognition test. But whether they have evolved mental time-travel and a personal identity through time is highly questionable. Although apes have good memory (e.g. Fouts and Fouts 1993) and planning skills (e.g. Boesch and Boesch 1984), analysis of the existing evidence suggests that they do not have an autobiographic memory and are not capable of anticipating future drives (Suddendorf 1994; Suddendorf and Corballis 1997). Chimpanzees have shown considerable skill in delaying gratification (e.g. de Waal 1982) and Boysen and Berntson (1995) showed how symbols can help them to gain executive control to override impulsive responses: Chimpanzees who had great difficulty in choosing a smaller pile of sweets in order to obtain a larger one found the task much easier when the piles were replaced with numerical symbols. Language and metamind evolution seem to be linked on several planes. Nevertheless, the fact remains that great apes did not invent symbols by which to override impulses or to communicate with others, even if they can use symbols which we invent and teach to them.

---

[3] Povinelli *et al.* (1996) have studied a variant of the mirror self-recognition test by showing children a three-minute-old video recording of themselves being unknowingly marked with a sticker on their foreheads. On being showed the video, most four-year-olds removed the mark, while only a minority of three-year-olds did, leading Povinelli (1995) to conclude that, although a present sense of self might exist by age two, a time-travelling sense of self does not emerge until about age four. However my own research, while confirming this result, showed that a similar paradigm involving the introduction of an unexpected object in the room, rather than an unexpected sticker in the child's hair, yielded the same pattern of results. This study raises doubts as to whether the video test has any bearing on the concept of self (Suddendorf, unpublished manuscript, 1997).

Metamind and symbols: the connection between self and others

Zaitchik (1990) showed that three-year-olds have problems understanding not only mental misrepresentations but also physical misrepresentations. In an analogue of the classical false-belief task, Zaitchik asked children about a photograph of a previous situation that had now been altered. Three-year-olds expected the photo to represent the current situation while four-year-olds understood that the photo was of the past situation. A metamind explanation of this finding may be that three-year-olds cannot yet disengage from the current situation and fail to understand the photo as representing the situation as it was before the change. On the other hand, these tasks might not require metarepresentation because the pictures are true representations of the past, rather than false representations of the present, and indeed performance is not correlated with false-belief tasks (Perner 1995). However, Thomas et al. (1994) showed that young children have severe problems with the representational nature of pictures. Appearance-reality distinctions for pictures were found to be as difficult as classic appearance-reality tasks. Young pre-schoolers have difficulty simultaneously representing the distinct properties of picture and depicted (Robinson et al. 1994).

Similarly, `false maps' and `false drawings' are only understood by around age four (Charman and Baron-Cohen 1992; Leslie and Thaiss 1992). Understanding these `false' symbolic representations might imply metamind, because the symbol needs to be brought into relation with an internal memory representation of the previous situation, and might require disengagement from the current perception and metarepresentational tagging as `past' or as `false'. Symbolic representation, like mental representation, is understood only once the individual acquires metamind. Parkin (1994) found a robust correlation between false-belief task performance (`where does P think X is?') and a parallel false-symbol task involving a misleading direction sign (`where does this sign show X is?'). There is some evidence suggesting that physical misrepresentation is understood earlier than mental misrepresentation (e.g. Robinson et al. 1996), which is not really surprising when one considers that the former has a primary reality that can be examined (e.g. a photo), while the latter has no directly accessible medium.

The most significant symbols are of course those involved in language. It is not surprising, therefore, that strong correlations between metamind tasks and language abilities have been reported. Fletcher-Flinn and Snelson (1997), for example, found an association between metalinguistic skills (that is, syllable segmentation and rhyming tasks) and false-belief task performance. Doherty (unpublished, cited in Perner 1995) found a similar association between synonym monitoring and false-belief tasks. In another study, Suddendorf and Fletcher-Flinn (1996) found a strong correlation between verbal intelligence as measured by the British Picture Vocabulary Scale (BPVS; Dunn et al. 1982) and false-belief understanding. One might argue that a particular level of linguistic skill is necessary to comprehend the false-belief task and that the correlation merely reflects this. However, Jenkins and Astington (1996) found an association with performance on the Test of Early Language Development (TELD, Hresko et al. 1981) that appeared to go beyond this suggestion. While a certain level of language skill is required to pass false-belief tasks (98% of the children who passed at least one task scored 14 or higher, while only 33% of children who failed on all tasks reached this level), the authors found a significant correlation between the number of different tasks passed and the TELD scores even when analysed only for those who were above the threshold of 14. This suggests a more fundamental relationship than passing a threshold. Some have argued that the ability to comprehend the recursive nature of syntax is also involved in false-belief tasks (e.g. Feldman 1988), others point to semantic abilities involved in understanding metarepresentational terms such as remember, forget or surprise (Olson 1988; Lyon and Flavell 1994). Furthermore, pragmatic aspects of communication might be linked to false-belief task performance (Baron-Cohen 1988). Whatever the exact connection to language, the two abilities are apparently linked. Bickerton (1995) has even argued that they are fundamentally the same.

Plaut and Karmiloff-Smith (1993) have suggested a reason for this association that ties in closely with my interpretation of metamind. They appeal to the symbolic representation entailed by language as a mediating factor for false-belief task performance. They argue that symbolic representations need to be generated by the child that can effectively override the immediate reality bias of the present experience. In other words, disengagement from the immediate present is fostered by symbolic representation. This is supported by the above-mentioned results of Boysen and Berntson's (1995) study on chimpanzees' increased executive control when confronted with numerals rather than treats.

Gestural communication might have preceded vocal language in evolution (Hewes 1973; Corballis 1992; Goldin-Meadow and McNeill, this volume) and this capacity might have already depended on metamind. In a recent study we investigated whether there is an association between metamind (as measured by false-belief understanding) and gestural representation in the absence of real objects (Suddendorf et al. in press). Earlier research has identified two successive levels in the development of gestural representation: body-part-as-object (BPO) and imaginary object (IO) pantomimes (Overton and Jackson 1973; Boyatzis and Watson 1993; O'Reilly 1995). We found that when asked to pretend to perform a common action such as brushing teeth with a toothbrush, most children without metamind substitute a body part (finger) for the object. Children with metamind, however, significantly more often act as if there were an object (toothbrush) to interact with. Modelling such `imaginary object' pantomimes to pre-metamind children did not improve their performance; they continued to use body parts as objects. Metamind appears to be important for imaginary object pantomime and thus gestural communication.

But correlations do not tell us much about causal connections. We do not know whether language influenced the evolution of metamind or vice versa. Moreover, language is a prime example of how it can be misleading to reason from ontogeny to phylogeny. Children grow up in a verbal environment. Our forebears, on the other hand, must have invented language. To invent a symbolic representational system one has to have metarepresentation, because one has to understand symbols as representations. It is not surprising, then, that apes like the bonobo Kanzi, who grow up in a language environment, can learn language up to a level comparable to a two-year-old child (Savage-Rumbaugh *et* al. 1993), while chimpanzees in the wild have not invented such a linguistic system at all. Although their collating mind enables them to use symbols and interpret pictures or videos (e.g., Premack and Woodruff 1978), apes might not understand them as representations, and can therefore not invent a symbolic language or draw pictures that represent something. Premack and Premack (1983) went to elaborate lengths to teach their chimpanzees to utilize maps or scale models. All these efforts turned out to be fruitless, as not even identical rooms led the apes to look at the same baited place in the second room after discovering the treat in the first. The very idea that one thing represents the situation of another was beyond their grasp. However, there are at least two studies in which primates were quite successful at using video information about their enclosure (Menzel et al. 1978; Vauclair 1996), but it seems plausible that these animals mistook the video for reality (like a window), rather than understood the representational nature of the video.

