

On Reading Others' Minds

Louise Röska-Hardy

In everyday human life we characteristically take each other to have thoughts, to make decisions and to pursue plans of action. We conceive of ourselves as thinkers, even if not exclusively so, and this has consequences for our understanding of thought, action and the social world. In philosophical terms, we take ourselves and other human beings to have minds, where the concept of a mind is the concept of that which is capable of consciously deliberating, believing, desiring, sensing, perceiving, purposefully moving and acting. This understanding informs our conception of ourselves and others as minded agents in the natural world, who can act in the light of their beliefs, desires and intentions. Through reference to beliefs, desires, emotions and other states of mind, we routinely construe others' behavior as action and we regard the states of mind we attribute to them as explanatory of their thoughts and actions. An understanding of thought and action in terms of states of mind pervades everyday social intercourse and provides the foundation for much of human culture. In brief, the capacity to understand oneself and others in terms of contentful beliefs, desires and other states of mind is crucial to typically human social cognition.

Human beings have evolved both biologically and culturally to be able to attribute states of mind to others. We as humans have developed the ability to think discursively, to be aware of what we are thinking, and to characterize thought and action with linguistic means. This ability is manifest in our everyday ways of describing and understanding our own and others' states of mind. Typically, we attribute states of mind by employing the resources of a natural language we know, e.g. by using a sentence to ascribe a mental state. Such attributions are essential to the way we justify, explain and anticipate thought and action in everyday situations. Any approach to mental attribution that claims to capture our *ordinary, everyday* attributions of states of mind must be able to account for the attributions we make using the sentences of a natural language. For it is these attributions that play a crucial role in our everyday, "folk-psychological" understanding of our own and others' behavior, on the one hand, and provide the candidates for our claims to know our own minds, on the other. Thus, if an account of our ability to understand others by reference to states of mind is to prove adequate, it must be able to accommodate the mental state attributions we make by means of language.

In recent years the capacity to interpret and anticipate the thoughts and behaviors of others in terms of states of mind has come to be called a 'theory of mind' and its exercise 'mentalizing' or 'mind-reading'. 'Mind-reading' thus refers to the process by which we come to understand others' thoughts and actions in terms of agency and states of mind. The human ability to mind-read raises a number of well-known interdisciplinary questions, which philosophers and psychologists have considered in the debate over theories of 'theory of mind'.¹ These questions concern the process of mental state attribution, the development or acquisition of such mind-reading skills and the contents of mental concepts. Three approaches have been influential among the philosophers and psychologists who have considered these questions: (1) theory-theory, (2) modularity theory and (3) mental simulation theory. The theory-theory

¹Cf. P. Carruthers and P. K. Smith, *Theories of Theories of Mind*, Cambridge 1996.

claims that we rely on an implicit psychological theory in forming beliefs about others' states of mind, a theory which is acquired and refined much like scientific theories in general. Modularity theory attributes our mind-reading ability to the maturation of an innate mental module, a genetically determined product of human evolution. Simulation theory maintains that we form beliefs about others' states of mind not by deploying a theory or relying on an innate module, but by using our own mental resources to simulate the other. Each of these approaches claims to offer an adequate account of our everyday ability to read others' minds. However, upon closer examination all three approaches exhibit serious shortcomings. It will be argued that as they stand, none of these approaches is adequate, especially if we focus on their accounts of the everyday mental concepts involved in mind-reading.²

1. Empirical vs. Conceptual Considerations

Citing states of mind as reasons for people's actions and thoughts comes so naturally to most of us that we forget that this way of understanding each other is a capacity that is not shared by all human adults, as cases of autism and Asperger's syndrome illustrate.³ Nor is the ability to mind-read, i.e. to understand others' thoughts and behavior by discriminating between states of mind, fully present at birth, although infants and pre-verbal children understand important aspects of the mind. Even those humans, who become proficient at mind-reading, evidence a distinct developmental pattern in the course of their maturation. Significantly, our nearest primate relatives, the great apes, lack the capacities that we humans bring to mental attribution and explanation. No nonhuman primate appears to understand the behavior of conspecifics intentionally, although all primates understand the behavior of conspecifics as animate and directed, e.g. chimpanzees have some understanding of others' intentions and knowledge.⁴ Nonhuman primates do not appear to possess a full-fledged 'theory of mind'. Human beings differ from nonhuman primates not only with respect to their theory of mind, but in having the concept of mind and concepts of states of mind, which enable them to understand their conspecifics as minded agents. The evidence from numerous empirical studies in psychopathology, developmental psychology, primatology and neuroscience points to mind-reading as an essentially human capacity.⁵

In the theory of mind debate philosophers have focussed on conceptual issues concerning the nature of our understanding of mind and the contents of everyday mental concepts, while developmental psychologists have studied the processes of mental state attribution in normal children experimentally and proposed accounts of the underlying mechanisms. Comparative psychologists have investigated nonhuman primates' ability to understand conspecifics' mental states in an effort to throw light on the evolution of mind-reading abilities. In view of

² For stylistic reasons 'states of mind' and 'mental states' will be used interchangeably. The use of the latter term should not be taken to imply the position called 'individualism' in the philosophy of mind, i.e. the position that states of mind are specifiable with reference to the internal states of an individual alone and without reference to context.

³ Cf. U. Frith 1989 *Autism: Explaining the Enigma*, Oxford; F. Happé 1994 *Autism: An Introduction to Psychological Theory*, London; S. Baron-Cohen 1995.

⁴ D. Povinelli *et al.*, 1990; M. Tomasello and J. Call, 1997.

