Part II
From Intentions to Institutions: Development and Evolution

THE ONTOGENY OF SOCIAL ONTOLOGY: STEPS TO SHARED INTENTIONALITY AND STATUS FUNCTIONS

Hannes Rakoczy and Michael Tomasello

Searle's (1995) book *The Construction of Social Reality* is one of the great works of late 20th-century philosophy. Like all great philosophical works, it provides fresh new insights and perspectives on phenomena that lie right before our eyes every minute of every day. After Searle's work, in our view, no right-minded thinker can seriously maintain a narrowly reductionist view of the nature of human experience. After *The Construction of Social Reality*, in our view, no right-minded thinker need take seriously philosophical theories that ignore the fundamentally social nature of many of the most important aspects of our daily lives.

The basic proposal is that humans live as much in a social and institutional world as they live in a physical one. In a football game, the teams and the score are as real as the grass and the dirt on which the game is played. Searle thus develops an ontological framework for the description of basically everything. We live in one world which at bottom consists of brute, observer-independent facts. At the very bottom of the brute facts are the physical facts, a level up the chemical, the biological, and among the latter the intentional facts.[1] Humans and other animals are intentional beings. That Jack believes Mary loves him, and that she hopes that he loves her are objective intentional facts about them. They are brute in the sense of holding independently of any particular observers. But humans (and according to Searle many other animals) not only are individual intentional beings, but they also enter into irreducibly collective we-intentionality with others. When Jack and Mary go dancing together, each intends not only to dance but that "we dance together". Where collective intentionality is involved, we deal with social facts, and again, social facts are brute, objective facts about groups of people. In addition, there is a specific class of collective intentionality that involves the collective application of constitutive rules and the assignment of some kind of deontic status to actions and objects—and here we leave the realm of brute facts and enter the realm of institutional reality, which is not observer-independent. That a group of people is a football team and that a certain move is offside are facts only within

[1] This prima facie looks like the usual levels picture invoking the notion of supervenience or a similar determination relation. Searle's formulation of this picture, though, is a bit idiosyncratic in that he talks of causation between lower and higher levels (e.g., 1983, 1992, 1995) which is inconsistent with a supervenience reading and most levels conceptions (see, e.g., Kim 1995). We will ignore these details in the following.

S.L. Tsohatzidis, (ed.), Intentional Acts and Institutional Facts, 113–137.

the collectively intentional practice of playing football. Take away the practice, for example, in a culture that does not play football, and there is no team and no offside left (though the grass and the dirt, and the people and their intentionality remain). Human social life, from language to money, marriage, and government, is mainly a matter of such institutional reality.

In developing his theory Searle in many places makes empirical claims. Specifically, he claims that of the main components in the construction of human social reality, social animals such as hyenas and chimpanzees possess the most basic, collective intentionality, and that they lack only the more normatively loaded status functions and constitutive rules—and the deontic powers these create. But we think that Searle has his facts about the animals wrong, and that this has some unfortunate consequences for his analysis of the origins of human social reality. The empirical evidence suggests that other animals, including our nearest primate relatives, do not really have shared or collective intentionality of the human kind—using either Searle's own definitions and concepts explicitly, or reasonable extensions or operationalizations of them.

Obviously, empirical facts do not determine conceptual distinctions, but more the other way around, and so being wrong about the facts is in most cases not lethal to a philosophical analysis. But in this case Searle's promiscuous attributions of collective intentionality belie a lack of appreciation, in our opinion, of the complexities and power of this very rare form of social interaction. In this chapter we attempt to spell out these complexities and this power by comparing the social interactions of human children to those of our nearest primate relatives, chimpanzees, with regard to social interactions involving: (1) collaboration, (2) communication, and (3) social learning. The resulting analysis of the skills and motivations which constitute collective intentionality also enables us to propose some plausible hypotheses about how human children move developmentally from fairly local social interactions involving shared intentionality to participation in full-blown institutional reality in the form of constitutive rules, status functions, and the deontology these create.

1. CHIMPANZEES AND HUMAN CHILDREN

Human children as well as many other animals engage in complex, socially coordinated behavior. Here we compare children to our closest primate relatives, chimpanzees, with an eye on the question of how to distinguish social coordination from collective intentionality[2] proper.