With metamind the child becomes capable of understanding that something is represented by self, others, and symbols. The world becomes an entirely different place to the child. The way the child thinks is revolutionized.

Metamind and thought

The tasks presented earlier already exemplify the new mental experiences of the child. There are dramatic new developments such as the conceptualization of one's own and others' minds, personal history and future, self concept and control. There are further ways in which the metamind changes the way the child interprets and understands the world. With metamind matures what Dennett (1987) calls the intentional stance, which is a way of interpreting the events in the world by invoking mental states such as intentions and beliefs, rather than physical or design explanations. Heider and Simmel (1944) showed in now classic experiments how humans tend to interpret the world in mental terms. They showed subjects a silent film in which geometric figures moved about and the subjects were asked to describe what they saw. Descriptions generally were based on ascribing intentions and agency to the shapes. Adopting the intentional stance is a helpful heuristic in describing and understanding the world. Explanations often appear to have a wider appeal when proposed as narratives and when they anthropomorphize.

Cultural mythologies are ample evidence for this inclination. This might well be due to the Machiavellian origin of human intelligence.

When this view is coupled with other aspects of metamind such as mental timetravel, it seems inevitable that teleology (cf. Kelemen, this volume) emerges. This is the idea that the things around us are here for some purpose. Adopting this view must have been of momentous adaptive value for our ancestors, as it encouraged the search for new uses of objects in the environment and thus must have fostered human control over events. Human cultures have experimented with many religious and natural technologies to influence significant events, sometimes successfully (e.g. making fire) and sometimes not (e.g. making rain). Predicting the future, one of the main features of metamind, is of course most reliable when one is shaping the future. The natural curiosity of our species had therefore been extended into a new era. In childhood, too, this teleological stance becomes increasingly (sometimes irritatingly) obvious when the child asks more and more questions about *why* or *what for*. Every detail of the world is investigated as to its function and purpose. Understanding the world on the explicit metalevel becomes one of our main preoccupations. You and I are devoting significant parts of our lives to this quest for knowledge-why else would I write and you read this chapter? And if you attempt to answer this question, you could be said-thanks to the virtue of recursive thought-to be on yet another teleological quest.

Complementary to generating questions through knowing `that' and `what' we *do not* know is generating answers through knowing `that' and `what' we *do* know. This metaknowledge can be very beneficial in the way knowledge is utilized. Metarepresentation and disengaging from present perception may be crucial for creative problem solving. Actively searching the knowledge base beyond currently activated areas of mental content appears to be at the very heart of divergent thinking. In our own research we have found that three- to four-year-old children who passed false-belief tasks were able to generate more items on creativity tests requiring them to generate ideas or items according to given criteria (Suddendorf and Fletcher-Flinn 1997). In a recent longitudinal study we then showed that children improve in divergent thinking once they pass theory of mind tasks (Suddendorf and Fletcher-Flinn, in press).

Divergent thinking, like language, requires informational access to varied domains of knowledge in the generative process of combining and recombining items into virtually infinite numbers of novel sequences. This generativity has been recognized as unique to the human condition (Corballis 1991). Metamind might have been crucial to the emergence of this important feat. Even apparently intelligent behaviour in animals is often characterized by a lack of transfer or flexibility. Non-human primates that act `intelligently' in one domain fail to transfer their knowledge to another-they have a 'laser-beam intelligence' (Cheney and Seyfarth 1990). With metamind this domain specificity may be overcome. "Flexible transfer of knowledge between different domains is one of the hallmarks of humans' relentless creativity and invention" (Suddendorf and Fletcher-Flinn 1997, p. 67).

Is metamind domain general or modular? Evidence from autism

While some authors agree with the proposal of an across-the-board cognitive change (e.g. Perner 1991; Gopnik 1993), others have suggested the existence of domain-specific modules. In particular a theory-of-mind module, specific to understanding *mental* representation, has been proposed (Leslie and Thaiss 1992; Baron-Cohen 1995, this volume). The argument for a specific mind-reading module derived from the finding that autistic children fail theory-of-mind tasks while passing other tasks that are seemingly identical except that they require physical rather than mental representations (e.g., false photographs, drawings, and models-Leekam and Perner 1991; Charman and Baron-Cohen 1992, 1995).

But are autistic children deficient only in understanding mental representations? Perhaps the physical representations in these experiments were not really equivalent to mental representations. Physical representations are not only propositions about referents but also external real things themselves. So while they might misrepresent the current situation, they are themselves true existing objects in their own right. Perhaps it is easier to learn about their properties without invoking metarepresentation because they, in contrast to mental states, have a physical existence. In other words autistic children, even though lacking metamind, might learn to pass these non-mental tasks indirectly without metarepresentation. For example, the autistic children might have learned that photographs are stable and unchanging in nature, while the three-year-olds generalize, apply their immature representational understanding, and believe that what is happening in reality is also happening in the photograph (Leekam and Perner 1991; Robinson et al. 1994). Most studies match control and autistic sample according to mental age. This disguises the fact that the autistic sample is much older and thus has had much more opportunity to learn contingencies between observables. High functioning autistic people appear to solve many tasks through indirect computation of what `comes naturally' to clinically normal people. For an autistic person, emotional states may not simply be perceived in somebody else's face, but need to be deduced on the basis of observed covariations between facial cues and behaviour. In other words, autistic children might learn via association and reward where normal children are applying pre-programmed shortcuts. Indeed, autistic children appear to tackle emotional and social problems in much the same dry way as they address mathematical problems (Sigman et al. 1995). In some respects autistic people seem to be behaviourists in a mentalists' society (cf. Baron-Cohen 1989).

While autistic children show evidence for a collating mind, through, for example, mirror self-recognition (Dawson and McKissick 1984) and pretend play (V. Lewis and Boucher 1988), there is a lack of evidence for metamind. Besides consistently failing classic theory-of-mind tasks (Baron-Cohen *et al.* 1985; Leslie and Frith 1988; BaronCohen 1989; Perner *et al.* 1989; Prior *et* al. 1990), autistic children do not seem to develop autobiographic memory (Boucher and Warrington 1976; Boucher and Lewis 1989; Powell and Jordan 1993), have severe problems with mental disengagement and executive control (Ozonoff *et* al. 1991; Hughes and Russell 1993; Hughes *et* al. 1994), lack generativity in pretend play (Jarrold *et al.* 1996), have problems imagining unreal things (Scott and Baron-Cohen 1996), have a gross language-acquisition deficit (Sigman and Ungerer 1984), are characteristically engaging in stereotyped and routinized actions rather than flexible planned behaviour (e.g. Prior and Hoffman 1990; Harris 1991), and are basically acultural (Loveland 1991). Autistic children seem to lack executive control and skills in regards to knowledge about self, about others, symbols and thought. In other words they appear to lack a metamind.

