Albert Newen and Anika Fiebich

A developmental theory of self-models: individual-cognitive and social-cognitive dimensions of self-consciousness

1. Introduction

Self-consciousness can be defined as the ability to have conscious experiences and on that basis to represent one's own states (processes and events) as one's own (Newen & Vogeley, 2003). This includes especially but not only mental phenomena. Self-consciousness can be instantiated in dealing with bodily features, too: I can express self-conscious states linguistically, e.g. by saying, "I am in pain", "I am wearing black shoes" or "I believe that my red-experience is untypical" (because I am red-green blind). The self can be defined as the bearer of self-conscious states. Since we are presupposing naturalism about self-consciousness (for a methodological discussion see Newen & Vogeley, 2008), the self is identified with a human being as a natural entity having the special ability of self-consciousness and thereby of establishing representations about herself. The content of these representations about oneself can be called a self-model. This use is close to that of Thomas Metzinger (2003). But – pace Metzinger – it is important to draw a distinction between the self (the human being having specific self-representations) and the self-model (the content of the self-representations a human being actually has or that are accessible to her).[1] The term "self" is used interchangeably with the term "I" to refer to a human being by characterizing the specific epistemic status that is connected with self-consciousness. In the case of competent speakers, the specific epistemic status is typically expressed by using the word "I". It has been shown that with the use of "I", we express an immediate self-representation (*de se representations*) that can and has to be distinguished from representations which are "de facto" about me – although I may not notice that (*de re representations* about me) (Perry, 1979). Bearing this naturalistic background in mind, we want to investigate how the contents of the self-representations (i.e. self-models) develop and to which extent they are constituted by social interactions. Before investigating this key question, we have to prepare the platform of the discussion (1) by arguing that the immediate self-representation, which is characteristic for self-consciousness, does not necessarily involve linguistic competencies, and (2) by introducing a background theory of mental representations which allows us to distinguish different levels of self-consciousness

---

[1] It has been shown that if someone is not drawing this distinction (like Metzinger (2003)), then he runs into trouble with basic linguistic evidences, e.g. that the word "I" refers in all its uses to the speaker of the utterance and not to the content of a bunch of self-representations. Metzinger accepts the semantically implausible claim that the word "I" is ambiguous by referring sometimes to the human beings making an utterance of the term "I" and sometimes to the content of self-representations. For further discussion of his theory see Newen (2003a).

and different kinds of self-models. For each self-model, we can then investigate to which extent it is constituted by individual-cognitive properties (representations of space, time, causality, quantity etc.) and/or by social-cognitive properties (social learning, preverbal and linguistic communication, theory of mind etc.). In the debate about the social dimension of self-consciousness, we will show that neither a pure concentration on the individual-cognitive properties (as paradigmatically put forward by Piaget (1970; 2003)) nor a pure concentration on the social foundations of our life (as paradigmatically developed by Mead (1934)) is adequate. Especially an adequate description of the early development of life needs a systematic consideration of both dimensions and their interaction. Therefore, we suggest a developmental theory of self-models.

2. The debate on the status of self-models: how social is our self-model?

In the tradition, we find an intensive and still continuing debate concerning the status of the self-model[2] which can be divided into two main schools of thought: One school claims that the self-model is completely determined by social-cognitive properties of the human being, i.e. by cognitive properties, which essentially rely on a system-system-interaction; the opponent school argues that the self-model is mainly constituted by individual-cognitive properties of a person. According to the latter, the self-model is just one cognitive phenomenon in the cognitive development of an individual (and her experiences), which can be essentially characterized independently from the social environment, only relying on a system-environment-interaction (i.e. an interaction with the physical environment).

A starting point of the debate was Baldwin's work (Baldwin, 1897). He claims that the human being as a social individual is a product of social life that follows from a "dialectic of personal maturation", that is the dialectic of giving and taking between the individuals and their fellow men. The interrelation of habit and accommodation (this means adaptation to experiences and change of habits) is an important influencing factor for human development. Baldwin subsumes this interrelation under the notion of "imitation", which he takes to provide a sufficient explanation for the development of the self, including moral, religious and aesthetic aspects. Since his writings fell into obscurity relatively quickly, we do not go into detail here. But nevertheless, they influenced the famous writings of other scientists (see Garz, 2006, for review). Mead (1934) offers the most prominent view defending the claim that the self-model is completely constituted by social features, e.g. by being member of a social group. He was also inspired by William James (1890). Mead's main consideration explains the origin of a self with the ability of symbolic interaction: the self-model is constituted by speech-behaviour in form of gestures and utterances ("symbolic interactionism"). In the same line, Habermas (1987) construes the whole genesis of the self-model as a social process. The second branch of considerations about the self is at least going back to Descartes (1641; 1992). He is a paradigmatic representative of a philosopher

---

[2] Although in the literature discussed below the researchers often speak about the *self*, they usually mean the *self-model*, i.e. the representational content that a human being develops about herself. Having this distinction in mind, in the quotations below it is often not marked explicitly unless we think that there is a danger of misunderstanding.

who thinks about the self in isolation from the social environment. Since we are presupposing a naturalistic framework to account for the self as a natural entity, which has the characteristic ability of self-consciousness, we consider neither dualistic theories nor theories, which deny the reality of the self (e.g. Hume's bundle theory of selves (Hume, 1748; 2006) or Metzinger's theory of phenomenal self-consciousness (Metzinger, 2003)). The most important representative of the claim that a self-model is constituted by individual-cognitive properties is Piaget. In his seminal work, he investigated the cognitive development of children. According to Piaget (1970; 2003), organization and adaptation are the two core concepts of cognition, which are essential to enable development at all. Organization is the tendency of all organisms to order and coordinate their life systematically. Adaptation (which contains equilibrium between assimilation and accommodation) is the tendency of all organisms to adapt to their particular environments. Piaget distinguishes four stages of cognitive development: 1) sensorimotor stage, 2) pre-operational stage, 3) concrete operational stage, and 4) formal operational stage (see Buggle, 1993, for review).[3] The central consideration concerning the self-model is that the self-model is just one cognitive phenomenon in the sequential cognitive development of an individual (and her experiences), which can be essentially characterized independently from the social environment. The self-model evolves as a by-product of the cognitive development. Since the whole cognitive development mainly relies on a system-environment-interaction, this also holds true for the development of self-models.

We will show that both lines of arguments are inadequate: Mead and the radical representatives of the social-cognitive theory of self-models are unable to account for the fact that there is a parallel development of individual-cognitive and social-cognitive properties that is relatively independent at the beginning of ontogeny. Piaget and the representatives of the individual-cognitive theory of self-models underestimate the dependency of the specific development of human cognition on social interaction. But the latter aspect is not relevant at the beginning of ontogeny. There is a shift of dominance in the constitutive elements of a self-concept from individual-cognitive to social-cognitive abilities during the first four years of life. Therefore, we will defend a *developmental theory of self-models,* which systematically accounts for both individual-cognitive as well as social-cognitive properties as constitutive elements of a self-model.

---

[3] Let us illustrate the core ideas of Piaget's theory shortly: The sensorimotor stage is separated into six stages, which the child passes by the age of two. During that period, he acquires different forms of perception and motor abilities as well as a goal-directed intelligence. Between the ages of two to seven, the child acquires the ability to speak on the preoperational stage (and thereby the ability to take up a position on objects and events) as well as the ability to pretend-play (symbol-play). Characteristic ways of thinking on this developmental stage are (moral) realism, animism and artificialism (Piaget, 1978; 2005). On the concrete operations stage, the child gains the ability to disassociate from direct experiences by the age of eleven and he becomes able to refer his thinking directly to events and objects, but not to hypotheses and propositions (e.g. the child doesn't understand the law of conservation of mass). From the age of eleven on, the child is able to think hypothetically and counterfactually and can release from the present.

2.1. A criticism on Mead's central arguments

Mead's core argument is twofold: first, he stresses the fact that we as human beings are from the very beginning dependent on living in a social group. Not only humans but also several kinds of animals live in groups, and the chances of survival at the early stages of life are dependent on being part of the group. If we want to account for the specific social and conventional properties of humans, it would be too simple to argue as follows: since the individual-cognitive properties could not have been come into existence without the individuals being part of a group, all mental properties are social-cognitive properties. According to this reasoning, all our cognitive properties would be characterized as social properties. This is a classification beside the point since we share even with animals not living in groups basic visual properties and abilities of spatial orientation which are paradigmatic cases of individual-cognitive properties since they evolve essentially in a system-environment-interaction but need not presuppose a system-system-interaction. The central question about the constitution of a self-model is the following: given the background condition that we are living in groups, it is still an open question which factors are dominant for the constitution of a self-model: either individual-cognitive properties (representations of space, time, causality, quantity) or social-cognitive properties (social learning, communication, theory of mind).