[2] Previously we (Tomasello and Rakoczy 2003) distinguished between 'shared intentionality' and 'collective intentionality', the former referring to here-and-now collaboration with specific others, the latter referring to broader, normatively structured societal interactions with others representing generalized others sensu Mead. This distinction is close to Searle's distinction between social and institutional reality. To avoid terminological confusion, though, in the present chapter we follow Searle in using 'collective intentionality' and 'shared intentionality' interchangeably, in the broad sense of we-intentionality creating social reality, with institutional reality being a specific subclass of social reality.

1.1. Collaboration

Humans share with many species the ability to act intentionally. Piaget (1936) demonstrated that from 8–9 months human infants engage in coordinated means-ends behavior, discriminating between success and failure (failure often leading to the choice of different means). Köhler (1925) and subsequent research has demonstrated that chimpanzees and many other animals act intentionally in this same way. Searle (1994, 2001) has long emphasized the continuity of human and other animal individual intentionality, defending our common sense practice of ascribing intentions to cats, dogs, and Köhler's apes against overly linguistic philosophical theories of what intentionality requires (such as Davidson's), theories that "overintellectualize the mind" (Hurley, 2001). At the same time he has stressed the specifically human ways of thinking that are linguistically and culturally constituted. Searle thus advocates a differentiated approach to the individual intentionality of humans and other animals, attempting to identify both similarities and differences.

When it comes to collective intentionality, however, we think Searle is too quick in stressing commonalities between humans and other species: "Humans share with many species of animals the capacity for collective intentionality" (2005: 6f.); "Animals hunting together, birds cooperating in building a nest, and presumably the so-called social insects, such as ants and bees, manifest collective intentionality" (1999: 121). Differences between humans and other species come only with regard to constitutive rules and status functions. But based on empirical research, we would draw a line already at collective intentionality, specifically, we would draw a line between coordinated social activities and collective intentionality.

By collective intentionality in these animal examples Searle means shared we-intentions of the form "we hunt together", "we build a nest together", etc., where each participant "is acting, and doing what he or she does, only as part of a collective action" (2005: 6). In contrast to reductively individualist accounts, Searle does not attempt to conceptually reduce we-intentions to complex individual intentionality of the form "I intend to X", "You intend to X", "I know that you intend to X and know that I intend to. . .". In fact, we-intentions according to Searle are both conceptually irreducible and biologically primitive.

But what are the criteria for ascribing such we-intentions, particularly in the case of prelinguistic and nonlinguistic creatures? More specifically: what counts as evidence for distinguishing such we-intentional collaboration from mere social coordination? Searle (1990, 1995) gives very illuminating analyses for the case of adult cooperators who can be asked about their individual and collective intentions, analyses that—in contrast to some individualist proposals—allow us to distinguish cooperation proper from complex coordinated social behavior plus mutual knowledge. Unfortunately, however, these analyses are unhelpful in the absence of language.

Bratman (1992) identified a trio of characteristic features of shared collaborative activities: (i) mutual responsiveness (to each other's behavior), (ii) commitment

to the joint activity (including coordinating and meshing sub-plans necessary for joint success), and (iii) commitment to mutual support of each other in the respective roles. Mere social coordination exhibits (i), but falls short of (ii) and (iii). Tomasello et al. (2005) recently adapted this analysis for interpreting preverbal collaboration, attempting to operationalize Bratman's criteria. First, if in collaboration the participants pursue a shared we-intention, then if one participant simply stops doing her part, the other(s) should be displeased with this and attempt to reengage her in the joint project. Second, if in pursuing their shared goal the participants grasp the roles that each participant plays (including themselves), then they should be able over repeated enactments of the joint activity to reciprocate and reverse roles as needed.

The question thus arises whether chimpanzees engage in these kinds of shared collaborative activities in the wild. Chimpanzees join one another in agonistic interactions within the group (so-called coalitions and alliances), and they act together to defend the group from predators and other chimpanzee groups. But in these interactions each individual does basically the same thing; they just do it at the same time and with responsiveness to one another's behavior. The most complex coordinated activity of chimpanzees is group hunting, in which two or more males seem to play different roles in corralling a monkey (Boesch and Boesch 1989). But in a reinterpretation of participant behavior over time in these hunts, some observers have characterized this activity as essentially identical to the group hunting of other social mammals such as lions and wolves (Cheney and Seyfarth 1990; Tomasello and Call 1997). Although it is a complex social activity, as it develops over time each individual simply assesses the state of the chase at each moment and decides what is best for it to do. There is nothing that would be called collaboration in the narrow sense of joint intentions based on coordinated plans. Rather, group hunting plausibly is an instance of complex social coordination only.