Even if mental attribution can be dissociated from other metamind skills in the autistic condition, metamind may still have evolved as a single capacity applied to varied modules. The theory-of-mind aspect might be a subsystem that can be disturbed while other parts remain intact. One reason to assume that metamind is a domain-general skill (or module) is the lack of evidence that any part of metamind exists in other species. The case would be further strengthened if it could be shown that the different components of metamind co-emerged in evolution.

Some words of caution

Although I have referred to `pre-metamind' and used expressions like "once the child acquires metamind", this development is of course not as abrupt as these words suggest. Children who pass some metamind tasks continue to fail others (Jenkins and Astington 1996). Further, the test-retest reliability of false-belief tasks has been judged to be low to moderate (Mayes *et* al. 1996)-although it depends of course on the age range considered, because two-year-olds will always fail and five, six, seven, or eight-year-olds will reliably pass. Thus, there appears to be no clear transition point; rather, we should assume that there is a transition period. Indeed, abilities entailed by metamind continue to develop throughout childhood, and indeed we appear never to stop learning about people's minds. It is all the more surprising, then, that we do find robust associations between false-belief-task performance and various other skills such as imaginary object pantomime and creativity. The gradual change seems to produce a qualitative shift during this transition period.

The stark difference between monkeys and great apes might also be softened in future by more intensive studies of the lesser apes. Non-primate species (such as parrots and sea mammals) might also turn out to show sophisticated skills, reminding us that evolution is more like a tree branching in all directions than a ladder pointing to heaven.

While the sum of correlational evidence paints a picture of a domain-general metamind faculty emerging around age four, this is not certain. Even partialling out age and 10 does not entirely eliminate the worrisome fact that most things develop at this age period-and that correlations between skills are to be expected. Indeed, most of the experimental data reviewed, especially the comparative research, is mined with qualifications and reservations. Overall, however, the trends in the data can be assembled into an account that makes evolutionary sense, even if the model I am advancing is in some respects merely a `just so' story. Ultimately, the findings from developmental and comparative research have to fit into a coherent evolutionary theory. The model I present is a first attempt.

When did metamind come on to the scene?

Since no other animal has developed metamind, we might reasonably assume that it evolved some time after the phylogenetic split from the line that led to modern chimpanzees some five million years ago. The prime candidates for introducing this new cognitive machinery are Homo erectus and ergaster. We do not have evidence suggesting any major changes in mental capacity among the australopithecines (5 to 2 million years ago), as there are no artefacts and their relative brain size did not exceed that of the apes (Wood 1992). At the other end of the time scale, the oldest representational art is about 35 000 years old and burials are up to 100 000 years old. The recent discovery of sophisticated 400 000 year-old spears in Germany (Thieme 1997) prove much earlier extensive planning skills (suggesting mental time-travel). The emergence of metamind thus appears to lie somewhere between two million and 400000 years ago.

Around two million years ago (perhaps even two and a half million, according to Semaw et al. 1997) we have the beginning of the Homo clade with slightly increased brain size and the first stone tools (Corballis, this volume). These early tools, however, seem to be within the capacity of chimpanzees (Wynn and McGrew 1989; Toth et al. 1993) and there are no other hints of metamind. This changes with the emergence of *H.* ergaster/erectus some two million years ago.

Donald (1991, 1993, this volume) proposed the perhaps most plausible psychological account of the emergence of *H.* erectus/ergaster. At the centre of what he calls mimetic culture stands a new level of motor skill, which allows the whole body to be used as a communication and representation device. Crucial to such bodily or gestural communication would have been an ability to represent imaginary objects. Imaginary object pantomime, as argued above, is associated with metamind. Another aspect emphasized by Donald is the capacity for autocueing. The individual became able to consciously stop, replay, and edit bodily movement and thereby gain voluntary access to memory. Intentional rehearsal and refinement enabled *H.* ergaster/erectus to begin developing intentional symbolic communication and culture. Autocueing is crucial to free recall and episodic memory which are also implicated in metamind (Perner and Ruffman 1995).

But what is the `hard' evidence for metamind in *H.* ergaster/erectus? The hardest evidence, literally, is that deriving from stone tools. While the Oldowan tools associated with *H.* habilis can be produced quickly, and were possibly used `on the spot', *H.* ergaster's Acheulian tool kit implies some premeditation. The production of a symmetrical bifacial hand-axe requires time and precision. A tool like that would not be manufactured for one-time, on-the-spot use and it has thus been taken as evidence for mental time-travel (Suddendorf 1994; Suddendorf and Corballis 1997). Not only must the future need be anticipated, but the persistence and uniformity of this tool culture implies instructed cultural learning based on metamind.

Another line of evidence is the relatively sudden first wave of migration of *H. erectus* from Africa across the old world. This suggests that *H. erectus* were capable of altering the various environments to their distinct needs, although they had not developed the Acheulian tool kit of their African relatives *H. ergaster*. In practice this might have meant quick adaptation to different food resources, clothing for colder climates, and eventually fire for heat, light, cooking, and defence (the earliest evidence of the use of fire is about 500 000 years old). These things could not have been achieved without a metamind and without the evolving mimetic culture with generative thought and gestural communication. The origins of human cultural universals such as morality, justice, religion, language, narrative, mythology etc., might then go as far back as the emergence of the mental machinery I have called metamind.

The arrival of *H. sapiens* on the scene some 150 000 years ago probably brought with it a vocal apparatus capable of transforming the gestural communication systems of their forebears into speech. But it took yet another representational revolution to catapult humanity into a position from which the extraordinary development to world domination could begin. Only five thousand years ago did we begin to utilize our understanding of mental and vocal representational systems to create lasting symbolic representations. We invented writing. This external memory storage allowed cultures to share their knowledge across space and time (cf. Donald 1991). The consequential rapid accumulation of information was embraced and utilized for the erection of magnificent cultures such as that of the Egyptians where science, art, agriculture, astronomy, religion, and architecture flourished to new heights. Our current technological leap is a continuation of this process of extended use of external memory storage into our computer age.

With the help of external memory storage we can still today reflect on the reflections of a long-dead Frenchman. Descartes' cogito implies a self-awareness firmly seated in the present. With an evolutionary approach that includes the human species, indeed the human mind, our self-awareness can go far beyond the immediately present. We can realize that we are only a link in our species' evolution with a long past and, hopefully, a long future.

Acknowledgement

I would like to thank Michael C. Corballis, Russell M. Genet, Kennedy McLachlan, Josef Perner, and Kevin Waugh for their valuable comments on earlier versions of this paper. They are not responsible for my views, however.

References

Amsterdam, B. K. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, 5, 297-305.

Anderson, J. R. (1986). Mirror-mediated finding of hidden food by monkeys *(Macaca tonkeana* and M. *fascicularis). Journal of Comparative Psychology, 100*, 237-42.