Mead argues that the social-cognitive properties are the essential features constituting a self-model as follows:

> Our contention is that mind can never find expression, and could never have come into existence at all, except in terms of a social environment: that an organized set or pattern of social relations and interactions (especially those of communication by means of gestures functioning as significant symbols and thus creating a universe of discourse) is necessarily presupposed by it and involved in its nature. [… it's our; A.N., A.F.] contention that mind develops and has its being only in and by virtue of the social process of experience and activity […] (Mead, 1934, p.223-224).

Mead's main concern was the role of verbal interaction (his so called "symbolic interactionism"). Mead characterized verbal gestures as "significant symbols" by which intelligence and thinking (as an implicit talk of an individual to himself by means of such gestures) can proceed at all. By the use of reflexive pronouns, a person becomes an object for himself, and the social process he belongs to becomes part of the experiences of the individual. Human intellect evolves from social processes; it is the product of social interaction. In addition to language, pretend-play is an important influencing factor for the development of a concept of personal identity. First of all, the child acquires the ability to pretend-play (where the child plays his own "role" as well as the role of someone else; e.g. he plays a police man and the arrested man successively). Afterwards, the child gains the ability to take part in a game (here, he puts himself not in the position of another person but rather in the positions of all other persons who participate in the game; e.g. playing a football game, he puts himself in the position of a goalie, striker etc.). Mead distinguishes between two perspectives of the self-model which interact and influence one another: "I"

and "me". The "me" is completely determined by the attitudes of other persons about myself (more correctly, which I take them to have). The "me" represents an organization of the society in which the individual has his clear position playing his conventional roles. In contrast to those expectations of the other person of myself, Mead postulates the "I", which stands for the creative and incalculable part of the self-model: "the 'I' is the response of the organism to the attitudes of the others; the 'me' is the organized set of attitudes of others which one himself assumes" (Mead, 1934, p.175). The "I" reacts to the "me" by taking a stance towards the estimated attitudes of other people about oneself. For example, watching how another person falls, one helps this person as best as one can (and is expressing by this behavior the cultivated social attitude of the "me"); but watching the fall was funny as well, and one has to restrain laughter (showing the naïve-social attitude of the "I"). Although Mead is famous for making this distinction, it does not matter much in our context since both aspects of his self-model are essentially constituted by social-cognitive properties, i.e. by a system-system-interaction.

The important aspect of his argument is the claim that a self-model presupposes a symbolic system or at least gestures, which are part of a symbolic communication. And since symbolic communication is a social-cognitive ability, the constitution of the self-model is essentially social. Our main critique concerns the fact that Mead does not offer reasons why self-models should presuppose linguistic competence. We will argue for non-linguistic self-consciousness in the next paragraph. If this is shown, there is no further argument that supports the social-cognitive theory of self-models. On the basis of this critique, we will establish a detailed investigation of the relevance of individual-cognitive and social-cognitive properties for the constitution of the self-model.

2.2. *Self-representations as a basis for self-consciousness are neither essentially involving linguistic representations nor consciousness*

There are two lines of argument, which show that the characteristic immediate self-representation does not involve language competence. According to Bermúdez' (1998) theoretical argument called "the paradox of self-consciousness", the connection of self-consciousness to linguistic competence leads to an unacceptable circularity. To avoid this problem, he argues that we have to posit non-conceptual self-consciousness as the basic form of self-consciousness, which is independent of and prior to linguistic competence. We receive the following circle: (1) in order to analyse our ability to have "I"-thoughts, we have to presuppose and analyse the capacity to form and understand "I"-sentences. (2) In order to analyse the capacity to form and understand "I"-sentences, we have to presuppose and analyse our ability to have "I"-thoughts. If we accept that self-consciousness can be already realized independently from linguistic competence (by denying (1)), then we have to account not only for conceptual forms of self-representation but also for non-conceptual ones. Furthermore, there is an empirical observation, which supports the independence of immediate self-representations from linguistic competence: non-linguistic animals and humans are both able to navigate in their environment. Spatial representation involves at least an implicit representation of the cognitive system to account for the relation to the environment. These egocentric spatial representations are a central aspect of all the abilities

of spatial navigation (Vosgerau, 2007). The same holds true for basic visual representations of objects: any visual representation of an object relies implicitly on the spatial relation between the observer and the objects observed. This is also a case of an implicit egocentric spatial representation. Furthermore, there is a special way of establishing an immediate self-representation in animate beings: it is established by our bodily feelings. The brain develops self-representations on the basis of our actions; e.g. the somatosensory cortex represents the body parts and their interrelations as well as the difference to the environment. Therefore, we have to presuppose immediate self-representations that are independent from linguistic abilities. We call them "non-conceptual immediate self-representations" or "non-conceptual self-representations" for short.

Do self-representations presuppose consciousness? Any instance of self-consciousness involves an immediate self-representation. But does any immediate self-representation lead to a case of self-consciousness? Are there examples of unconscious non-conceptual self-representations? The ability of spatial orientation, which we observe in desert ants (homing behaviour) as well as in some robots, involves an implicit egocentric spatial representation that is causally relevant to realize the behaviour. Since consciousness is not always realized in these cases, non-conceptual self-representations do involve neither linguistic representations nor consciousness.[4] To account for such basic self-representations in a general framework, we need an adequate account of mental representations.

3. Varieties of representation and misrepresentation

Mental representations can be characterized as involving representational vehicles (i.e. brain states in the case of human beings and nonhuman animals[5]), the represented entity and the representational relation which holds between the representational vehicle and the represented entity. We have argued elsewhere (Newen & Vogeley, 2003; Newen & Vosgerau, 2007) that it is very fertile to distinguish five different forms of representation by cognitive capacities according to developmental psychology. Each form of representation has essentially distinctive structural features (Newen & Bartels, 2007). These structural features straightforwardly specify criteria of adequacy for each level of representation that determine misrepresentations in each case. We suggest five levels of representation where the following criteria of adequacy are developed on the basis of our own earlier work (Newen & Vosgerau, 2007; Vosgerau, 2009). We start with two kinds of non-conceptual representations constituted by causal relations or by systematic correlations as criteria of adequacy:

(1a)   Non-conceptual sensory-based representations involve a causal relation between a stimulus and a brain state.

---

[4] Considerations on consciousness are just a minor topic here. In another paper, we argue that content and consciousness are orthogonal to each other (Vosgerau, Schlicht & Newen, 2008).

[5] We are not discussing the cases of robots or animals without brains in this paper, although the whole structure remains open for a use of the theory of representation in these cases, too.

(1b)   Non-conceptual contingency representations are based on the detection of systematic correlations between movements and the sense input of the representing system.

Here, the core idea of systematic contingencies put forward by Noe (2005) is integrated into the framework: systematic changes in the sense input can be "foreseen" and attenuated so that stable representations become possible despite the constantly changing input. In addition, during the human ontogeny, conceptual representations develop when the structure of the representations is reorganized such that the representations include systematic distinctions between objects and properties (or between events and properties).

(2)   Conceptual representations are defined to involve (i) an object-property (or event-property) distinction, (ii) a relative stimulus-independence, and (iii) the embedding of the relevant property-representation (the concept) into a minimal semantic net (Newen & Bartels, 2007). According to this definition, concepts are still independent from natural language competence. A conceptual representation is adequate if the object (or event) is classified – by generalizing and systematizing the property – as part of or constitutive for the correct category.

(3)   Propositional representations are combinations of concepts (which are not related to each other like determinable and determinates) satisfying the generality constraint (Evans, 1982): if a cognitive system has the concepts F and G as well as the object representations a and b, then it must be able to produce systematically all varieties of combinations: Fa, Fb, Ga, Gb. Propositional representations can be activated absolutely independent from any specific stimuli while conceptual representations are only relatively stimulus-independent. A propositional representation has truth- (or satisfaction-) conditions, and it is adequate if those are satisfied by facts.

(4)   Meta-representations are necessarily involved in an explicit representation of a propositional attitude involving a subject, an attitude and a propositional content. The so-called theory of mind capacity presupposes such representations. Theory of Mind (ToM) is defined as the ability to attribute mental states, especially propositional attitudes like desires and beliefs to other human beings. A certain level of linguistic ability was found to be a crucial presupposition for infant's ToM-development (Jenkins & Astington, 1996; Astington & Jenkins, 1999). Representations of attitudes are meta-representations of states of affairs; e.g. if a language-competent intentional system makes the ascription "Mary believes that the chocolate is in the kitchen", then this assertion should be represented as the belief-relation between the propositional content that the chocolate is in the kitchen and the subject Mary having this particular belief. A meta-representation is adequate if it can be used to explain the behavior of a subject relying on the folk-psychological explanation (standard belief-desire-explanation).