⁵ Cf. S. Baron-Cohen, H. Tager-Flusberg and D. Cohen (eds.), *Understanding Other Minds: Perspective from Developmental Cognitive Neuroscience*, Oxford 2000; P. Carruthers and A. Chamberlain (eds.), *Evolution and the Human Mind: Modularity, Language and Metacognition*, Cambridge 2000.

the importance of empirical issues concerning processes of mental state attribution and their acquisition or development, it might seem that philosophy has little to contribute to the topic of theory of mind. One might argue that scientific theory construction, hypothesis testing and experimental results are called for, not conceptual inquiry and reflection on beliefs and forms of thought. However, this conclusion would be too quick. Empirical inquiry must generally be informed by reflection on its conceptual presuppositions and foundations. However, in the study of mind, in particular, conceptual considerations are inescapable and indispensable, because our everyday concepts of states of mind play a constitutive role in defining the subject matter. Our idea of what mind is largely depends on our ordinary concepts of states of mind – beliefs, desires, hopes, fears, intentions and the like -- and to a large extent these concepts determine what is to be investigated in the study of mind. Thus, if empirical psychological theories of mind are to be adequate, we must be able to recognize ourselves and our attributions of states of mind in their descriptions, concepts and structures.

2. Mind-Reading and Mental State Attributions

What does human mind-reading involve? How do we recognize and understand others' states of mind? One place to start an examination of mind-reading is with the everyday attributions we as mature adults make in the course of explaining and anticipating others' thoughts and actions. Commonly, we characterize both ourselves and others in psychological or mental terms, e.g. as '*believing* that Zeche Zollverein is in Essen' or '*wanting* to visit the Gasometer in Oberhausen'. The ability to characterize oneself and others by means of such predicates is a central feature of the psychological competence of normal adults. It underwrites the self-attribution of beliefs, desires, emotions and other conscious, occurrent states of mind and subserves the attribution of states of mind to others. In attributing a state of mind, an attributor characterizes a mental state token as being of a particular type, e.g. '*believing* that Zeche Zollverein is in Essen' or '*wanting* to visit the Gasometer', in contradistinction to other mental state types. In the case of sensations and perceptions, the modality must be specified. In the case of beliefs, desires and other "propositional attitudes", one must characterize both the mental state type, e.g. '*believing*', and the content, 'that Zeche Zollverein is in Essen'. Plausibly, the phrases '*hoping* that Zeche Zollverein is in Essen' and '*believing* that Zeche Zollverein is in Essen' characterize different mental states and express different mental state concepts, because the mental state types are different. Whereas, '*believing* that Zeche Zollverein is in Essen' and '*believing* that the Tetraheder is in Bottrop' are different mental states, because their contents differ, even though the mental state type is the same. If the attribution of a mental state, whether to oneself or another, consists in characterizing the state as being of one type or another, then attributing a mental state involves a conceptual act, a judgement. On this view the notion of a mental state type like believing, hoping, wanting, etc. is the notion of a mental concept.

We usually attribute states of mind to others by using a sentence of a language we know, such as "She intends to climb the Tetraheder ", "He wants to visit the Gasometer" or "She has a headache".⁶ Such attributions generally exhibit a grammatical subject denoting the person to

⁶ Even though attributions of states of mind are usually linguistically formulated, they need not be explicitly verbalized, e.g. an attribution need not issue in a linguistic utterance. The linguistic character of such attributions stems from the fact that many attributed states of mind concern 'propositional attitudes', e.g. beliefs, desires, hopes, fears, etc., -- states of mind that are identified by their propositional content.

whom the state is attributed, a psychological matrix verb characterizing the psychological or mental kind type and a grammatical complement specifying what the mental state concerns, i.e. an embedded sentence, an infinitive construction or a pronoun, which expresses the content. In the case of psychological states characterized by means of a sentential complement, e.g. beliefs, desires and other “propositional attitudes”, the attributor must characterize both the content of the state and the mental state type. Moreover, the propositional content must be specified in a “fine-grained” manner, if it is to contribute to an understanding of another’s thoughts and actions. With the help of natural language constructions, we are able to make fine-grained distinctions – as many as the semantic resources of the language allow. This is important, because it is imperative to capture the way others conceive of objects, facts and situations, if we want to understand their particular thoughts and actions. For example, ‘wanting to buy a book by Mark Twain’ and ‘wanting to buy a book by Samuel Clemens’ will have different consequences for thought and action, if a person does not know that ‘Mark Twain’ is the *nom de plume* of Samuel Clemens. Setting the question of content to one side, it seems clear that everyday attributions of states of mind involve mental concepts.

The claim that concepts of states of mind play a central role in mind-reading on the part of normal human adults, does not entail that processes at the subpersonal level or nonverbal capabilities at the personal level are irrelevant to our mind-reading abilities.⁷ A number of nonverbal and nonconceptual capabilities at the subpersonal and personal level have been identified as prerequisites for mind-reading -- intermodal and intersubjective mapping, neonate imitation, recognizing goal-directed intentional movement, gaze following, joint or shared attention and still others have been proposed on the basis of neurological research in nonhuman primates, e.g. neural mirror matching systems⁸ However, these are presented as necessary for mind-reading, but not as sufficient. Thus, something more is required. Even where there is evidence for early nonmentalistic understanding of others’ actions based on perception, it is conceded on all sides that reading others’ minds cannot be explained by this understanding alone. Perception-based understanding of action is a form of body-reading and behavior-reading rather than mind-reading.⁹ The idea that we “see” states of mind in movement and body posture is plausible only for a limited number of states of mind -- pain, depression, and other emotional states. In the case of those we know well, e.g. our spouses, siblings or close friends, our ability to perceptually “read” body posture and behavior may be quite reliable, despite mistakes. However, as soon as we consider casual acquaintances, not to mention complete strangers or members of other cultures, the perceptual claim is problematic. In these cases it becomes clear that perception of body posture and behavior alone do not suffice for mind-reading. The same bodily movement could be part of a mugging, a jocular greeting among youths or a saving Heimlich maneuver. There is no direct route from a perceived bodily movement to specific, contentful states of mind, as critics of classical behaviorism emphasized, because a bodily movement or posture does not wear its description