Asendorpf, J. B. and Baudonniere, P. M. (1993). Self-awareness and other-awareness: mirror self-recognition and synchronic imitation among unfamiliar peers. *Developmental Psychology,* 29,88-95.

Asendorpf, J. B., Warkentin, V., and Baudonniere, P. M. (1996). Self-awareness and other-awareness II: mirror self-recognition, social contingency awareness, and synchronic imitation. *Developmental Psychology,* 32, 313-21.

Astington, J. W. (1994). *The child's discovery of the mind.* Fontana Press, London.

Astington, J. W. and Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition and Emotion, 9*, 151-65.

Baillargeon, R. (1987). Object permanence in 3.5- and 4.5-month old infants. *Developmental Psychology,* 23, 655-64.

Baillargeon, R. and DeVos, J. (1991). Object permanence in young infants: further evidence. *Child Development*, 62, 1227-46.

Baldwin, D. A. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20, 395-418.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders, 18,* 379-401.

Baron-Cohen, S. (1989). Are autistic children behaviorists? An examination of their mental-physical and appearance-reality distinctions. *Journal of Autism and Developmental Disorders, 19,* 579-600.

Baron-Cohen, S. (1995). *Mindblindness.* MIT Press, Cambridge, MA.

Baron-Cohen, S., Leslie, A., and Frith, U. (1985). Does the autistic child have a `theory of mind'? *Cognition*, 21, 37-46.

Barresi, J. and Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences, 19,* 107-54.

Bauer, P. J. (1996). What do infants recall of their lives? Memory for specific events by one-to two-year-olds. *American Psychologist*, 51, 29-41.

Bickerton, D. (1995). *Language and human behavior*. University of Washington Press, Seattle. Bischof, N. (1985). *Das Rdtzel Odipus* [The Oedipus riddle]. Piper, Munich.

Bischof-Kohler, D. (1989). *Spiegelbild und Empathie* [Mirror image and empathy]. Hans Huber, Bern.

Bischof-Kohler, D. (in press). Zusammenhange zwischen Kognitiver, motivationaler und emotionaler Entwicklung in der fruhen Kindheit und im Vorschulalter [Relationships between cognitive, motivational and emotional development in early childhood and in preschoolers]. *In Lehrbuch Ebtwicklungs psychologie* (ed. H. Keller). Huber, Bern.

Boesch, C. (1992). New elements of a theory of mind in wild chimpanzees. *Behavioral and Brain Sciences*, 15, 149-50.

Boesch, C. and Boesch, H. (1984). Mental map in wild chimpanzees: an analysis of hammer transports for nut cracking. *Primates*, 25, 160-170.

Boucher, J. and Lewis, V. (1989). Memory impairments and communications in relatively able autistic children. *Journal of Child Psychology and Psychiatry*, 30, 99-122.

Boucher, J. and Warrington, E. K. (1976). Memory deficits in early infantile autism: some similarities to the amnesic syndrome. *British Journal of Psychology*, 67, 73-87.

Boyatzis, C. J. and Watson, M. W. (1993). Preschool children's symbolic representation of objects through gestures. *Child Development*, 64, 729-35.

Boysen, S. T. and Berntson, G. G. (1995). Responses to quantity: perceptual versus cognitive mechanisms in chimpanzees *(Pan troglodytes). Journal of Experimental Psychology: Animal Behavior Processes*, 21, 82-6.

Brown, D. E. (1991). *Human universals*. McGraw-Hill, New York.

Brownell, C. A. and Carriger, M. S. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Development*, 61, 1164-74.

Bullock, M. and Luetkenhaus, P. (1988). The development of volitional behavior in the toddler years. *Child Development*, *59*, 664-74.

Byrne, R. W. (1994). The evolution of intelligence. In *Behavior and evolution* (ed. P. J. B. Slater and T. R. Halliday), pp. 223-65. Cambridge University Press.

Byrne, R. W. and Whiten, A. (ed.) (1988). *Machiavellian intelligence*. Clarendon Press, Oxford. Byrne, R. W. and Whiten, A. (1990). Tactical deception in primates: the 1990 database. *Primate Report*, *27, 1-101.*

Byrne, R. W. and Whiten, A. (1992). Cognitive evolution in primates. *Man*, *27,* 609-27. Carruthers, P. and Smith, P. K. (1996). *Theories of theories of mind.* Cambridge University Press. Chandler, M. and Hala, S. (1994). The role of personal involvement in the assessment of early false belief skills. In *Children's early understanding of mind* (ed. C. Lewis and P. Mitchell), pp. 403-25. Erlbaum, Hove.

Charman, T. and Baron-Cohen, S. (1992). Understanding drawings and beliefs: a further test of the metarepresentation theory of autism. *Journal of Child Psychology and Psychiatry, 33,* 1105-12.

Charman, T. and Baron-Cohen, S. (1995). Understanding photos, models, and beliefs: a test of the modularity thesis of theory of mind. *Cognitive Development*, *10,* 287-98.

Cheney, D. L. and Seyfarth, R. M. (1990). How *monkeys see the world.* University of Chicago Press.

Clements, W. A. and Perner, J. (1997). When actions really speak louder than words-but only implicitly: young children's understanding of false belief in action. Unpublished manuscript. Concar, D. (1996). Into the mind unborn. *New Scientist*, *152,* 40-5.

Corballis, M. C. (1991). *The lopsided ape.* Oxford University Press.

Corballis, M. C. (1992). On the evolution of language and generativity. *Cognition*, *44,* 197-226. Daechler, M., Bukatko, D., Benson, K., and Myers, N. (1976). The effects of size and color cues on the delayed response of very young children. *Bulletin of the Psychonomic Society, 7,* 65-8. Dawson, G. and McKissick, F. C. (1984). Self-recognition in autistic children. *Journal of Autism and Developmental Disorders*, *14,* 383-94.

deBlois, S. T. and Novak, M. A. (1994). Object permanence in rhesus monkeys (Macaca mulatta). *Journal of Comparative Psychology*, *108,* 318-27.

Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences, 4,* 568-70. Dennett, D. C. (1987). *The intentional stance.* MIT Press, Cambridge, MA.

Dennett, D. C. (1995). *Darwin's dangerous idea.* Simon & Schuster, New York. de Waal, F. B. M. (1982). *Chimpanzee politics.* Jonathan Cape, London.

de Waal, F. B. M. (1989). *Peacemaking among primates.* Harvard University Press.

DeCasper, A. J. and Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9,* 133-50.

Diamond, A. (1985). Development of the ability to use recall to guide action, as indicated by infants' performance on AB. *Child Development, 56,* 868-83.

Dohl, F. (1970). Zielorientiertes Verhalten beim Schimpansen [Goal-directed behavior in chimpanzees]. *Naturwissenschaft und Medizin, 34,* 43-57.

Donald, M. (1991). *Origins of the modern mind.* Harvard University Press.

Donald, M. (1993). Human cognitive evolution: what we were, what we are becoming. *Social Research, 60,* 143-70.

Dretske, F. (1995). *Naturalizing the mind.* MIT Press, London.