Natural observations of human children, in contrast, suggest that during the second year they begin to engage in collaborative games with complementary roles and turn-taking structure, and in collaborative instrumental activities with clearly differentiated roles (Brownell and Carriger 1990; Eckerman and Didow 1996; Ratner and Bruner 1978; Ross and Lollis 1987). Natural observations on their own, however, are hard to interpret and need to be complemented by experimental data.

In older experimental studies with apes (e.g., Crawford 1937; Chalmeau 1994), the most complex behavior that could be extracted was something like two chimpanzees pulling a heavy object in parallel, and during this activity almost no communication among partners was observed (Povinelli and O'Neill 2000). More recently, Warneken et al. (2006) directly compared children (18 and 24 months old) and human-raised chimpanzees on experimental tasks that required collaboration with an experimenter. Some of these tasks had an instrumental problem-solving structure (toys for the children/food for the chimps could be obtained through an apparatus that had to be operated together), others

were ends-in-themselves, collaborative games. During the collaborative acts, the experimenter was programmed to interrupt the shared activity at a predetermined point (by ceasing to act in his role) to see if the subjects would attempt to reengage his participation.

Children engaged proficiently in both kinds of collaborative tasks, and during the interruption periods they communicated appropriately with the experimenter (e.g., pointed to his side of the apparatus) and tried to reengage him in the collaborative activity. Chimpanzees, in contrast, engaged with the experimenter in a coordinated way only in the instrumental tasks where there was a food reward to be gained, but not in the noninstrumental game acts. Furthermore, even in the instrumental tasks, when the partner interrupted his engagement, they never communicated with him or tried to reengage him in the collaborative activity (they rather tried on their own or disengaged).

With specific regard to the grasp of the role structure of joint acts, Carpenter et al. (2005) recently placed 1.5-year-old children in the context of novel collaborative games with an experimenter, where the child first learned to play one role and the experimenter the other complementary one. These children demonstrated their ability for role-reversal imitation by spontaneously acting out the other role when appropriate. Human-raised chimpanzees, in contrast, though they had acquired some imitative abilities through enculturation, revealed little competence for such role-reversal imitation (Tomasello and Carpenter 2005).

These results suggest that even very young children pursue shared we-intentions with others. In the experiments they were committed to continuing the joint activity, as indicated by their reengagement attempts, and they understood the roles each participant played and how they were interconnected, as indicated by their pointing to the other's part and by reversing roles. Chimpanzees' behavior, in contrast, is more plausibly interpreted as social coordination with others without the pursuit of a shared we-intention. These findings thus suggest that humans develop a nascent ability to engage in collaboration as the basic form of collective intentionality in the second year of life, before language acquisition has begun in earnest, and this development seems to be a uniquely human achievement.

1.2. Communication

Human communication, needless to say in a volume on John Searle, is a cooperative enterprise. And obviously humans engage in a very specific form of cooperative communication, namely language, which no other animal does. But though language is special, it is not a primitive. Rather it has its ontogenetic roots in pre-linguistic social and communicative activities, as pragmatist theories of language acquisition have stressed (e.g., Bruner 1983; Tomasello 1999, 2003). These activities themselves, we would now like to argue, are characterized by nascent collective intentionality, and are uniquely human.

It is true that chimpanzees, like many species, engage in varied and complex communicative behavior. Particularly, they gesture to one another regularly.

While some of their gestures are relatively inflexible displays invariably elicited by particular environmental events, an important subset are learned by individuals and used flexibly—such things as 'arm-raise' to elicit play or 'touch-side' to request nursing. We know that such gestures are learned because in many cases only some individuals use them, and indeed several observers have noted the existence of idiosyncratic gestures used by only single individuals (Goodall 1986). And their flexible use has been repeatedly documented in the sense of a single gesture being used for multiple communicative ends, and the same communicative end being served by multiple gestures (Tomasello et al. 1985, 1989). Flexible use is also evident in the fact that apes only use their visually based gestures such as 'arm-raise' when the recipient is already visually oriented toward them—so-called audience effects (Tomasello et al. 1994, 1997; Kaminski et al. 2004).