⁷ Pace S. Gallagher in: “The Practice of Mind: Theory, Simulation of Primary Interaction?”, *Journal of Consciousness Studies* 8 (2001), 83-108.

⁸ A. Gopnik and A. Meltzoff 1997; S. Baron-Cohen 1995; D. A. Baldwin and J. A. Baird 2001; A. Leslie 1994; C. D. Frith and U. Frith 1999, A. Meltzoff and W. Prinz 2001; M. Tomasello 1999; Rizzolatti *et al.* 1988, 1996, 2001.

⁹ For criticisms of the perceptual approach advocated by Shaun Gallagher 2001, *cf.* Röska-Hardy 2001.

on its sleeve.¹⁰ Furthermore, there is no direct correlation between types of states of mind and types of behavior. Different, even contradictory, pairings between states of mind and movements or positions are possible. Finally, perceived bodily movements and behavior rarely issue from or depend on a single state of mind, but on several so that one can't move directly from what is perceived to a specific state of mind. These observations suggest that body-reading and behavior-reading alone do not suffice for understanding others' states of mind, even though they may contribute to it.

Reading others' minds in a typically human fashion appears to require concepts of states of mind. What accounts do theory-theory, modularity theory and simulation theory offer of the concepts involved in mind-reading?

3. Theory-Theory

According to the theory-theory, our mind-reading ability rests on a body of implicit general knowledge or theory, which is acquired by learning. In attributing mental states to others, we inferentially apply this folk-psychological theory of mind, which consists in a set of generalizations or laws for the deployment of mental state concepts. Depending on the particular version of theory-theory, the theory involved is conceived as a set of generalizations or laws for the deployment of mental concepts, e.g. in philosophical or analytical functionalism, or as a theory analogous to any other scientific theory.¹¹ Versions differ as to whether the theory in question is acquired through general learning mechanisms or through a process of theory formation analogous to scientific theorizing. However, all versions maintain that mental state attribution involves theory-mediated inference. Theory-theorists hold that we attribute mental states by inferring their occurrence from observation of behavior and environmental events, i.e. by recognizing their causal-explanatory role in accordance with the theoretical generalizations comprising the theory.

On theory-theory, the concepts of mental states employed in mental state attributions are theoretical concepts, which are postulated as a way of explaining and predicting thought and action. A theoretical concept is the concept of a state defined in terms of causal-inferential relationships to publicly observable events in the environment and to overt behavior. Consequently, grasping a mental concept involves mastering theoretical generalizations about causal or inferential relations and mental state representations, e.g. beliefs and desires. On one influential version, mental states such as beliefs and desires are defined functionally in terms of their causal-inferential relations to events in the environment, to other mental states and to bodily behavior.¹² Hence, the concept of a mental state is the concept of a state apt to cause or be caused by certain events or certain types of behavior, i.e. the concept of a particular functional role. On this account, specifying the nature or content of a mental state concept involves generalizations that make reference to dispositions, causal interactions or subjunctive considerations, as Alvin Goldman has pointed out.¹³ It is these generalizations – specifically,

¹⁰ P. Geach 1957, chapter 8.

¹¹ D. Lewis, 1970, 1972; P. Churchland 1988; 1988; A. Gopnik, 1993; J. Perner, 1991, H. Wellman 1990, A. Gopnik and H. Wellman 2000.

¹² D. Lewis, 1966, 1970, 1972; P. Churchland, 1970, 1988.

¹³ A. Goldman 1993.

the causal-functional role the terms play in the generalizations of the theory – that determine the contents of mental concepts.

The theoretical knowledge account of theory-theory entails that an attributor must possess the concept of a particular mental state (type) in order to attribute the mental state to another person. That is to say, at some level one must grasp the mental concept as it is defined by the generalizations of the theory in order to make a corresponding mental state attribution. Thus, even though the attributor's grasp of the theory and its concepts may be "tacit" or "implicit", it is incumbent on the theory-theory to explain how attributors can acquire mastery of the mental state concepts so defined. In particular, theory-theory must explain how *ordinary* attributors acquire and deploy such concepts, because it claims that knowledge of a folk-psychological theory grounds all mental state attributions.

Another version of theory-theory traces the source of our ability to mind-read to a theory of mind that is constructed from evidence in the course of childhood development in a way comparable to the development of scientific theories.¹⁴ Often termed the 'child as scientist' view, this version emphasizes parallels between cognitive development and theory change in science. It is claimed that children's implicit theories of mind reflect the distinctive pattern of explanation, prediction and interpretation that is the hallmark of theoretical structures. Mental state concepts are construed as theoretical constructs, which are posited to provide causal explanations and facilitate predictions. Such concepts are viewed as theoretical, because they are taken to undergo theoretical revision in children's cognitive development similar to theoretical concepts in science. Children's transitions from one developing theory of mind to the next are understood in analogy to transitions in scientific theorizing as in the development of the heliocentric theory of planetary movement from Copernicus to Kepler. Briefly, children start with nonrepresentational notions of desire and perception, as illustrated in the false-belief tasks, and on the basis of evidence and experience revise and develop increasingly accurate theories. By the age of four normal children's theory of mind involves genuinely representational mental concepts like belief just as normal adults' implicit theory of mind does.