(5)   Finally, we distinguish iterative meta-representations, which are involved in an explicit representation of a second-order attitude ascription like "Peter hopes that Mary believes that he will come". An iterative meta-representation is adequate if it can be used to explain thinking about social relations on the basis of folk-psychological explanations like "She believes that Peter desires that p, but I believe

that Peter desires that not p; therefore, I will inform her about Peter's desires to prevent her from bringing the wrong present."

In order to establish these five levels of representation as significant forms of representation, we argue that we can distinguish five different types of misrepresentation. If we look at a cat, which is standing upright on its back feet on a kitchen floor, we develop one of the five different representations depending on our interests and our abilities: if I do not have any conceptual abilities, I still will be able to see the cat (but not as a cat). At the level of non-conceptual sensory-based representations, I just develop a figure-ground distinction of the scene, and thereby the information processing can be inadequate by constructing the false figure-ground relation. On the basis of non-conceptual contingency representations, I generate a detailed three-dimensional structure of the figure, which is inadequate if the cat-figure that I correctly notice is falsely structured in its details. A second type of misrepresentation might happen at the level of conceptual representations, e.g. if I have the figure representation, which is even in its details correctly structured, but I classify the object as a squirrel instead of a cat. The third type of misrepresentation involves an inadequate propositional representation characterizing the situation: I am wrong if I represent my understanding of the visual scene by claiming "The cat is lying on the floor". At the fourth level, we acquire the ability to ascribe propositional attitudes: I can utter "I believe that the cat is standing upright" and "Peter believes that the cat is lying on the floor" since the cat changed its position immediately after Peter left the room. I can of course misrepresent Peter's attitude not knowing that he can still see the cat through a little window. The final level of representation involves second order ascriptions of attitudes: a misrepresentation happens if I falsely interpret the attitude ascriptions of someone else. If I model Anna's beliefs about Peter presupposing that she is sharing the beliefs I have, I can go wrong. This would be the case if Anna – but not I – noticed that Peter still can see the cat. I am wrong by claiming "Anna believes that Peter believes that the cat is lying on the floor". The same scenery can lead to representations at very different levels having their own standards of misrepresentation.

The central presupposition is now that these kinds of representation are used by humans not only to develop representations of the external world and other persons but also about oneself. Therefore, we are able to distinguish five levels of self-representation and the respective forms of self-acquaintance and self-consciousness. Each form of self-consciousness is the basis for a human being to construct an implicit self-model (self-schema) or an explicit self-model (self-image). So, we can now characterize five types of self-models for which we then can investigate to which extent they are constituted by individual-cognitive and/or social-cognitive properties.

4. Varieties of self-models: self-schema and self-images

We describe five different kinds of self-models: first an unconscious self-representation called a "non-conceptual self-schema" and then four kinds of consciously represented "self-images" – conceptual, propositional, meta-representational and iterative meta-representational self-images. Each self-model is constituted by a characteristic type of self-consciousness and the minimally stable content represented about oneself on the basis of

instances of this type of self-consciousness. The stability of the content presupposes memory abilities which increase step by step during ontogeny.

Let us characterize these forms of self-models: As soon as babies have developed functioning sense organs, they are able to represent implicitly their own bodily states and start to distinguish them from states of the external world. This clearly happens when babies start to grasp objects by the end of the third month. On the basis of perception-action-loops, we develop the basic form of consciousness of one's own states that we call *phenomenal self-acquaintance*. It is used to establish a minimal stable representation of one's own body, which constitutes a non-conceptual self-schema. A detailed explication of this unconscious self-schema can be given by Gallagher's theory of body schema and body image: a body schema "involves certain motor capacities, abilities, and habits that both enable and constrain movements and the maintenance of posture" (Gallagher, 2005, p.24). Our unconscious body schema is e.g. responsible for walking through a house with low ceilings without bumping the head.

> A body image consists of a system of perceptions, attitudes, and beliefs pertaining to one's own body. In contrast, a body schema is a system of sensory-motor capacities that function without awareness or the necessity of perceptual monitoring (Gallagher, 2005, p.24).

The body schema is the characteristic example of a first relatively stable self-model. Since it remains completely unconscious, we call it "self-schema". The conscious representation of one's own body relies on the conscious representation of bodily properties and postures. There are two strategies to reach such a conscious representation:

(1) The subject's perceptual experience of her own body does not involve any conceptual representations; therefore, we can call the result a "body percept" (Gallagher, 2005, p.25).

(2) A representation of one's own body on the basis of conceptual classifications constitutes a conceptual body image involving conceptual classifications of the body (body concept). While the body percept is still a case of a non-conceptual self-model, the body concept is a case of a conceptual self-model.

Conceptual self-models are the basic type of self-images since they involve essentially a conscious classification which is attached to oneself. On the basis of such a conceptual self-consciousness, we develop a conceptual self-image. It is constituted as a unity of classifications that involves not only concepts of bodily states and properties but also concepts of mental as well as social phenomena. Although these conceptual classifications are explicit, the reference to oneself is still implicit; i.e. it is constituted by an immediate reference determined by my introspection (while the reference to external objects is constituted by the causal relation to external objects on the basis of sense experiences). Both kinds of reference determination may remain implicit. On the basis of a causally or introspectively determined reference, children learn to attach concepts like "sad", "ball", "bird", "singing", etc. without having an explicit self-concept.

A conceptual self-image is extended into a propositional self-image if it includes not only classifications but furthermore whole propositions about oneself. To represent a proposition about myself in the relevant indexical mode, I need an explicit immediate self-

representation, i.e. a self-concept. We usually learn to express our explicit self-concepts at first by uttering our own proper name and then adequately by using the term "I". On the basis of such a propositional self-consciousness, which essentially involves a self-concept, we develop a propositional self-image. It is constituted by a unity of explicit self-ascriptions of states and properties while they are represented as parts of events or situations, e.g. "I am playing football in the stadium", "I am making a cake together with dad". This is a first basic person-model including characteristic dispositions of persons and social roles. It includes often stable representations of one's own desires, but furthermore, the subjects do not take into account other propositional attitudes.

A meta-representational self-image moreover systematically involves self-ascriptions of propositional attitudes (as instances of meta-representational self-consciousness). The subject must have learned to deal with the so-called false belief task, which is usually successfully managed by four-year-old kids. On this basis, the infant starts to develop an autobiography including a bunch of characteristic beliefs, desires, hopes, fears, etc. This development is essentially correlated with the constitution of the autobiographical memory. Therefore, we can characterize the self-image on this level as a complete person-model about oneself including propositional attitudes.

A final stage in our picture developed here is the iterative meta-representational self-image: between the age of 7 and 9, children acquire the ability to make correct second-order ascriptions like "John believes that Mary hopes that the train is arriving in time" (Wimmer & Perner, 1983). The iterative meta-representational self-image is constituted by second-order self-ascriptions of propositional attitudes, e.g. "Anna believes that I think that Hans is a nice person, but in fact I do not like Hans." This high-level self-image is the presupposition for a distinguished communication about social interactions. Therefore, we can characterize this self-image as an intersubjectively reflected person-model involving second-order propositional attitudes.

In the following overview, we present these five types of representation and the instances of self-consciousness that rely on the respective kinds of self-representation. Furthermore, we give an overview of these self-models, which are constituted as unities of instances of self-consciousness on each level (see table below).