Dunn, J. (1991). Young children's understanding of other people: evidence from observations within the family. In *Children's theories of mind* (ed. D. Frye and C. Moore), pp. 97-114. Lawrence Erlbaum, Hillsdale, NJ.

Dunn, L. M., Dunn, L., Whetton, C., and Pintillie, D. (1982). *British picture vocabulary scale.* Nfer-Nelson, London.

Ettlinger, G. and Wilson, W. A. (1990). Cross-modal performance: behavioral processes, phylogenetic considerations and neuronal mechanisms. *Behavioral Brain Research, 40,* 169-92.

Feldman, C. F. (1988). Early forms of thought about thoughts: some simple linguistic correlates of mental state. In *Developing theories of mind* (ed. J. W. Astington, P. L. Harris, and D. R. Olson), pp. 126-137. Cambridge University Press.

Filion, C. M., Washburn, D. A.; and Gulledge, J. P. (1996). Can monkeys (Macaca mulatta) represent invisible displacement? *Journal of Comparative Psychology, 110,* 386-95.

Flavell, J. H. (1993). The development of children's understanding of false belief and the appearance-reality distinction. *International Journal of Psychology, 28,* 595-604.

Flavell, J. H., Green, F. L., and Flavell, E. R. (1986). Development of knowledge about the appearance-reality distinction. *Monographs of the Society for Research in Child Development,* 51, (1, Serial No. 212).

Flavell, J. H., Green, F. L., Wahl, K. E., and Flavell, E. R. (1987). The effects of question clarification and memory aids on young children's performance-reality tasks. *Cognitive Development, 2,* 127-44.

Fletcher-Flinn, C. M. and Snelson, H. (1997). The relation between metaphonological abilities, metarepresentation and cognition. *New Zealand Journal of Psychology, 26,* 20-8.

Fouts, R. S. and Fouts, D. H. (1993). Chimpanzees' use of sign language. In *The great ape project,* (ed. P. Cavalieri and P. Singer), pp. 28-41. Fourth Estate, London.

Frankfurt, H. G. (1988). *The importance of what we care about.* Cambridge University Press. Freeman, N. H., Lacohee, H., and Coulton, S. (1995). Cued-recall approach to 3-year-olds' memory for an honest mistake. *Journal of Experimental Child Psychology, 60,* 102-15.

Frye, D., Zelazo, P. D., and Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10,* 483-527.

Gagliardi, J. L., Kirkpatrick-Steger, K. K., Thomas, J., Allen, G. J., and Blumberg, M. S. (1995). Seeing and knowing: knowledge attribution versus stimulus control in adult humans *(homo sapiens). Joumal of Comparative Psychology, 109,* 107-14.

Gallup, G. G., Jr. (1970). Chimpanzees: self recognition. *Science, 167,* 86-7.

Gardner R. and Gardner, B. (1969). Teaching sign language to a chimpanzee. *Science, 165,* 664-72.

Gibson, E. J. (1995). Are we automata? In *The self in early infancy* (ed. P. Rochat), pp. 3-15. Elsevier, Amsterdam.

Goodall, J. (1986). *The chimpanzees of Gombe: patterns of behavior.* Harvard University Press. Gopnik, A. (1993). How we know our own minds: the illusion of first-person knowledge of intentionality. *Behavioral and Brain Sciences, 16,* 1-14.

Gopnik, A. and Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development, 59,* 26-37.

Gopnik, A. and Graf, P. (1988). Knowing how you know: young children's ability to identify and remember the source of their belief. *Child Development, 59,* 1366-71.

Greenfield, P. M. and Savage-Rumbaugh, E. S. (1990). Grammatical combination in Pan paniscus: processes of learning and invention in the evolution and development of language. In *`Language' and intelligence in monkeys and apes* (ed. S. T. Parker and K. R. Gibson), pp. 540-78. Cambridge University Press.

Haake, R. J. and Somerville, S. C. (1985). Development of logical search skills in infancy. *Developmental Pschology, 21,* 176-86.

Harris, P. L. (1991). The work of imagination. In *Natural theories of mind: evolution, development, and simulation of everyday mind-reading* (ed. A. Whiten), pp. 283-304. Basil Blackwell, Oxford.

Hayes, C. (1951). *The ape in our house.* Harper and Brothers, New York.

Heider, F. and Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57,* 243-59.

Herman, L. M., Pack, A. A., and Morel-Samuels, P. (1993). Representational and conceptual skills of dolphins. In *Language and communication: comparative perspectives* (ed. H. L. Roitblat, L. M. Herman, and P. E. Nachtigall), pp. 403-42. Lawrence Erlbaum, Hillsdale, NJ. Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology, 14,* 5-24.

Heyes, C. M. (1993). Anecdotes, training, trapping, and triangulation: do animals attribute mental states? *Animal Behaviour, 46,* 177-88.

Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences, 21,* 101-48.

Howe, M. L. and Courage, M. L. (1993). On resolving the enigma of infantile amnesia. *Psychological Bulletin, 113,* 305-26.

Howe, M. L. and Courage, M. L. (1997). The emergence and early development of autobiographical memory. *Psychological Review, 104,* 499-523.

Hresko, W. P., Reid, D. K., and Hammill, D. D. (1981). *The Test of Early Language Development (TELD).* Pro-Ed, Austin.

Hughes, C. and Russell, J. (1993). Autistic children's difficulty with mental disengagement from an object: its implications for theories of autism. *Developmental Psychology, 3,* 498-510. Hughes, C., Russell, J., and Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia, 32,* 477-92.

Humphrey, N. K. (1976). The social function of intellect. In *Growing points in ethology* (ed. P. P. G. Bateson and R. A. Hinde), pp. 303-17. Cambridge University Press.

Humphrey, N. K. (1986). *The inner eye.* Faber and Faber, London. Humphrey, N. K. (1992). *A history of the mind.* Chatto and Windus, London. Hyatt, C. W. and Hopkins, W. D. (1994). Self-awareness in bonobos and chimpanzees: a comparative perspective. In *Self-awareness in animals and humans* (ed. S. T. Parker, R. W. Mitchell, and M. L. Boccia), pp. 248-53. Cambridge University Press.

James, W. (1890). *Principles of psychology, Vol.* 1-2. Holt, New York.

Jarrold, C., Carruthers, P., Smith, P. K., and Boucher, J. (1994). Pretend play: is it metarepresentational? *Mind and Language, 9,* 445-68.

Jarrold, C., Boucher, J., and Smith, P. K. (1996). Generativity deficits in pretend play in autism. *British Journal of Developmental Psychology, 14,* 275-300.

Jenkins, J. M. and Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology, 32, 70-8.* Jolly, A. (1966). Lemur social behavior and primate intelligence. *Science, 153,* 501-6.

Jolly, A. (1985). *The evolution of primate behavior,* (2nd edn). MacMillan, New York. Kagan, J. (1981). *The second year*. Harvard University Press.

Kohler, W. (1927). *The mentality of apes* (trans. E. Winter). Routledge and Kegan Paul, London. (Original work published 1917).

Kummer, H., Anzensberger, G., and Hemelrijk, C. K. (1996). Hiding and perspective taking in long-tailed macaques *(Macaca fascicularis). Journal of Comparative Psychology, 110,* 97-102. Kundera, M. (1991). Immortality. Grone Weidenfeld, New York.