As an account of mind-reading, the functionalist theory-theory account of mental concepts is confronted with three fundamental objections.¹⁵ The first concerns the lawlike generalizations posited by the approach, which allow us to infer behavior or thought predictively or retrodictively. Importantly, the contents of mental state concepts are supposed to derive from their role in these theoretical generalizations. However, only a few examples of such generalizations or laws have been explicitly formulated and these are problematic.¹⁶ This is surprising, if such generalizations are the source of our everyday competence in mind-reading. Moreover, the laws and generalizations which have been explicitly formulated appear to be conceptual in character, not empirical and contingent.¹⁷ A second difficulty stems from the theory-theory's specification of the contents of mental state concepts in terms of a network of relations between environmental or proximal stimuli like perception and internal states, between mental states and other types of mental states, internally construed, and between

¹⁴ A. Gopnik and A. Meltzoff 1997, Perner 1991, A. Gopnik and H. Wellman 1992, 2000.

¹⁵ A third objection to the theory-theory account concerns our ability to self-attribute mental states, *cf.* Röska-Hardy 2000.

¹⁶ *Cf.* P. Churchland 1970, A. Goldman 1989.

¹⁷ Röska-Hardy 2001.

internal mental states and behavioral outputs. This specification entails that mental state concepts have subjunctive or counterfactual properties, as Alvin Goldman has emphasized.¹⁸ How does an ordinary attributor grasp such properties in mind-reading, given that he has no information about the immediate causes of the other's thought or behavior and need not have any information about its effects? Consequently, it is unclear how attributors can deploy mental concepts defined in terms of such lawlike generalizations in mind-reading. This is a serious defect, since on theory-theory the lawlike generalizations determine both the contents of mental state concepts and the meanings of the mental state predicates employed in attributions. Without some account, it is unclear how attributors are able to master the "theoretical" concepts of the theory in order to attribute states of mind. Finally, lawlike generalizations by their very nature apply to classes of phenomena; they are not designed to capture particular cases. However, in attributing mental states, we generally are concerned to understand a specific thought or an action of a particular individual in a concrete situation. For example, we often want to understand why someone did something, even if there is no chance that she will ever do it again. The particularity involved in understanding others' thought and action suggests that the lawlike generalizations of theory-theory must be supplemented in order to account for our everyday mind-reading ability.

Difficulties also face the child-as-scientist version of theory-theory. In particular, it is amazing that all normal children converge on the same theory of mind within a similar time-frame. Moreover, empirical data indicates that developmental patterns and end-state mind-reading ability are quite similar among normal humans and even across cultures.¹⁹ If children develop a theory of mind on the basis of evidence like a scientist, one would expect greater variation in the resulting theories. As modularity theorists have noted, this version of theory-theory construes children as outstanding theorists where the mind is concerned, because the concepts of belief, desire and other "propositional" attitudes are quite complex. However, children do not seem to excel theoretically in other areas.²⁰ These shortcomings are often taken to suggest that our capacity to mind-read is based on a body of implicit general knowledge that is inherited as an innate module comparable to Noam Chomsky's language module.

4. Modularity Theory

Although modularity theories are often construed as versions of theory-theory, there are good reasons for treating the two separately. For modularity theorists the core process in mentalistic attribution is not theorizing. Modularity theories construe cognitive structures like the apparatus of mental state attribution as the result of innate modules, not as the product of learning or theorizing; they are not the consequence of an *acquired* theory. According to modularity theory, the mind-reading abilities involved in the theory of mind are due to innately specified modules and result from innately specified developmental processes. These are not developed from evidence in the course of ontogenetic development. Rather, they are created from pre-determined representations of input, triggered by experience from the environment. On this view the theory of mind is a domain-specific ability, supported by an

¹⁸ A. Goldman 1993.

¹⁹ G. Segal 1996.

²⁰ A. Leslie and D. Roth 1993.

innate, encapsulated and domain-specific module, whose function is segregated from the other intellectual capacities of the individual.²¹

Leading proponents of modularity, e.g. Noam Chomsky, Jerry Fodor, Simon Baron-Cohen and Alan Leslie, differ significantly in their respective positions and specific claims about innate modules.²² However, the central idea of modularity theory is that the contents of everyday mental concepts in the theory of mind are part of a special purpose body of knowledge in a mental module, which is innate and matures through a process of ontogenetic development. Although modularity theory, like theory-theory, construes mental state concepts as abstract theoretical postulates, embedded in causal laws, it claims that such concepts are part of our innate endowment as humans. Specifically, mental state concepts like belief and desire are the product of an innate structure dedicated to interpreting behavior in terms of beliefs and desires. Consequently, the development of mind-reading is genetically determined and not based on empirical theory construction.