5. The constitution of self-models

5.1. Individual-cognitive versus social-cognitive properties: a characterization

In order to investigate the question to which extent a self-model depends on social interaction, we distinguish two kinds of cognitive properties which can be constitutive for the development of a self-model: a) *social-cognitive properties* on the one hand, which evolve during social interaction (system-system-interaction), and b) *individual-cognitive properties* on the other hand, which emerge in dealing with oneself and inanimate objects (system-environment-interaction). We are not presupposing that there is a sharp boundary between both kinds of properties. Nevertheless, we can offer paradigmatic cases to illustrate the distinction: whereas the abilities of pre-verbal and verbal communication, social learning and the understanding of other humans as mental beings (theory of mind) rank clearly

Levels of self-consciousness and of self-models

Forms of representation and the age of acquisition	Types of self-consciousness (or self-acquaintance)	Types of self-models (self-schema and self-image)
non-conceptual representations (even before birth; birth to 9 months)	phenomenal self-acquaintance	the non-conceptual self-schema as an implicit unity of bodily experiences (body schema, on the one hand, and body image relying only on body percepts, on the other)
conceptual representations (9 to 24 months)	conceptual self-consciousness	the conceptual self-image as a unity of object-property features, e.g. body image relying on body concepts; not only bodily properties, also mental and social properties can be included.
propositional representations (2 to 4 years)	propositional self-consciousness	the propositional self-image as a basic person-model constituted as a unity of states, properties and roles which are represented as parts of complex situations
meta-representations (4 years onwards)	meta-representational self-consciousness	the meta-representational self-image as a complete person-model especially involving propositional attitudes
iterative meta-representations (9 years onwards)	iterative meta-representational self-consciousness	the iterative meta-repres. self-image as a reflected person-model especially involving second-order propositional attitudes

among the social-cognitive properties, the ability to estimate causal and spatiotemporal relations as well as an understanding of quantity belong to the individual-cognitive properties. As these examples already indicate, we can observe both kinds of properties on a pre-reflexive, intuitive level (realized by non-conceptual representations) and on a reflexive, inferential level (realized by conceptual/propositional representations). We are now able to investigate the role of social interaction for the constitution of a self-model by carefully analyzing the relevance of individual-cognitive and social-cognitive properties for the constitution of a self-model on each of the five levels. To benefit from detailed description from the perspectives of cognitive development we presuppose, in line with other *theories of self-consciousness* (Neisser, 1988; Bermudéz, 1998; Metzinger, 2003 etc) that there are several especially important features closely connected with self-consciousness: 1) perspectivity (that can be understood social-cognitively as the belief-perspective or individual-cognitively as the spatiotemporal perspective of an individual), 2) the *unity of experiences* (in the sense of a multimodal integration of sensory information and in the sense of a transtemporal unity of a subject) and 3) the *sense of ownership and agency.*

Therefore we account for central aspects in the development of these features. Furthermore, we try to characterize important changes in the structure and content of self-models.

We aim to establish two claims: first, we show that properties of both kinds are observable at each level. Second, we argue that at the beginning of the development, the self-model is essentially constituted by individual-cognitive properties whereas later on – from the age of two onwards – the self-model gradually becomes more and more dominated by social-cognitive properties. To prove this claim, we now present a detailed description of the central observations concerning this issue in developmental psychology.

5.2. The parallel development of individual-cognitive and social-cognitive properties from a developmental perspective

On the different stages of consciousness in ontogeny, the child acquires various cognitive competencies that can be divided into individual-cognitive abilities and social-cognitive abilities. We now have a closer look at these stages of self-models to establish our first thesis that individual-cognitive as well as social-cognitive properties are constitutive for the development of a self-model at each level. In the next paragraph, we add arguments to prove the more advanced thesis according to which there is a shift of dominance in the constitutive elements of a self-model from individual-cognitive to social-cognitive properties. Let us begin with a detailed presentation of the cognitive development.

The non-conceptual self-schema (involving causal relations): from birth to 3rd month

From birth to 3 months, human babies make sensorimotor experiences which are right from the beginning organized in a unity of experience. During the first weeks of life, they learn to modify their reflexes in order to adapt them to the environment (Piaget, 1970; 2003, Sensorimotor Stage (SS), Level I). The baby starts to "structure" his physical and social environment. At this level, it is not very useful to distinguish individual-cognitive as opposed to social-cognitive abilities since the abilities are primarily organized around nutrition such that both dimensions are inseparable. A rich sensory input from a physical environment supports the cognitive development at this early age. By the fourth month, babies combine single reflexes to a complex behaviour such as grasping an object and putting it into the mouth (Piaget, 1970; 2003, SS, Level II). A well-known social interaction that takes place at this very early age is neonate imitation (Meltzhoff & Moore, 1977), i.e. even newborn babies are able to imitate the facial expressions of another person. A basic implicit self-representation seems to be inborn or at least very early developed to account for neonate imitation. Since it does not make much sense to separate individual-cognitive and social-cognitive abilities at this age, we start to discuss the open question at the next level: to which extent is their self-model constituted by individual-cognitive or social-cognitive abilities?

The non-conceptual self-schema (contingency relations): from 3rd month to 9th month

At this age, we can clearly distinguish both dimensions of properties. Let us start with summarizing some individual-cognitive properties: Baillargeon (1987, 1993) shows that 3-month-olds understand that objects continue to exist even when these objects are not longer visible ("object permanence"). In dealing with inanimate objects, the child also establishes a first representation of spatial relations (especially the spatial organization of a grasping distance versus a non-grasping distance) and reaches a first stage of having a spatial perspective. At the same time, children acquire a first comprehension of time order (Haith et. al., 1993). They have then acquired a simple grasp of space and time. Based on this grasp, 6 to 10-month-olds gain an insight into causal relations (Leslie & Keeble, 1987; Cohen & Oakes, 1993). Furthermore, children already show a clear feeling of their own preferences at this age – and behave in accordance with this feeling in order to evoke pleasant effects; e.g. 4 to 8-month-olds shake the rattle in order to evoke the pleasant noise (Piaget, 1970; 2003, SS, Level III). Thereby, they develop a basic feeling of agency of their own actions. This feeling of agency is connected with a registration of the impact of one's own behaviour on the environment, while the environment includes inanimate objects as well as other persons which are influenced in social interaction.

The feeling of agency is an important presupposition for the social-cognitive properties that develop at this age: together with a first understanding of one's own influence on the partners during a social interaction the babies register that there are other subjects, which are agents as well. 2 to 3-month-olds are the more engaged in social interaction the more the communication partner reacts on their behaviour. This reaction is pleasant for the child and induces a feeling of preference in him so that the baby shows his pleasure by smiling at the mother (social smile) (Murray & Trevarthen, 1985). The baby perceives his mother (and other humans as well) as an active and reactive being and thereby has a first (yet not language-based) impression of "the other"; this impression is however purely perception-based and includes no understanding of others as mental beings and no full-fledged causal understanding (i.e. an understanding of invisible forces that guide behaviour and action patterns). Already at this stage, contact to other human beings elicits a social learning effect since by observing the behaviour of other humans, 6-month-olds acquire new behavioural patterns (Collie & Hayne, 1999).

To summarize: At this ontogenetic stage, the child establishes a self-environment interaction. The most important individual-cognitive abilities are the development of a basic feeling of agency and the registration of "object-permanence". Both features are individual-cognitive. Furthermore, the structure of the implicit self-representation is a dyadic one: the self is only registered in relation to an environment. We will see that on the next ontogenetic stage a triadic structure of the self-model is established (self-other-object).

The conceptual self-image (from 9th month to 2nd year)

In the literature, the so-called "9-month-revolution" is well described. Concerning the individual-cognitive properties, a new level of understanding spatial relations emerges. Benson and Uzgiris (1985) revealed that for 10-month-olds, motional experiences are

important for the development of spatial orientation. The child's spatial orientation is improved so well by the 14th month that the child isn't just aware of his own spatial perspective but rather being also able to recognize the visual perspective of others (Sodian, Thoermer & Metz, 2007). This important competence implies a well-developed grasp of the own spatial perspective, on the one hand, and the ability to distinguish between one's own perspective and that of the other, on the other hand. Social-cognitively, this competence involves a rudimentary understanding of others as mental beings. Another remarkable individual-cognitive ability that develops during this period is the recognition of oneself in a mirror by the 18th month (Lewis & Brooks-Gun, 1979). Furthermore, an 11-month-old child shows a clear comprehension of causal relations by understanding that during a collision, a big object has the greater power than a smaller one to push the other away (Kotevsky & Baillargeon, 1994). Bauer and Fivush (1992; Bauer, 1995) show that the causal understanding facilitates memory and imitation abilities in 1 to 2-year-olds. From the age of one on, children make simple plans (Willats, 1990) that often implicitly presuppose an understanding of causal relations.

Concerning the social-cognitive properties: from the 18th month on, the child does not only imitate the actual behaviour of another person but rather the behaviour apparently intended by the other to execute in order to achieve a particular goal (Meltzhoff, 1995). According to Piaget, the imitation of the behaviour of others delayed in time is the first indication of a stable mental representation (Piaget, 1970; 2003, SS, Level VI). The child has a rudimentary understanding of other humans as mental beings; he understands that the behaviour of other persons is guided by their desires to achieve specific goals. 9 to 12-month-olds are capable to distinguish between humans and inanimate objects (Poulin-Dubois, 1999). The social interaction becomes significantly more important when between 9 and 15 months, children acquire the ability of "joint attention"; i.e. they are able to register that the other person is attentive to the same object oneself is looking at (Adamson & Bakeman, 1991, Gauvain, 2001). This involves a new structure of the self-model, because the self is represented not only in relation to an object in the environment but also in relation to an additional subject. Furthermore, joint attention is a crucial presupposition for language acquisition (Baldwin, 1991) and social referencing (Campos & Stenberg, 1981).