Lalonde, C. E. and Chandler, M J. (1995). False belief understanding goes to school: on the social-emotional consequences of coming early or late to a first theory of mind. *Cognition and Emotion, 9,* 167-85.

Leahy, T. H. (1994). Is this a dagger I see before me? Four theorists in search of consciousness. *Contemporary Psychology, 39,* 575-81.

Leekam, S. R. and Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40,* 203-18.

Leslie, A. (1987). Pretence and representation in infancy: the origin of `theory of mind'. *Psychological Review, 94,* 412-26.

Leslie, A. and Frith, U. (1988). Autistic children's understanding of seeing, knowing, and believing. *British Journal of Developmental Psychology, 6,* 315-29.

Leslie, A. and Thaiss, L. (1992). Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition,* 43, 225-51.

Lewis, M. and Brooks-Gunn, J. (1979). *Social cognition and the acquisition of self.* Plenum, New York.

Lewis, M., Sullivan, M. W., Stranger, C., and Weiss, M. (1989). Self development and self-conscious emotions. *Child Development*, 60, 146-56.

Lewis, V. and Boucher, J. (1988). Spontaneous, instructed and elicited play in relatively able autistic children. *British Journal of Developmental Psychology,* 6, 325-39.

Lillard, A. S. (1993). Pretend play skills and the child's theory of mind. *Child Development,* 64, 348-71.

Lillard, A. S. (1998). Ethnopsychologies: cultural variations in theory of mind. *Psychological Bulletin,* 123, 3-32.

Lorenz, K. (1973). *Die Rueckseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens* [Behind the mirror: A search for a natural history of human knowledge]. Piper, Munich.

Loveland, K. A. (1991). Social affordances and interaction: II autism and the affordances of the human environment. *Ecological Psychology,* 3, 99-119.

Lyon, D. L. and Flavell, J. H. (1994). Young children's understanding of `remember' and `forget'. *Child Development*, 65, 1357-71.

Marten, K. and Psarakos, S. (1994). Evidence of self-awareness in bottlenose dolphins *(Tursiops truncatus).* In *Self-awareness in animals and humans,* (ed. S. T. Parker, R. W. Mitchell, and M. L. Boccia), pp. 361-79. Cambridge University Press.

Marten, K. and Psarakos, S. (1995). Using self-view television to distinguish between self-examination and social behavior in the bottlenose dolphin *(Tursiops truncatus). Consciousness and Cognition*, 4, 205-24.

Mayes, L. C., Min, A., Tercyak, K. P., Cicchetti, D. V., and Cohn, D. J. (1996). Test-retest reliability for false-belief tasks. *Journal of Child Psychology and Psychiatry,* 37, 313-19. Meador, D. M., Rumbaugh, D. M., Pate, J. L., and Bard, K. A. (1987). Learning, problem solving, cognition, and intelligence. In *Comparative primate biology, Vol.* 2B (ed. J. Erwin), pp.17-83. Alan R. Liss, New York.

Meltzoff, A. N. (1995). Understanding the intentions of others: re-enactment of intended acts by 18-month-old children. *Developmental Psychology,* 31, 838-50.

Meltzoff, A. N. and Gopnik, A. (1989). On linking nonverbal imitation, representation, and language learning in the first two years of life. In *The many faces of imitation in language learning* (ed. G. E. Speidel and K. E. Nelson), pp. 23-51. Springer, Berlin.

Menzel, E. W. (1974). A group of young chimpanzees in a one-acre field. In *Behavior of nonhuman primates, Vol.* 5, (ed. A. M. Schrier and F. Stollnitz), pp. 83-153. Academic Press, New York.

Menzel, E. W., Premack, D., and Woodruff, G. (1978). Map reading by chimpanzees. *Folia Primatologica,* 29, 241-49.

Menzel, E. W., Savage-Rumbaugh, E. S., and Lawson, J. (1985). Chimpanzee *(Pan troglodytes)* problem solving with the use of mirrors and televised equivalents of mirrors. *Journal of Comparative Psychology,* 99, 211-17.

Miles, L. (1990). The cognitive foundations for reference in a signing orang-utan. In *`Language' and intelligence in monkeys and apes* (ed. S. T. Parker and K. R. Gibson), pp. 511-39. Cambridge University Press.

Mischel, W., Shoda, Y., and Rodriguez, M. L. (1989). Delay of gratification in children. *Science,* 244,933-8.

Mitchell, P. (1997). *Introduction to theory of mind.* Arnold, London.

Moore, C. (1996). Theories of mind in infancy. *British Journal of Developmental Psychology,* 14, 19-40.

Moore, C., Pure, K., and Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development, 61,* 722-30.

Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., and MacCallum, F. (1995). Conflicting desires and the child's theory of mind. *Cognitive Development, 10,* 467-82.

Moore, C., Barresi, J., and Thompson, C. (in press). The cognitive basis of future-oriented prosocial behavior. *Social Development.*

Natale, F. and Antinucci, F. (1989). Stage 6 object-concept and representation. In *Cognitive structure and development in nonhuman primates* (ed. F. Antinucci), pp. 97-112. Lawrence Erlbaum, Hillsdale, NJ.

Natale, F., Antinucci, F., Spinozzi, G., and Poti, P. (1986). Stage 6 object concept in non-human primate cognition: a comparison between gorilla *(Gorilla gorilla gorilla)* and Japanese macaque *(Macaca fuscata). Journal of Comparative Psychology, 100,* 335-9.

Nelson, K. (1992). Emergence of autobiographical memory at age 4. *Human Development, 35,* 172-7.

O'Connell, S. M. (1995). Empathy in chimpanzees: evidence for theory of mind? *Primates, 36,* 397-410.

O'Neill, D. K. and Gopnik, A. (1991). Young children's ability to identify the sources of their beliefs. *Developmental Psychology, 27,* 390-7.

O'Reilly, A. W. (1995). Using representations: Comprehension and production of actions with imagined objects. *Child Development, 66,* 999-1010.

Olson, D. R. (1988). On the origin of beliefs and other intentional states in children. In *Developing theories of mind* (ed. J. W. Astington, P. L. Harris, and D. R. Olson), pp. 414-26. Cambridge University Press.

Olson, D. R. (1993). The development of representations: the origins of mental life. *Canadian Psychology, 34,* 1-14.

Overton, W. F. and Jackson, J. P. (1973). The representation of imagined objects in action sequences: a developmental study. *Child Development, 44,* 309-14.

Ozonoff, S., Pennington, B. F., and Rogers, S. (1991). Executive function deficits in high-functioning autistic children: relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32,* 1081-105.

Parker, S. T. (1990). Origins of comparative developmental evolutionary studies of primate mental abilities. In *`Language'and intelligence in monkeys and apes* (ed. S. T. Parker and K. R. Gibson), pp. 3-64. Cambridge University Press.

Parker, S. T. and Gibson, K. R. (1979). A developmental model for the evolution of language and intelligence in early hominids. *Behavioral and Brain Sciences, 2,* 367-408.