On modularity theory mental state concepts or specialized mind-reading processors are construed as innate and merely triggered by certain types of experience and maturation. Theorists like Jerry Fodor hold that the mind is composed of independently functioning, genetically specified modular systems, but emphasize innate concepts, maintaining that mental state concepts -- like all other concepts -- are present at birth, while modularists like Baron-Cohen reverse the emphasis. Others follow Alan Leslie in proposing a distinct Theory of Mind module which begins to develop with the onset of pretend play at 18-months of age.²³ On Fodor's view, development of mind-reading abilities is mainly an increase of information processing capacities, which allow children to better use what they already know, while innate processor theories like Baron-Cohen's postulate prespecified processors which create mandatory interpretations of human behavior in terms of mental states. These processors take human behavior or information about the other person as input and generate explanations for that behavior, predictions of what the person will do, etc. as output. Baron-Cohen posits four processors: an intentionality detector that determines agentive movement in terms of goals and desires, an eye-direction detector, a shared-attention mechanism and a ToM Mechanism to link agents via mental attitudes to propositions.²⁴ Alan Leslie's theory is similar, but involves three modules: a ToM Mechanism, a module to impute agency and a module to interpret physical motion.

For example, Alan Leslie postulates several different modules that come on line sequentially in children's developing theory of mind.²⁵ He hypothesizes that a domain-specific mechanism or information processing device, the ToM Mechanism (ToMM), computes certain data structures called metarepresentations, which specify attitude, agent, an anchor (an aspect of the real situation) and a pretend or imagine state. Modularity theories like Leslie's ToMM hypothesis cite our innate endowment as members of the human species as the source of our mind-reading abilities, but offer no further elaboration of mental state concepts

²¹ A. Leslie 1987, 1999; Baron-Cohen 1995; Segal 1996; Fodor 1992, 1994.

²² However, they are all influenced by Noam Chomsky's model of a language acquisition device that is innately preprogrammed to recognize language and test linguistic input against a constrained and preordained set of test hypotheses about language structure, *cf.* Chomsky 1986.

²³ A. Leslie 1987, 1992.

²⁴ S. Baron-Cohen 1995.

²⁵ A. Leslie 1987, A. Leslie and D. Roth 1993, 1994.

as such. For example, Leslie does not explain how the concepts of belief or desire are represented in the postulated metarepresentations or how specific mental state concepts differ from each other, e.g. belief *vs.* desire. Thus, beyond modularity theory's claim that mental concepts are innate, the content of mental concepts remains unclear.

Taking stock, modularity theories posit genetically specified mechanisms that operate over mental state concepts like belief, desire, fear and intention in order to account for our mind-reading capacities. They construe mental state concepts as part of the innate endowment of our species, but they offer no explication of the contents of these concepts. At the very least, modularity theories need to say more about the concepts involved in mind-reading and their relation to language in order to account for our everyday capacity to attribute states of mind to others.

5. Simulation Theory

The proponents of mental simulation deny that attributions of states of mind rely on theoretical knowledge and inference alone. Simulationists take the first-person point of view to be essential to states of mind and mental state concepts, despite quite different views on the nature of simulation, e.g. on whether it involves analogical inference or employs "ascent routines". They also differ over whether simulation requires the prior possession of mental state concepts. Despite the differences in detail, however, simulation theorists hold that one uses one's own mental resources to attribute mental states by pretending or imagining oneself to be in the other's position and then generating the thoughts or actions attributed to the other within a mental simulation without benefit of theory.

On Alvin Goldman's cognitive-scientific simulation approach, one attributes mental states to others by using one's own cognitive and inferential mechanisms to match or replicate those of the other person.²⁶ For example, to determine what decision another person will make, an attributor pretends to have the beliefs and desires that she takes the other to have and feeds these into her own practical-reasoning system, which generates further mental states as outputs. The "decision" output of the simulation is then attributed to the other person by analogical inference. In this version of simulation the attributor's cognitive system operates over pretend or surrogate states; it is run "off-line" and consequently does not issue in action. Instead, the output of the simulation process provides the basis for attributing a decision or mental state to the other. In contrast to theory-theory, Goldman's approach does not require that an attributor have any knowledge of the system employed or of its properties. He must only "possess" such a mechanism and be able to "feed in" appropriate inputs and to construe its outputs in mental terms. Simulation is "process-driven", not "theory-driven". His approach is premised on the assumption that pretend mental states and non-pretend mental states are homologous as well as on the assumption that attributors and attributees are similar with respect to practical and cognitive mental processes. Simulation terminates in an analogical inference from oneself to the other; the attributor infers the mental state of the other from her recognition of the pretend state issuing from her simulation of the other. Hence, the end product of a simulation is a judgement, which involves classifying an occurrent mental state as a token of a particular mental state type and consequently requires mental concepts. In Goldman's opinion, the simulation approach does not provide an account of such concepts.

²⁶ A. Goldman 1989, 1993, 2000.

Goldman favors a first-person approach to mental concepts and proposes an account of mental state concepts in terms of introspection or self-monitoring. In his view we have a first-person, introspective understanding of mental state concepts which rests on direct, non-inferential access to our conscious mental states. He has suggested that mental concepts such as desire, belief, etc. are understood partly in terms of intrinsic, non-dispositional characteristics of conscious experience – perhaps qualitative or phenomenological characteristics, which are introspectively accessible to the subject of the experience. He hypothesizes that subjects can detect these characteristics and distinguish them from one another, i.e. that they are epistemically identifiable, and suggests that such internally detectable characteristics underlie our grasp of mental concepts.²⁷ Recently Goldman has advanced a “dual representation hypothesis” concerning mental concepts.²⁸ He proposes that people develop two sorts of mental representations for some mental states, e.g. desire. The idea is that people come to understand certain behavioral representations and representations of certain inner characteristics as representations of one and the same sort of state. Goldman cites research on resonance phenomena in the ventral pre-motor cortex -- ‘the mirror matching mechanism’ – in support.²⁹