To summarize: this period is essentially determined by a first explicit understanding of spatial perspective, self-recognition in the mirror and the social ability of "joint attention". The latter introduces a new structure and enables to acquire a lot of new information (not only about the other person but also) about oneself as being related to the other. Therefore the social cognition gains in importance for the structure and content of the self-model.

The propositional self-image (from 2nd year to 4th year)

During this period, the child acquires various social-cognitive as well as individual-cognitive properties. First, the individual-cognitive ones: the child acquires a more sophisticated notion of time. From the age of four on, he estimates time-lags between events correctly as long as these are less than 60 days (see below). 3-4-year-olds build scripts that classify the typical course of an events such as a birthday party (Fivush & Hamond, 1990; Nelson & Hudson, 1988). The child also has a more sophisticated notion of space and his

own spatial perspective. From the age of 2 on, the child acquires a so-called "linking-competence"; i.e. he can carry along his own position so that he finds back to the starting point (Müller & Wehmer, 1988; Gallistel 1990). From the age of 3 on, children are able to express their knowledge about the difference between humans and inanimate objects verbally; and this knowledge includes a registration of invisible processes like inheritance and digestion (Wellman & Inagaki, 1997).

2.5-year-olds already have a well-developed understanding of tool-use that implies a comprehension of the causal relation between the features of the tool and the probability of success for applying these tools (Chen & Siegler, 2000) as well as an understanding of physical functions in general. At the same time, children begin to play social games requiring the understanding of social roles in general, such as "mother soothes her baby" (O'Reilly & Bornstein, 1993).

The final remark already produced a switch from individual-cognitive to social-cognitive properties: 3-year-olds understand that experiencing an event reveals beliefs about this particular event – but the mere physical closeness to an observer of that event doesn't (Pillow, 1988). This understanding implies the comprehension that it is me that has to observe or experience an event in order to get beliefs about this particular event. From the 2nd year on, children already begin being engaged in pretend-plays such as pretending a banana to be a telephone (Rakoczy, 2006). Furthermore, they gain a first understanding of regularities in games and heavily insist on acting according to these regularities (Rakoczy et. al., 2008). They also develop an understanding of shared intentions: cooperative behaviour presupposes shared intentions ("we-intentionality", Rakoczy, 2008a, p.101f.). This period ends with the development of the important ability of having a theory of mind, which is the core feature of the next stage: The theory of mind ability is defined as the ability to distinguish my own beliefs from those of someone else. Passing the false belief task provides clear evidence for possessing a theory of mind ability. 2-year-olds develop a basic psychological theory that includes a comprehension of the aims and desires of others (Wellman & Gelman, 1998). The child knows that another person acts according to her desires although these desires might diverge from his own desires (Astington, 1993), yet he is not able to attribute a false belief to another person (Wellman & Wooley, 1990). There are several social learning effects on this stage: the theory of mind development depends on social factors such as family size (Jenkins & Astington, 1996), the number of older siblings (Ruffman et. al., 1998), how well-developed the own language competencies are (Astington & Jenkins, 1999) and how often the child is engaged in pretend-plays (Youngblade & Dunn, 1995). Interestingly, 3-year-olds pass the false belief task in the role of a cheat so that their already developed ability to lie facilitates the comprehension of the false belief of another person (Sullivan & Winner, 1993). The children learn to develop basic person-models including dispositions to act and to desire something. But the person-model does not systematically include further propositional attitudes.

To summarize: the central elements in this period are the development of basic linguistic competences, the understanding of regularities, pretend-play, we-intentionality and a first understanding of the desires of other persons without passing the false belief task. There is certainly also a significant improvement in causal understanding of tool-use. However, the cognitive changes caused by the new social-cognitive abilities are dramatic at this period.

The understanding of regularities includes a new structure of the self-model such that the self is related to a whole group by expecting regularities to hold in this group: we receive a self-group-convention structure. Furthermore, the social learning and theory of mind competences are crucial preconditions for the further development of a complex social cognition. This will be argued for in the next paragraph.

The meta-representational self-image (from 4 years onwards)

The core feature of the four-year-revolution is a social-cognitive property: the development of the theory of mind ability. A subject that passes the false belief task clearly has an explicit representation of other humans as mental beings with their own propositional attitudes. The children learn to develop complete person-models systematically including propositional attitudes of several types (beliefs, desires, hopes, fears etc.). This changes the way of social interaction radically. Furthermore, preschoolers attribute emotions to the Heider and Simmel animations, i.e. animated movements of geometrical figures which seem to be human-like (Berry & Springer, 1993).

There are also important developments of the individual-cognitive dimension: the four-year-old child understands how physical and biological processes work, e.g. that plants are able to heal whereas objects like a scratched chair aren't (Backscheider et. al., 1993). From the age of four on, the child comprehends the time-lag between two events in the past as long as the interval does not exceed the duration of 60 days (Friedman, 1991). Children in school start systematically learning a language, mathematics and all the standard culture-dependent knowledge and abilities (e.g. history). The crucial point is that this knowledge and these abilities are then all trained under a specific social condition, i.e. school and teaching conditions. Although abilities like mathematics can intuitively be characterized as individual-cognitive properties, from the age of six onwards, all these abilities are essentially acquired in a social learning situation. Therefore, they are no longer strictly separable from properties, which are basically acquired in a system-environment-interaction. When systematic teaching becomes part of the cognitive development, the social-cognitive properties become dominant. In this trend, the so-called iterative meta-representational self-image is a further stage in ontogeny that the child passes by the age of nine. Then the self-image includes second-order self-ascriptions of propositional attitudes. Since we are not aiming at a full description of the cognitive development, we stop the analysis here – although it is clear and worthy of mention that the period of puberty consists in a further radical progress of the type of self-images involved.

After we have shown that we can find characteristic new properties of the individual-cognitive as well as of the social-cognitive dimension at each ontogenetic level, we now want to prove the second thesis.

5.3. The shift of dominance in the constitution of a self-model: from individual-cognitive to social-cognitive properties

Our second thesis is that there is a shift of dominance from individual-cognitive to social-cognitive properties that are constitutive for the development of the structure and the content

of a self-model during ontogeny. In order to prove our second thesis concerning the shift of significance, we are going to work out the main lines of cognitive development described above, stress the change of the structure of the self-models and compare the cognitive development of human children with the cognitive development of animals. The main reason to do the latter is the following: if we can work out those properties that distinguish humans and animals in their cognitive development, then these are significantly cultural-cognitive properties that are responsible for the development of the sophisticated human culture. They are a subclass of the social-cognitive properties. On the other hand, those cognitive properties, which we share with animals, can be characterized as non-cultural-cognitive properties, because they do not involve any normative rules. It remains an open question how much of these non-cultural properties are individual-cognitive and how much are social-cognitive properties.

In a simplified view, we can distinguish three important culmination points in the ontogeny discussed here: the 9-month-revolution, the 2-year-revolution and the 4-year-revolution.[6] During the first nine months, babies develop perceptions and goal-directed actions structurally similar to those of a lot of mammals. Since animals like rats, cats, dogs, chimpanzees, etc. share the physiological organization of the visual system and the motor system to a great extent with humans, we have to admit that we share a lot of individual-cognitive properties with animals (including nonsocial animals), which are constituted by basic perceptual abilities. The social-cognitive abilities that we acquire quite early in ontogeny (like neonate imitation and the first understanding of others as active and reactive beings) are a first basic step towards a social constitution of a self, but they do not become dominantly crucial for the cognitive system. The 9-month-revolution is a first significant step towards the shift of dominance of the social-cognitive properties: the children acquire the ability of joint attention and thereby switch from a dyadic interaction with an object to a triadic interaction with another subject and an object: this is the most significant change of the self-model at this stage which includes a new structural relation to other subjects. Furthermore, children at this stage acquire the ability to represent someone else's perspective and start to understand the basic intentions of others. They develop concepts to classify the world, and they are able to pass the "mirror-rouge test" (i.e. looking in a mirror, they remove a rouge-spot on their cheek – what reveals their ability of self-recognition in the mirror). This is clearly a strong enrichment of the representation of the self and of others, which we have to account for in a perception-goal psychology (Baron-Cohen, 1993). Cognitive systems are represented as intending agents with a spatial perspective. Interestingly, we still find these abilities to a great extent in the animal kingdom: chimpanzees are able to recognize the perspective of a conspecific, and they are able to imitate the behaviour of humans (Call & Tomasello, 2008). Even more specifically, they are able to grasp why a human act has been done in a special mode (Call & Tomasello, 2008, p.188), which in turn presupposes that they notice the intention of others. Finally, we have strong evidence that grey parrots as well as chimpanzees are able to form concepts (Newen & Bartels, 2007), and it is well-known that several species are able to recognize themselves

---

[6] According to our view, we can add the 9-year-transformation while the latter is not part of our discussion.

in the mirror, including chimpanzees, elephants and magpie (pica pica) (Prior, Schwarz & Güntürkün, 2008).