Parker, S. T. and Gibson, K. R. (1990). *`Language' and intelligence in monkeys and apes.* Cambridge University Press.

Parker, S. T., Mitchell, R. W., and Boccia, M. L. (eds.). (1994). *Self-awareness in animals and humans.* Cambridge University Press.

Parkin, L. (1994). *Normal and autistic children's theory of representation.* Unpublished doctoral thesis, University of Sussex.

Patterson, F. (1991). Self-awareness in the gorilla Koko. *Gorilla, 14,* 2-5.

Patterson, F. and Linden, F. (1981). *The education of Koko.* Holt, Rinehart and Winston, New York.

Pepperberg, I. M. (1990). Conceptual abilities of nonprimate species, with an emphasis on an African grey parrot. In *`Language' and intelligence in monkeys and apes* (ed. S. T. Parker and K. R. Gibson), pp. 469-507. Cambridge University Press.

Pepperberg, I. M. (1993). Cognition and communication in an African grey parrot *(Psittacus erithacus):* studies on nonhuman, nonprimate, nonmammalian subject. In *Language and communication: comparative perspectives* (ed. H. L. Roitblat, L. M. Herman, and P. E. Nachtigall), pp. 221-48. Lawrence Erlbaum, Hillsdale, NJ.

Pepperberg, I. M. and Kozak, F. A. (1986). Object permanence in the African grey parrot *Psittacus erithacus). Animal Learning and Behavior, 14,* 322-30.

Pepperberg, I. M., Garcia, S. E., Jackson, E. C., and Marconi, S. (1995). Mirror use by African grey parrots. *Journal of Comparative Psychology,* 109, 182-95.

Perner, J. (1991). *Understanding the representational mind.* MIT Press, Cambridge, MA.

Perner, J. (1995). The many faces of belief: reflections on Fodor's and the child's theory of mind. *Cognition*, 57, 241-69.

Perner, J. and Ruffman, T. (1995). Episodic memory and autogenetic consciousness: developmental evidence and a theory of childhood amnesia. *Journal of Experimental Child Psychology,* 59,516-48.

Perner, J., Frith, U., Leslie, A., and Leekam, S. (1989). Exploration of the autistic child's theory of mind. *Child Development*, 60, 689-700.

Peskin, J. (1992). Ruse and representations: on children's ability to conceal information. *Developmental Psychology, 28,* 84-9.

Peskin, J. (1996). Guise and guile: children's understanding of narratives in which the purpose of pretense is deception. *Child Development*, 67, 1735-51.

Piaget, J. (1951). *Play, dreams and imitation in childhood.* Heinemann, Melbourne.

Piaget, J. (1952). *The origins of intelligence in children.* International University Press, New York. Piaget, J. (1954). *The construction of reality in the child.* Basic Books, New York.

Pillemer, D. B., Picariello, M. L., and Pruett, J. C. (1994). Very long-term memories of a salient preschool event. *Applied Cognitive Psychology, 8,* 95-106.

Plaut, D. C. and Karmiloff-Smith, A. (1993). Representational development and theory-of-mind computations. *Behavioral and Brain Sciences*, 16, 70-1.

Povinelli, D. J. (1989). Failure to find self-recognition in Asian elephants *(Elephans maximus)* in contrast to their use of mirror cues to discover hidden food. *Journal of Comparative Psychology, 103,* 122-31.

Povinelli, D. J. (1995). The unduplicated self. In *The self in early infancy* (ed. P. Rochat). North-Holland-Elsevier, Amsterdam.

Povinelli, D. J. and Cant, J. G. H. (1995). Arboreal clambering and the evolution of selfconception. *Quarterly Review of Biology,* 70, 393-421.

Povinelli, D. J. and Eddy, T. J. (1996). What young chimpanzees know about seeing. *Monographs of the Society for Research in Child Development,* 61, v-247.

Povinelli, D. J., Nelson, K. E., and Boysen, S. T. (1990). Inferences about guessing and knowing by chimpanzees *(Pan troglodytes). Journal of Comparative Psychology, 104,* 203-10. Povinelli, D. J., Nelson, K. E., and Boysen, S. T. (1992). Comprehension of social role reversal by chimpanzees: evidence for empathy? *Animal Behaviour*, *43,* 633-40.

Povinelli, D. J., Parks, K. A., and Novak, M. A. (1992). Role reversal by rhesus monkeys, but no evidence of empathy. *Animal Behaviour*, *44,* 269-81.

Povinelli, D. J., Landau, K. R., and Perilloux, H. K. (1996). Self-recognition in young children using delayed versus live feedback: evidence for a developmental asynchrony. *Child Development,* 67, 1540-54.

Powell, S. D. and Jordan, R. R. (1993). Being subjective about autistic thinking and learning to learn. *Educational Psychology, 13,* 359-70.

Pratt, C. and Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development*, 61, 973-82.

Premack, D. (1988). `Does the chimpanzee have a theory of mind?' revisited. In *Machiavellian intelligence* (ed. R. W. Byrne and A. Whiten), pp. 160-79. Clarendon Press, Oxford. Premack,

D. and Dasser, V. (1991). Perceptual origins and conceptual evidence for theory of mind in apes and children. In *Natural theories of mind: evolution, development and simulation of everyday mind-reading* (ed. A. Whiten), pp. 253-66. Basil Blackwell, Oxford.

Premack, D. and Premack, A. J. (1983). *The mind of an ape. W. W.* Norton, New York.

Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 4,* 515-26.

Prior, M. and Hoffman, W. (1990). Neuropsychological testing of autistic children through an exploration of frontal lobe tests. *Journal of Autism and developmental Disorders, 20,* 581-90. Prior, M., Dahlstrom, B., and Squires, T. L. (1990). Autistic children's knowledge of thinking and feeling states in other people. *Journal of Child Psychology and Psychiatry, 51,* 587-601. Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? *Behavioral and Brain Sciences, 1,* 592-3.

Robinson, E. J., Nye, R., and Thomas, G. V. (1994). Children's conceptions of the relationship between pictures and their referents. *Cognitive development, 9,* 165-91.

Robinson, E. J., Riggs, K. J., and Samuel, J. (1996). Children's memory for drawings based on a false belief. *Developmental Psychology, 32,* 1056-64.

Rovee-Collier, C. (1997). Dissociations in infant memory: rethinking the development of implicit and explicit memory. *Psychological Review, 104,* 467-98.

Ruffman, T., Olson, D. R., Ash, T., and Keenan, T. (1993). The ABCs of deception: do young children understand deception in the same way as adults? *Developmental Psychology, 29,* 74-87.

Rumbaugh, D. M., Savage-Rumbaugh, E. S., and Washburn, D. A. (1994). Learning, prediction, and control with an eye to the future. In *The development of future-oriented processes* (ed. M. M. Haith, J. B. Benson, R. J. Roberts, and B. Pennington), pp. 120-38. University of Chicago Press.

Russell, J., Jarrold, C., and Potel, D. (1994). Executive factors in preschoolers' strategic deception. *British Journal of Developmental Psychology, 12,* 301-14.

Saltmarsh, R., Mitchell, P., and Robinson, E. (1995). Realism and children's early grasp of mental representation: belief-based judgements in the state change task. *Cognition, 57,* 297-325.