This suggestion notwithstanding, Goldman’s claim that fully grasping concepts of “mental representational states” like beliefs partly involves latching representations or conceptual structures onto introspectible characteristics of mental states is problematic. First, if introspective detection of phenomenological characteristics is necessary in order to possess mental state concepts, how does the boot-strapping get started? How does one initially identify a mental state token as having a certain, non-relational categorial property? Moreover, which properties qualify a state as a *belief* in contradistinction to a *hope*? Secondly, how does a person re-identify an inner property *p* of a mental state type as *p*? Thirdly, how does one come to associate the expressions of a natural language with introspectively accessible features of mental states? Finally, what reason is there to assume that others have mental states similar to one’s own? This assumption is central for Goldman’s introspectionist, analogical inference account of mind-reading. Even if we assume that the analogical inference from oneself to the other is not problematic, it remains the case that on Goldman’s account the introspectively accessible properties of one’s mental states stand in no relation to one’s environmental circumstances, perceptual situation or behavior. If the contents of mental concepts are primarily determined by reference to intrinsic, categorial properties, then behavior and changes in the environment would seem to play no systematic role. However, the latter seem to be one reason for assuming that mental concepts apply to others as well as to ourselves. For we take certain events in the environment and certain behaviors as indicative of others’ mental states, as Wittgenstein emphasized with his idea of outward criteria. Moreover, if “external” factors have no systematic relation to the contents of mental state concepts, how does an attributor identify the initial inputs for a simulation routine, i.e. the mental states of the person to be simulated?

Since Goldman only claims that introspective access is necessary for mental concepts, not that it is sufficient, much depends on how he augments the introspective core of his

²⁷Goldman hypothesizes that we may recognize a particular mental state as a token of a mental state type much as we recognize a token object, say a chair, as an object of a certain type, by using current available information about the token to “match” it in non-cognitive fashion with a representation or pattern stored in memory.

²⁸ V. Gallese and A. Goldman 1998, A. Goldman 2000a, 2000b.

²⁹ G. Rizzolatti 1988, 1996, 2001.

simulation account. However, it is clear that resources beyond first-person introspective access are needed to provide an adequate account of the content of everyday mental state concepts and their role in our capacity to read others' minds.

In contrast to the introspection-simulation account, Robert Gordon's simulation approach eschews introspective access and analogical inference in accounting for everyday our attributions of mental states.³⁰ He elaborates his simulation approach to reading others' minds in tandem with an account of mental self-attribution in terms of 'ascent routines'. He contends that ascent routines allow us to ascribe beliefs and other mental states to ourselves accurately and reliably as well as non-inferentially and without appeal to evidence, e.g. without consulting environmental circumstances or behavior, as the theory-theory would have it, or introspectible properties, as Goldman suggests. For example, one answers a question concerning a mental state, such as "Do you believe that Mickey Mouse has a tail?", by asking oneself an "outward-looking" question about the world, "Does Mickey Mouse have a tail?" and then "ascending" with the answer from the "object-level" ("Yes, Mickey Mouse has a tail") to a level which concerns the mental state of belief ("Yes, I believe that Mickey has a tail"). Gordon calls this procedure an 'ascent routine' because it answers a question about belief, i.e. a question about a mental state that is about whether I *believe p*, by answering a question "at a lower semantic level", namely a question which is directly about *p* (M.M. has a tail). According to Gordon, the ascent routine allows one to move from "*p*" to the self-ascription of a mental state without exercising introspection, reasoning or inference. In his opinion, the ascent routine procedure reflects the way adults usually determine whether or not they believe that *p*; they simply ask themselves the question whether or not *p*. He claims that even those who possess the concept of belief rely on an ascent routine when answering questions about what they believe. He hypothesizes that there are ascent routines for propositional attitudes other than belief and also extends the ascent routine proposal to the ascription of pain. Thus, the ascent routine procedure forms the basis for making self-attributions of mental states.

Gordon's version of simulation does not require an analogical inference from oneself to another. Rather, in simulating another person, the attributor "recenters her cognitive map" on the other so that the first-person pronoun 'I' refers "exclusively" to the individual on whom the attributor's egocentric map has been recentered. Gordon describes this recentering as "imaginatively transforming ourselves into other 'first persons'", as "identifying with" an individual within simulation or as "becoming" the other. After the imaginative transformation into the other, a simulator directly attributes the belief or decision generated within the scope of simulation to the other *via* an ascent routine, thus obviating the need for an analogical inference from herself to the other. Attributing mental states to another is, in essence, a case of mental "self"-attribution to oneself-as-the-other within the context of simulation. This is because embedding an ascent routine within a simulation of another allows one to attribute mental states to the other directly by ascent routine. Thus, for Gordon an egocentric shift on the part of the attributor with an ensuing ascent routine lies at the core of mind-reading.

Gordon claims that the ascent routine procedure indicates how one might acquire the concept of a mental state like belief as well as the capacity to make genuine, comprehending attributions of mental states, because it provides a way of reconceptualizing pains and beliefs as having a mental location. However, this is not the case. The meanings of mental predicates and a grasp of mental concepts must be presupposed in order to get the ascent routine started.

³⁰ R. Gordon 1986, 1992, 1995, 2000.