Although chimpanzees manage the understanding of intended actions and pass the mirror-rouge test, they do not acquire the central abilities of the 2-year-revolution: the understanding of regularities/norms, pretend play and shared intentionality ("we-intentionality") (Rakoczy 2008, 2008a). These abilities are the presuppositions of cooperative behaviour and communicative interaction. Especially the understanding of normative regularities involves a new structure of the self-model which can be described as a self-group-convention structure. The most important study that has been done in this area is a systematic comparison of 105 children being 2.5 years old with 105 chimpanzees[7] (worked out at the MPI Leipzig in the group of Tomasello). Both groups performed the same tasks, which have been separated into two classes: (1) physical tasks which demand the understanding of space (e.g. spatial memory, object permanence, rotation), of quantities[8] and of causality (including tool-use and an understanding of tool-properties), and (2) social tasks that include social learning, communication and theory of mind. The result of the study reveals that chimpanzees have equally well-developed individual-cognitive abilities like 2.5-year-old human children whereas the human children are significantly better in performing the social tasks (Hermann et al., 2008), especially in the social learning condition. This finding strongly supports the cultural intelligence hypothesis according to which the specific cognitive development of humans is grounded in the development and widespread use of the social abilities mentioned above. From the age of two on, when the specific human development starts, we do not have more intelligence than animals, but we develop these social skills that are crucial to start a unique cognitive progress.

The 4-year-revolution, which essentially includes the theory of mind ability: this involves the explicit representation of attitudes as part of the self-model which is separated from explicit person-models of others. This ability is inexistent in the animal kingdom (not to speak of the 9-year-transformation), at least according to the present systematic investigations. The simplified view suggests that there is a strong boundary between human and animal competences. This is definitely not the case. All these specific human developments seem to have precursors in the animal kingdom. But within the human society, the social skills are developed to such a high degree, and they are used so intensely that this influenced and still influences the whole cognitive development.

---

[7] We ignore the further test dimension, which is reported in the study by including a third test group: 32 orangutans. The interesting result is that sometimes orangutans are slightly worse than humans and chimpanzees even in dealing with physical tasks.

[8] In the description above, we took into account abilities of spatial orientation and of causal understanding. We did not describe the development of an understanding of quantity. But concerning this individual-cognitive ability, there exists an analogue story about successive steps of acquiring a full-blown concept of number. This is worked out in detail by the group of Spelke: She argues that the ability to think and reason about numbers is due to two basic abilities, which we share with a lot of animals, i.e. one "system for representing large, approximate numerical magnitudes and a second system for the precise representation of small numbers of individual objects. These systems account for our basic numerical intuitions, and serve as the foundation of the more sophisticated numerical concepts that are uniquely human" (Feigenson, Dehaene & Spelke, 2004, p.307).

The transformation from the non-conceptual self to the meta-representational self is a process in which social-cognitive properties gradually gain in importance such that from the age of 4 on, the cognitive development is essentially influenced by social-cognitive abilities, which seem to be specifically human. According to Carey, the specific cognitive development of humans mainly relies on two factors: on the development of language as an extremely efficient instrument of representation and on an unfolding of a special learning strategy, which she characterizes as "bootstrapping" (Carey, 2004). The important aspect for her argument is that both characteristic features are essentially social-cognitive properties, which pave the specific way for developing new concepts – that is, the foundation of our complex thinking.

Further support for such a transformation is given by the observation that humans (and nonhuman primates) seem not only to be endowed with general learning abilities but also with a small number of domain-specific core systems of knowledge: according to Kinzler and Spelke, humans have core systems of representations of objects, number, space and action. They speculate whether we also have to assume a core system of understanding social partners (Kinzler & Spelke, 2007). It has been shown that the four established core systems can be found across very different human cultures, and it seems that we still share those with nonhuman primates. There are evidences that we have to presuppose a core system of representing social communication partners in humans, but there is – in accordance with the studies cited above – no evidence so far that such a system exists in nonhuman primates. Hence, the development of human cognition is essentially triggered by social-cognitive abilities which enable humans to exceed the realm of core knowledge and establish systematic knowledge which is going more and more beyond perceptual evidence and that, although it might be counterintuitive at first glance, is nevertheless very successful in application. This advanced system of human knowledge starting with propositional representations is radically dependent on conventions and cultural habits.

Our picture has to be refined and clarified when we take a deeper look at different cultures: how can we account for the fact that Western Societies in their advanced development take individual freedom and idiosyncratic unfolding of a person for being so important? Isn't this observation incompatible with the claim of the shift of significance from individual-cognitive to social-cognitive properties? It is not: rather we have to distinguish between individual-cognitive properties, which emerge during a system-environment-interaction, and individualistic features of a person, which develop during a system-system-interaction. If I develop a self-image including a lot of individualistic features, then this is essentially developed in dealing with and dissociating from the society. Thus, such a self-image is still constituted by social-cognitive properties. Supports for this view are the intercultural studies of Markus and Kitayama (1991, 1998). They argue that we can distinguish between individualist and collectivist cultures since in each culture, we find a respective self-image. People with an "independent" self-image, which we typically find in the United States and Western-Europe, focus on individualistic features representing themselves as being essentially separate from others and emphasizing internal attributes like personal traits, skills, motives, and values. People with an "interdependent" self-image, which we typically find in Asian cultures, represent themselves as part of a group and thereby stress the close connection to other people (family, colleagues, religious or political

groups, etc.).[9] We can account for this important distinction between cultures: it is a distinction presupposing a self-image that is constructed at least on the level of propositional representations that in turn has been argued to mark the level at which the specific human cognitive development starts.

This observation is furthermore supported by systematic investigations of the development of three central cognitive dimensions, namely emotions, agency and intentionality. First, we developed a new and general classification of types of emotions according to which there is a lot of evidence to distinguish them exactly in line with the classification above (namely: pre-emotions, basic emotions and then primary and secondary cognitive emotions while primary cognitive emotions are essentially involving propositional representations (Zinck & Newen, 2008)). Second, concerning the phenomenon of agency (i.e. registering that oneself is the agent of an action), we developed a distinction between the feeling of agency (with underlying non-conceptual representations), the judgment of agency (with underlying propositional representations) and the specific social, normative dimension of responsibility (Synofzik, Vosgerau & Newen, 2008a, 2008b). Furthermore, intentionality is also fruitfully described on the basis of such a distinction of levels of representation: nonconceptual motor intention, joint attention based on object representations, joint intentions (we-intentionality based on understanding propositions), intentions manifested in propositional attitudes (Schlicht, 2008). All three phenomena, emotions, agency as well as intentionality, can be fruitfully described in the framework of several levels of representation, which we use to describe different self-models.

Finally, we can support our claim by important observations concerning autism: It is a well-known fact that autistic people suffer from a severe deficit in theory of mind abilities (a fortiori in all more complex social abilities), but they only have a minor deficit in individual-cognitive abilities. They clearly have a fundamental understanding of the physical environment but lack a fundamental understanding of the social environment with radical consequences for their whole cognitive and social development (Frith, 2003). It has been shown that autistic people develop a self-concept, which is mainly characterized by individual-cognitive features (gender, age, height, etc.) and does involve only very few social-cognitive features (like being a member of a social group or being helpful for others). Contrary to this, most of the self-ascriptions of people without mental disorders are explicitly social (Lee & Hobson, 1998). In the same line, it has been shown that autistic people have an intuitive understanding of physics but lack an intuitive folk psychology (Baron-Cohen et al., 2001). Autistic people are able to manage sabotage (presupposing primarily physical knowledge and performance rather than social thinking) but not to deceive other persons by lying (presupposing some understanding of the mental states of others) (Sodian & Frith, 1992). These observations support the claim that a full-blown self-concept is essentially dependent on social-cognitive abilities, especially including the theory of mind ability.