Savage-Rumbaugh, E. S. and McDonald, K. (1988). Deception and social manipulation in symbol-using apes. In *Machiavellian Intelligence* (ed. R. W. Byrne and A. Whiten), pp. 224-37. Clarendon Press, Oxford.

Savage-Rumbaugh, E. S., Murphy, J., Sevick, R. A., Brakke, K .E., Williams, S. L., and Rumbaugh, D. M. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development, 58,* v-221.

Schusterman, R. J., Gisiner, R., Grimm, B. K., and Hanggi, E. B. (1993). Behavioral control by exclusion and attempts at establishing semanticity in marine mammals using match to sample paradigms. In *Language and communication: comparative perspectives* (ed. H. L. Roitblat, L. M. Herman, and P. E. Nachtigall), pp. 249-74. Lawrence Erlbaum, Hillsdale, NJ.

Scott, F. J., and Baron-Cohen, S. (1996). Imagining real and unreal things: evidence of a dissociation in autism. *Journal of Cognitive Neuroscience, 8,* 371-82.

Semaw, S., Renne, P., Harris, J. W. K., Feibel, C. S., Bernor, R. L., Fesseha, N., *et al.* (1997). 2.5-million-year-old stone tools from Gona, Ethiopia. *Nature, 385,* 333-336.

Siegal; M. (1995). Becoming mindful of food and conversation. *Current Directions in Psychological Science, 4,* 177-81.

Sigman, M. and Ungerer, J. A. (1984). Cognitive and language skills in autistic, mentally retarded, and normal children. *Developmental Psychology, 20,* 293-302.

Sigman, M., Arbelle, S., and Dissanayake, C. (1995). Current research findings on childhood autism. *Canadian Journal of Child Psychiatry, 40,* 289-98.

Suarez, S. D. and Gallup, G. G., Jr. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution, 10,* 175-88.

Suddendorf, T. (1994). *Discovery of the fourth dimension: mental time travel and human evolution.* Master's thesis, University of Waikato, Hamilton, New Zealand. Internet: http://cogprints.soton.ac.uk

Suddendorf, T. *(1997). Children's understanding of the relation between delayed video representation and current reality: a test for self-awareness?* Unpublished manuscript, University of Auckland, New Zealand. Internet: http://cogprints.soton.ac.uk

Suddendorf, T. and Corballis, M. C. *(1997).* Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs, 123, 133-67.*

Suddendorf, T. and Fletcher-Flinn, C. M. *(1996).* Some correlates of the emerging theory of mind. *Australian Journal of Psychology, 48, (suppl.), 26-7.*

Suddendorf, T. and Fletcher-Flinn, C. M. *(1997).* Theory of mind and the origins of divergent thinking. *Journal of Creative Behavior, 31, 59-69.*

Suddendorf, T. and Fletcher-Flinn, C. M. (in press). Children's divergent thinking improves when they understand false beliefs. *Creativity Research Journal,* special issue.

Suddendorf, T., Fletcher-Flinn, C. M., and Johnston, L. (in press). *Pantomime and theory of mind. Journal of Genetic Psychology.*

Taylor, T., Esbensen, B. M., and Bennett, R. T. *(1994).* Children's understanding of knowledge acquisition: the tendency for children to report that they have always known what they have just learned. *Child Development, 65, 1581-604.*

Thieme, H. *(1997).* Lower Palaeolithic hunting spears from Germany. *Nature, 385, 807-10.* Thomas, G. V., Nye, R., and Robinson, E. J. *(1994).* How children view pictures: children's responses to pictures as things in themselves and as representations of something else. *Cognitive Development, 9, 141-64.*

Tomasello, M., Kruger, A. C., and Ratner, H. H. *(1993).* Cultural learning. *Behavioral and Brain Sciences, 16, 495-552.*

Toth, N., Schick, K. D., Savage-Rumbaugh, S., Sevcik, R. A., and Rumbaugh, D. M. *(1993).* Pan the toolmaker: investigations into stone tool-making and tool-using capabilities of a bonobo *(Pan paniscus). Journal of Archeological Science, 20, 81-91.*

Tulving, E. *(1985).* How many memory systems are there? *American Psychologist, 40, 385-398.* Vauclair, J. *(1996). Animal cognition.* Harvard University Press, London.

Visalberghi, E. and Limongelli, L. *(1994).* Lack of comprehension of cause-effect relationships in tool-using capuchin monkeys *(Cebus apella). Journal of Comparative Psychology, 108, 15-22.*

Welch-Ross, M. K. *(1995).* An integrative model of the development of autobiographical memory. *Developmental Review, 15, 338-65.*

Wellman, H. M. *(1990). The child's theory of mind.* MIT Press, Cambridge.

Wheeler, M. A., Stuss, D. T., and Tulving, E. *(1997).* Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin, 121, 331-54.*

Whiten, A. *(1996).* Imitation, pretence and mindreading. In *Reaching into thought: the minds of the great apes* (ed. A. Russon, K. A. Bard and S. T. Parker), pp. *300-24.* Cambridge University Press.

Whiten, A. and Byrne, R. W. *(1988).* Tactical deception in primates. *Behavioral and Brain Sciences, 11, 233-73.*

Whiten, A. and Byrne, R. W. *(1991).* The emergence of metarepresentation in human ontogeny and primate phylogeny. In *Natural theories of mind: evolution, development, and simulation of everyday mind-reading* (ed. A. Whiten), pp. *267-81.* Blackwell, Oxford.

Whiten, A. and Ham, R. *(1992).* On the nature and evolution of imitation in the animal kingdom: reappraisal of a century of research. In *Advances in the study of behavior, Vol. 21,* (ed. P. J. B. Slater, J. S. Rosenblatt, C. Beer, and M. Milinski). Academic Press, New York. Wimmer, H. and Hard, M. *(1991).* Against the Cartesian view of mind: young children's difficulty with own false beliefs. *British Journal of Developmental Psychology, 9, 125-38.* Wimmer, H. and Perner, J. *(1983).* Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13, 103-28.*

Wimmer, H., Hogrefe, G. J., and Perner, J. (1988). Children's understanding of informational access as source of knowledge. *Child Development, 59,* 386-96.

Wood, B. (1992). Origin and the evolution of the genus *Homo. Nature, 355*, 783-90.
Woolley, J. D. and Bruell, M. J. (1996). Young children's awareness of the origins of their mental representations. *Developmental Psychology, 32,* 335-46,
Wynn, T. and McGrew, W. C. (1989). An ape's view of the Oldowan. *Man, 24,* 383-98. Zahn-Waxler, C., Radke-Yarrow, M., and King, R. A. (1979). Child-rearing and children's prosocial imitations toward victims of distress. Child *Development, 50,* 319-30.
Zaitchik, D. (1990). When representations conflict with reality: the preschooler's problem with false beliefs and `false' photographs. *Cognition, 35*, 41-68.
Zelazo, P. D., and Jacques, S. (1997). Children's rule use: representation, reflection and cognitive control. *Annals of Child Development, 12,* 119-76.
Zelazo, P. D., Frye, D., and Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, 11, 37-63.