One has to understand that the initial question is about belief, not about hope or desire. Without further elaboration the ascent routine procedure does not supply the foundation for an account of mental state concepts. Furthermore, the crucial assumption that one can mentally “transform” oneself into the other, “recenter” one’s egocentric map or “become” the other needs to be clarified in order to provide an account of mental state attribution. What do these metaphors amount to? Finally, how does one “recenter” or “mentally transform oneself into the other”? Recentering cannot generally be accomplished directly, e.g. on the basis of spatial or temporal perspective-taking, because an attributor will only rarely be in the same physical environment as the person she is simulating and never have the same personal history. And, as hermeneuticists from Schleiermacher to Gadamer have emphasized, our beliefs, desires, and other thoughts determine not only how we view “the world”, but also what we take to be “the facts”. In essence, one must take into account those features of a situation which are salient for the other in order to mentally simulate the other. This requires substantive assumptions about the other’s beliefs, desires, intentions, etc., which must be specified in contentful terms, i.e. treated as a belief that *p*, a hope that *q*, and so forth in order to generate further thoughts or decisions. Clearly, more must be said in order to account for our everyday ability to read others’ minds in terms of simulation with ascent routines.

6. Conclusion

In view of the deficits of theory-theory, modularity theory and simulation theory, taken individually, a richer account of our everyday capacity to understand others in terms of intentional action and states of mind is required. However, the suggestion that we sometimes use theory and sometimes simulation in attributing states of mind to others, does not present an adequate alternative, because it does not provide an account of how we ordinary attributors understand the mental state concepts involved in making mental state attributions. Thus, it leaves out a central aspect of our typically human way of reading others’ minds. Such hybrid approaches are also silent on developmental issues. A more systematic approach will integrate developmental claims about mind-reading abilities with theoretical claims, for developmental claims constrain claims about end-state mind-reading capacity. For example, one could hardly claim that our ability to make mental state attributions rests on simulation, while maintaining that we acquire this ability by theory construction.

What direction should an adequate account of the human ability to mind-read take? If we consider our everyday mental state attributions and the mental state concepts they involve, the shortcomings of theory-theory, modularity theory and simulation theory offer a clue. In all three approaches social interaction and the mastery of a natural language are hardly considered. The important role of the social matrix in cognitive development is obscured by focussing on the individual attributor and his or her cognitive abilities, which are construed in terms of internal representations. The crucial role of language use and its systematic connections to social contexts is also ignored. A different approach to mind-reading is suggested, if we take interaction in social contexts and the ability to use natural languages seriously, as L. Wittgenstein, L. Vygotsky and J. Bruner have.³¹ Such an approach need not deny that human beings are born with innately determined capacities and cognitive dispositions, quite the contrary; certain social skills and certain aspects of language may be

³¹ L. Wittgenstein 1953/1958; L. Vygotsky 1962, 1978; J. Bruner 1990, M. Tomasello 1999.

largely innately determined. Capacities like intermodal and intersubjective mapping, imitation, recognizing intentional movement, gaze following and joint attention clearly subserve the development of social competencies and underpin the acquisition of language. However, they are not sufficient to explain typically human mind-reading, if the ability to understand others in terms of thoughts and actions is a cognitive achievement, for concepts of states of mind have yet to enter. Given the crucial importance of language in human cognition, more attention should be devoted to the intersection of pre-verbal social competencies and the acquisition of language in social contexts by those who aim to account for our ability to mind-read – philosophers and psychologists alike. Here it is important for all parties to the theory of mind debate not only to consider the empirical data on cognitive development and deficits, while critically reviewing the theoretical commitments of the respective investigators, but also to relate proposed accounts to our everyday way of attributing mental states. Exploring the connection between social interaction, language acquisition and the ability to understand others' thoughts and actions by reference to mental states is an interdisciplinary project, which promises to yield a more adequate account of the way we read others' minds.

*Literature**

Baldwin, D.A. and J.A. Baird 2001 "Discerning intentions in dynamic human actions" in: *Trends in Cognitive Sciences* 5, 171-8.

Baron-Cohen, S. 1995 *Mindblindness. An Essay on Autism and the Theory of Mind*. Cambridge, MA.

Baron-Cohen, S., A. Leslie and U. Frith 1985 "Does the autistic child have a Theory of Mind?" in: *Cognition* 21, 37-46.

Bruner, J. 1990 *Acts of Meaning* Cambridge, MA.

Chomsky, N. 1980 *Rules and Rerpresentations*, New York.
1986 *Knowledge of Language: its Nature, Origins and Use*, New York.

Churchland, P. 1970 "The Logical Form of Action Explanation" in: *Philosophical Review* 79, 214-236.
1988 *Matter and Consciousness*, Cambridge, MA.

Fodor, J. 1983 *The Modularity of the Mind*, Cambridge, MA.
1987 *Psychosemantics*, Cambridge, MA.
1992 "A Theory of the Child's Theory of Mind" in: *Cognition* 44, 283-96.

Gallese, V. and A. Goldman 1998, "Mirror Neurons and the Simulation Theory of Mind-Reading"
in: *Trends in Cognitive Sciences* 2, 493-501.

Geach, Peter, *Mental Acts*, London 1957.

Goldman, A. 1989 "Interpretation Psychologized" in: *Mind and Language* 4: 161-85.
1993 "The Psychology of Folk Psychology" in: *Behavioral and Brain Sciences* 16, 15-28.
2000a "Folk Psychology and Mental Concepts", *Protosociology*. 14, 4-25.
2000b "The Mentalizing Folk" in: *Metarepresentation: A Multidisciplinary Perspective*, D. Sperber (ed.), Oxford.

Gopnik, A. 1993 "How We Know Our Own Minds: the Illusion of First-Person Knowledge of Intentionality"
in: *Behavioral and Brain Sciences* 16, 1-14.

Gopnik, A. and A. Meltzoff, 1997 *Words, Thoughts and Theories*, Cambridge, MA.