To summarize: there is evidence from empirical anthropology, animal studies and developmental psychology as well as from studies in autism that in the cognitive development of humans, there is a shift of dominance from individual-cognitive properties

---

[9] This is of course an average observation which does not allow drawing immediate conclusion about an individual living in a specific culture.

to social-cognitive properties concerning the development of a self-model: There is a change in structure and content which supports the claim: The change in structure is a development of the self-model from stage 1 to 4: We characterized the self-model as constructed according to (1) a dyadic self-object-relation, (2) a triadic self-other-object-relation and (3) a self-group-convention relation. (4) With the theory of mind ability the self-model is then constructed according to the difference between explicit self-models and a plurality of person-models involving explicit attitude representations. A radical shift in dominance happens during the transformation from conceptual self-images (step 2) to meta-representational self-images (step 4). There is a significant period around the age of two which seems to mark the main shift in dominance from individual-cognitive to social-cognitive properties. We share a non-conceptual self-schema with several types of nonhuman animals, we also share at least partly a conceptual self-image with nonhuman primates whereas the propositional and meta-representational self-images are typically human due to the increasing relevance of social-cognitive properties from the age of two onwards. The acquisition of a complex self-model (involving a system of self-ascribed propositional attitudes) clearly seems to be essentially linked to having social-cognitive abilities.[10]

6. Advantages of our view

The theory of different self-models evolving during ontogeny is supported by or at least compatible with several other theories: Damasio (1999) distinguishes a proto-self, a core self and an extended (autobiographical) self. The idea of a proto-self takes into account even more basic considerations than we thought of here: in order to survive, any biological system must establish an immune system to separate between those materials that are tolerable for the system and those which are not. The activity of the immune system already signals the existence of a biological self, which supports the idea of a proto-self. Damasio's core self as well as Gallagher's minimal self are essentially matching with our characterization of a non-conceptual self-schema (partly involving the features which we separated as belonging to the conceptual self-image), and Damasio's extended self is identical with our meta-representational self-image. The latter is often characterized as a "narrative self". Our view can account for so-called narrative theories of the self according to which any self-model is constituted by narratives; e.g. Dennett (1991) conceives of the self as a "centre of narrative gravity," i.e. the self-model is the abstract intersection point of the multiple stories about oneself. The meta-representational self-image as introduced above essentially relies on the theory of mind ability. It is constituted by whole stories involving propositional attitude ascriptions which I develop about myself. These narratives constitute my autobiography (including mental dispositions and propositional attitudes). So we can identify the "narrative self" with our meta-representational self-image. Contrary to theories of "narrative self", we are able to characterize a lot of different and more primitive selves,

---

[10] Especially later developments are essentially involving social interaction: In puberty children search for their "cultural identity" by explicitly relating or separating themselves to different kinds of groups in a society.

and our theory is also open for more advanced selves, which we expect to develop in puberty.

To conclude: human self-consciousness has an individual-cognitive and a social-cognitive dimension, which develop parallel to each other at the beginning of ontogeny. It can be shown that we have to distinguish different levels of self-models and that a deeper look at the development reveals a shift of dominance in the constitutive elements of a self-model: while in early life, the non-conceptual self-schemata as well as the conceptual self-images are primarily constituted by individual-cognitive properties, from the propositional self-image on, social-cognitive properties gain in importance. From the age of 4 on (when the child has a meta-representational self-image), the self-model is essentially constituted by social-cognitive properties, which include language-competence and learning strategies that are specifically human and that determine the ongoing social enculturation of our cognition. In the debate about the dimensions of self-consciousness, we argue that neither a pure concentration on the individual-cognitive properties – as paradigmatically put forward by Piaget – nor a pure concentration on the social foundations of our life as Mead preferred is adequate. Especially the early years of life need a systematic consideration of both dimensions to account for the shift of dominance indicating the essential role of social-cognitive abilities for a full-blown self-model. Therefore, we suggest a *developmental theory of self-models*.

7. References

Adamson, L.B. & Bakeman, R. (1991): The development of shared intention during infancy. In: R. Vasta (Ed.): Annals of child development, Vol. 8, 1-41. London: Kingsley.

Astington, J.W. (1993): The child's discovery of the mind. Cambridge, MA: Harvard University Press.

Astington, J.W. & Jenkins, J.M. (1999): A longitudinal study of the relation between language and theory-of-mind development. Developmental Psychology 35, 1311-1320.

Backscheider, A.G., Shatz, M. & Gelman, S.A. (1993): Preschoolers' ability to distinguish living kinds as a function of regrowth. Child Development 64, 1242-1257.

Baillargeon, R. (1987): Object permanence in 3.5 and 4.5-month-old infants. Developmental Psychology 23(5), 655-664.

Baillargeon, R. (1993): The object concept revisited: New directions in the investigations of infants' physical knowledge. In: C.E. Granrud (ed.): Visual perception and cognition in infancy, 265-315. Hillsdale, NJ: Erlbaum.

Baldwin, J. M. (1897): Social and ethical interpretations in mental development. London: University Press.

Baldwin, D.A. (1991): Infants' contribution to the achievement of joint reference. Child Development 62, 875-890.

Baron-Cohen, S. (1993): From attention-goal psychology to belief-desire psychology: The development of a theory of mind, and its dysfunction. In: S. Baron-Cohen, H.

Tager-Flussberg & D.J. Cohen (eds.): Understanding other minds: Perspectives from autism, 59-82. Oxford: Oxford University Press.

Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V. & Lawson, J.(2001): Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. Journal of Developmental and Learning Disorder 5, 47-78.

Benson, J.B. & Uzgiris, I.C. (1985): Effect of self-initiated locomotion on infant search activity. Developmental Psychology 21, 923-931

Bermúdez, J. L. (1998): The Paradox of Self-Consciousness. Cambridge, MA: MIT Press.

Berry, D. S. & Springer, K. (1993): Structure, Motion, and Preschoolers' Perception of Social Causality. Ecological Psychology, 5(4), 273-283.

Buggle, F. (1993): Die Entwicklungspsychologie Jean Piagets. Stuttgart: Kohlhammer.

Call, J. & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. Trends in Cognitive Science, 12, 187-192.

Campos, J.J. & Stenberg, C.R. (1981): Perception, appraisal, and emotion: The onset of social referencing. In: M.E. Lamb & L.R. Sherrod (eds.): Infant social cognition: Empirical and theoretical considerations, 273-374. Hillsdale, NJ: Erlbaum.

Carey, S. (2004): Bootstrapping and the origins of concepts, Daedalus, 59-68.

Chen, Z. & Siegler, R.S. (2000): Across the great divide: bridging the gap between understanding of toddlers' and older children's thinking. Malden: Blackwell.

Cohen, L.B. & Oakes, L.M. (1993): How infants perceive a simple causal event. Developmental Psychology 29, 421-433.

Collie, R. & Hayne, H. (1999): Deferred imitation by 6- and 9-month-old infants: More evidence for declarative memory. Developmental Psychobiology, 35, 83-90.

Damasio, A. R. (1999): The feeling of what happens. New York: Harcourt Brace.

Dennett, D.C. (1991): Consciousness Explained. New York: Little Brown.

Descartes, R. (1641, 1992): Meditationes de Prima Philosophia. Meditationen über die Grundlagen der Philosophie. Philosophische Bibliothek, Band 250a, lateinisch-deutsche Ausgabe. Hamburg: Meiner.

Evans, G. (1982): The Varieties of Reference, hg. v. J. McDowell, Oxford: Oxford Univ. Press.

Feigenson, L., Dehaene, S. & Spelke, E.S. (2004): Core Systems of number, Trends in Cognitive Sciences, 8, 307-314.

Fivush, R. & Hammond, N.R. (1990): Autobiographical memory across the preschool years: Toward reconceptualizing childhood amnesia. In: R. Fivush & J.A. Hudson (eds.): Knowing and remembering in young children, 223-48. Cambridge: Cambridge UniPress.

Friedman, W.J. (1991): The development of children's memory for the time of past events. Child Development 62, 139-155.

Frith, U. (2003): Autism – Explaining the Enigma. Oxford: Blackwell Publishing.

Garz, D. (2006): Sozialpsychologische Entwicklungstheorien. Von Mead, Piaget und Kohlberg bis zur Gegenwart. Wiesbaden: VS Verlag für Sozialwissenschaften.

Gallagher, S. (2005): How the Body Shapes the Mind. Oxford: Clarendon.

Gallistel, C. R. (1990): The organization of learning. Cambridge, MA: MIT Press.

Gauvain, M. (2001): The social context of cognitive development. New York: Guilford Press.

Habermas, J. (1987): The Theory of Communicative Action, 2 Vols. Cambridge, MA: Polity.

Haith, M.M., Wentworth, N. & Canfield, R.L. (1993): The formation of expectations in early infancy. In: C. Rovee-Collier & L.P. Lipsitt (eds.): Advances in infancy research. Norwood, NJ: Ablex.

Herrmann, E., Call, J., Hernández-Lloreda, M.V., Hare, B. & Tomasello, M. (2007): Humans have evolved specialized skills of social cognition: The Cultural Intelligence Hypothesis. Science, 317, 1360-1366.

Hume, D. (1748; 2006): An enquiry concerning human understanding. Ed. by Beauchamp, T.L. Oxford: Oxford University Press.

James, W. (1890): Principles of Psychology. 2 Vol. London: Dover Publications.

Jenkins, J.M. & Astington, J.W. (1996): Cognitive Factory and Family Structure Associated With Theory of Mind Development in Young Children. Developmental Psychology, 1996, 32 (1), 70-78.

Kinzler, K.D. & Spelke, E.S. (2007), Core Systems in Human Cognition. Progress in Brain Research, 164, 257-264.

Kotovsky, L. & Baillargeon, R. (1994): Calibration-based reasoning about collision events in 12-month-old infants. Cognition 51, 107-129.

Lee, A. & Hobson, R.P. (1998): On Developing Self-concepts: A Controlled Study of Children and Adolescents with Autism. Journal of Child Psychology and Psychiatry 39 (8), 1131-1144.

Leslie, A.M. & Keeble, S. (1987): Do six month-old infants perceive causality? Cognition, 25, 265-288.

Lewis, M. & Brooks-Gunn, J. (1979): Social cognition and the acquisition of self. New York: Plenum Press. Markus, H., & Kitayama, S. (1991): Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review 98, 224-253.

Markus, H., & Kitayama, S. (1998): The cultural psychology of personality. Journal of Cross Cultural Psychology 29(1), 63-87.

Mead, G.H. (1934): Mind, Self, and Society. Ed. by Charles W. Morris. Chicago: University of Chicago Press.

Metzoff, A. & Moore, M.K. (1977): Imitation of facial and manual gestures by human neonates. Science 198; 75-78.

Meltzhoff, A. (1995): Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. Developmental Psychology 31, 838-850.

Metzinger, T. (2003): Being no one. Cambridge, MA: MIT Press.

Müller, M. & Wehner, R. (1988): Path integration in desert ants Cataglyphis fortis. Proceedings of the National Academy of Sciences 85, 5287-5290.

Murray, L. & Trevarthen, C. (1985): Emotional regulation of interactions between two-month-olds and their mothers. In: T.M. Field & N.A. Fox (eds.): Social perception in infants, 177-197. Norwood, NJ: Ablex.

Nelson, K.E. & Hudson, J. (1988): Scripts and memory: functional relationships in development. In: F.E. Weinert & M. Perlmutter (eds.): Memory development: Universal changes and individual differences, 147-167. Hillsdale, NJ: Erlbaum.

Newen, A. (2000): Selbst und Selbstbewußtsein aus philosophischer und kognitionswissenschaftlicher Perspektive. In: A. Newen & K. Vogeley (eds.): Selbst und Gehirn, 17-53. Paderborn: mentis.

Newen, A. & Vogeley, K. (2003): Self-Representation: The Neural Signature of Self-Consciousness. Consciousness & Cognition 12, 529-543.

Newen, A. (2003a): Ist eine kognitive Selbstbezugnahme naturalisierbar? in: Haas-Spohn (ed.): Intentionalität zwischen Subjektivität und Weltbezug, Paderborn: mentis, 461-475.

Newen, A. & Bartels, A. (2007): Animal Minds: The Possession of Concepts. Philosophical Psychology, 20(3), 283-308.

Newen, A. & Vosgerau, G. (2007): A Representational Theory of Self-Knowledge. Erkenntnis, 67, 337-353.

Newen, A. & Vogeley, K. (2008): Paradigmen in der Philosophie des Geistes. Die Grenzen der Transzendentalphilosophie und die Zukunft des Naturalismus, in: Spät, P. (Hg): Zur Zukunft der Philosophie des Geistes, 93-124, Paderborn: mentis.

Newen, A. & Schlicht, T. (forthcoming): Understanding Other Minds: Developing the Person Model Theory.

Noe, A. (2005): Action in Perception. Cambridge, MA: MIT Press.

O' Reilly, A.W. & Bornstein, M.H. (1993): Caregiver-child interaction in play. In: M.H. Bornstein & A.W. O'Reilly (eds.): The role of play in the development of thought, 55-66. San Francisco: Jossey-Bass.

Perry, J. (1979): The problem of the essential indexical, Nous 13, 3-21.

Piaget, J. (1970; 2003): Meine Theorie der geistigen Entwicklung. Weinheim: Beltz.

Pillow, B.H. (1988): The development of children's beliefs about the mental world. Merril-Palmer Quarterly 34, 1-32.

Poulin-Dubois, D. (1999): Infants' distinction between animate and inanimate objects: The origins of naïve psychology. In: P. Rochat (ed.): Early social cognition: Understanding others in the first months of life, 257 – 280. Mahwah, NJ: Erlbaum.

Prior, H., Schwarz, A. & Güntürkün, O. (2008): Mirror-Induced Behavior in the Magpie (Pica pica): Evidence of Self-Recognition, PLoS Biology, 6(8): e202.

Rakoczy, H. (2006): Pretend play and the development of collective intentionality. Cognitive Systems Research 7, 113-127.

Rakoczy, H., Warneken, F., Tomasello, M. (2008): The Sources of Normatively: Young children's awareness of the normative structure of games. Developmental Psychology 44(3), 875-881.

Rakoczy, H. (2008a): „Du, Ich, Wir": Zur Entwicklung sozialer Kognition bei Mensch und Tier. In: R.I. Schubotz (ed.): Other Minds. Die Gedanken und Gefühle Anderer, 93-110. Paderborn: mentis.

Ruffman, T., Perner, J., Naito, M., Parkin, L. & Clements, W.A. (1998): Older (but Not Younger) Siblings Facilitate False Belief Understanding. Developmental Psychology, 34(1), 161-174.

Schlicht, T. (2008): Ein Stufenmodell der Intentionalität, in: Spät, P. (Hg): Zur Zukunft der Philosophie des Geistes, 59-91. Paderborn: mentis.

Schlicht, T., Springer, A., Volz, K., Vosgerau, G., Schmidt-Daffy, M., Simon, D. & Zinck, A. (forthcoming): Self as a cultural construct? An argument of levels of self-representations. Philosophical Psychology.

Sodian, B, Thoermer, C. & Metz, U. (2007): Now I see but you don't: 14-months-olds can represent another person's visual perspective. Developmental Science 10, 199-204.

Sodian, B. & Frith, U. (1992): Deception and sabotage in autistic, retarded and normal children. Journal of Child and Psychiatry 33(3), 591-605.

Sullivan, K. & Winner, E. (1993): Three-year-olds' understanding of mental states: The influence of trickery. Journal of Experimental Child Psychology 56, 135-148.

Synofzik, M., Vosgerau, G. & Newen, A. (2008a): I move, therefore I am: A new theoretical framework to investigate agency and ownership, in: Consciousness & Cognition 17, 411-424.

Synofzik, M., Vosgerau, G. & Newen, A.(2008b): Beyond the comparator model: A multifactorial two-step account of agency, Consciousness & Cognition 17, 219-239.

Vosgerau, G. (2007): Conceptuality in Spatial Representation. Philosophical Psychology 20, 349-365.

Vosgerau, G. (2009): Mental Representation and Self-Consciousness. Paderborn: mentis

Vosgerau, G., Schlicht, T. & Newen, A. (2008): Orthogonality of phenomenality and content. American Philosophical Quarterly 45, 329-348.

Walter, S. (2000): Autismus – Erscheinungsbild, Ursachen und Behandlungsmöglichkeiten. Giessen: Köhler.

Wellman, H.M. & Gelman, S. (1998): Knowledge acquistion in foundational domans. In: W. Damon & R.S. Siegler (eds.): Handbook of child psychology, Vol. 2: Cognition, Perception and Language. New York: Wiley.

Wellman, H. M. & Wooley, J. D. (1990): From simple desires to ordinary beliefs: The early development of everyday psychology. Cognition 35, 245-275.

Wellman, H.M. & Inagaki, K. (eds.) (1997): The emergence of core domains of thought: children's reasoning about physical, psychological, and biological phenomena. San Francisco: Jossey-Bass Publishers.

Willats, P. (1990): Development of problem solving strategies in infancy. In: D. Bjorklund (Ed.), Children's strategies: Contemporary views of cognitive development, 23-66, Hillsdale, NJ: Erlbaum.

Wimmer, H. & Perner J. (1983): Beliefs about beliefs. Representation and constraining function of wrong beliefs in young children's understanding of deception. Cognition 13, 103-128.

Youngblade, L.M. & Dunn, J. (1995): Individual differences in young children's pretend play with mother and siblings: Links to relationships and understanding of other people's feelings and beliefs. Child Development 66, 1472-1492.

Zinck, A. & Newen, A (2008): Classifying Emotion: A Developmental Account. Synthese 162(1), 1-25.