Gopnik, A. and H. Wellman 1992 "Why the Child's Theory of Mind really is a Theory" in: *Mind and Language*

** I would like to thank Brigitte Blockhaus und Gesine Worm of KWI for their assistance in procuring literature.

- 7, 145-171.
 2000 "The Theory Theory" in: D. Sperber (ed.), *Metarepresentation: A Multidisciplinary Perspective*, Oxford.
- Gordon, R. 1986 "Folk Psychology as Simulation" in: *Mind and Language* 1, 158-71.
 1995 "Simulation without Introspection or Inference from Me to You" in: *Mental Simulation*, M. Davies and T. Stone (eds.), Oxford.
 1996 "'Radical' Simulationism" in: *Theories of Theories of Mind*, P. Carruthers and P. K. Smith (eds.), Cambridge.
 2000 "Sellars's Ryleans Revisited" in: *Protosociology*. 14, 102-114.
- Heal, J. 1995 "How to Think about Thinking" in: *Mental Simulation*, M. Davies and T. Stone (eds.), Oxford.
 1996 "Simulation and Cognitive Penetrability" in: *Mind and Language* 11, 44-67.
 1998 "Co-cognition and Off-line Simulation: Two Ways of Understanding the Simulation Approach" in: *Mind and Language* 13, 477-498.
 1998 "Understanding Other Minds from the Inside", in: O'Hear (ed.), *Current Issues in the Philosophy of Mind*, Cambridge and in: *Proto-Sociology* 14.
 2000 "Other Minds, Rationality and Analogy" in: *Proceedings of the Aristotelian Society*, Supplementary vol. LXXIV.
- Leslie, A. 1987 "Pretense and Recognition: the 'Origins' of the Theory of Mind" in: *Psychological Review* 94, 412-26.
 1992 "Autism and the 'Theory of Mind' Module", *Current Directions in Psychological Science* 1, 18-21.
 1994 "Pretending and Believing: Issues in the Theory of ToMM" in: *Cognition* 50, 211-38.
 2000 "How to acquire a 'representational theory of mind'", in: *Metarepresentation: A Multidisciplinary Perspective*, D. Sperber (ed.), Oxford.
- Leslie, A. and D. Roth 1993 "What Autism teaches us about Metarepresentation" in: S. Baron-Cohen, H. Tager-Flusberg and D. Cohen (eds.), *Understanding Other Minds: Perspectives from Autism*, Oxford.
- Lewis, D. 1966 "An Argument for the Identity Theory" in: *Journal of Philosophy* 63, 17-25.
 1970 "How to Define Theoretical Terms" in: *Journal of Philosophy* 67, 427-446.
 1972 "Psychophysical and Theoretical Identifications" in: *Australasian Journal of Philosophy* 50: 249-58.
- Meltzoff, A. N. 1995 "Understanding the Intentions of Others: Re-enactment of intended acts by 18-month-old children" in: *Developmental Psychology* 31, 838-50.
- Meltzoff A. N. and M. Moore 1983, "New born infants imitate adult facial gestures" in: *Child Development* 54, 702-9.
- Meltzoff, A. N. and W. Prinz, W. 2001, *The Imitative Mind: Development, Evolution and Brain Bases*, Cambridge.
- Perner, J. 1991 *Understanding the Representational Mind*, Cambridge, MA.
- Povinelli, D. and K. Nelson, "Inferences about guessing and knowing in chimpanzees" in: *Journal of Comparative Psychology* 104, 203-20.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G. and Matelli, M. 1988 "Functional organization of interior area 6 in the macaque monkey: II. Area F5 and the control of distal movements, in: *Experimental Brain Research* 71, 491-507.
- Rizzolatti, G., Fadiga, L. Matellie, M. Bettinardi, V., Paulesu, E., Perani, D. und Fazio, G. 1996 "Localization of grasp representations in humans by PET: 1. Observation vs. Execution" in: *Experimental Brain Research* 111, 246-252.
- Rizzolatti, G., Fogassi, L., Gallese, V. 2001 "From Mirror Neurons to Imitation: Facts and Speculations" in: A. Meltzoff and W. Prinz (eds.), *The Imitative Mind: Development, Evolution and Brain Bases*, Cambridge.
- Röska-Hardy, L. 2000 "Self-Attribution and Simulation Theory" in: *Protosociology* 14, 115-144.

- 2001 "Zuschreibungen und Personenverstehen" in: Th. Rentsch (ed.), *Sprache, Erkenntnis, Verstehen. Grundfragen der theoretischen Philosophie der Gegenwart*, Dresden.
- Scholl, B. J. and A. Leslie 1999 "Modularity, Development and 'Theory of Mind'" in: *Mind and Language* 14:131-53.
- Segal, G. 1996 "The Modularity Theory of Mind" in: P. Carruthers and P. Smith (eds.), *Theories of Theory of Mind*, New York.
- Tomasello, M. 1999 *The Cultural Origins of Human Cognition*, Cambridge, MA.
- Vygotsky, L. 1962 *Thought and Language*. Cambridge, MA.
1978 *Mind in Society: The Development of Higher Psychological Processes*, M. Cole, V. Johns-Steiner, S. Cribner and E. Souberman (eds.), Cambridge, MA.
- Wellman, H. M. 1990 *The Child's Theory of Mind*, Cambridge, MA.
- Wittgenstein, Ludwig 1953/1958 *Philosophical Investigations/Philosophische Untersuchungen*, New York.
1958/1972 *The Blue and Brown Books*, Oxford.