Though these gestures are flexible, intentional acts, they are all tied to concrete action contexts and center around moving others to behave in certain ways. While human infants use these kinds of gestures as well, from around 1 year they engage in communication which goes beyond such proto-imperative acts: they show objects to others, with no obvious aim apart from looking at them together; they point out things to other people simply to share attention and interest with them[3]; and they point for ignorant others to inform them about states of affairs (proto-assertive or proto-informative, Liszkowski et al. 2006).

Apes, in contrast, do not show objects to conspecifics, nor do they point for others. They do not do these things even though they are capable of gesturing flexibly, as indicated above, and though they can track what other individuals do and do not see. Thus, they follow the gaze direction of conspecifics to relatively distal locations (Tomasello et al. 1998); they follow another's gaze direction around and behind barriers to locate specific targets (Tomasello et al. 1999); and they even seem to know the content of what another sees, as individuals act differently depending on whether a competitor does or does not see a potentially contestable food item (Hare et al. 2000, 2001). So the question arises: If chimpanzees have the ability to gesture flexibly and they also know something about what others do and do not see—and there are certainly occasions in their lives when making someone see something would be useful—why do they not sometimes attempt to direct another's attention to something the other does not see by means of a pointing gesture or something equivalent? Some might object that

[3] For example: Child and another person sit in a room and something interesting happens which both see, e.g., lightning outside. The child points to the event, looking back and forth between the other person and the event. Recent experimental evidence shows that 1-year-olds want the other person not only to look at the event or at them, but to attend to the referent and share their attitude about it—if the other does not do so, infants keep on pointing (Liszkowski et al. 2004). Misleadingly, this form of pointing was originally called 'proto-declarative' in a seminal work by Bates et al. (1975). Following Searle's taxonomy of illocutionary act types, it is not so clear what this kind of pointing is proto- for. It most closely resembles expressives in that emotions are expressed and shared, and there is no obvious direction of fit.

they do do this on occasion in some experimental settings, but this in fact only deepens the mystery. The observation is that captive chimpanzees will often "point" (whole arm with open hand) to food so that humans will give it to them (Leavens and Hopkins 1998) or also, in the case of human-raised apes, to currently inaccessible locations they want access to (Savage-Rumbaugh 1990). This means that apes can, in unnatural circumstances with members of the human species, learn to do something in some ways equivalent to pointing (in one of its functions). And yet there is not a single reliable observation, by any scientist anywhere, of one ape pointing for another.

The reason why chimpanzees do not point and human infants do, we think, is that chimpanzee communication lacks the nascent collective intentionality we find in infant communication. Infants' communication—both production and comprehension—is embedded in frames of joint attention and action ("we are looking at this together", "we are doing that together") which make meaning and interpretation possible. Consider this very simple scenario, often called the object choice task. Tomasello et al. (1997) had one person, called the "hider", hide food or a toy from the subject in one of three distinctive containers. Later, a second person, called the "helper", showed the subject where it was by tilting the appropriate container toward them, so that they could see the prize, just before their attempt to find it. After this warm-up period in which he defined his role, the helper began helping not by showing the food or toy but by giving signs, one of which was pointing (with gaze alternation between subject and bucket as an additional cue to his intentions). The apes as a group were very poor (at chance) in comprehending the meaning of the pointing gesture, even though they were attentive and motivated on virtually every trial. (Itakura, et al. 1999, used a trained chimpanzee conspecific to give a similar cue, but still found negative results.) Human 2-year-old children, in contrast, performed very skillfully in this so-called object choice task. Subsequent studies have shown that apes are also generally unable to use other kinds of communicative cues (see Call and Tomasello 2005, for a review), and that even prelinguistic human infants of 14 months of age can comprehend the meaning of the pointing gesture in this situation (Behne et al. 2005).

It is important to recall that apes are very good at following gaze direction in general (including of humans), and so their struggles in this kind of task do not emanate from an inability to follow the directionality of the pointing/gazing cue. Rather, it seems that they do not understand the meaning of this cue—they do not understand either that the human is directing their attention in this direction intentionally or why she is doing so. As evidence for this interpretation, Hare and Tomasello (2004) compared this pointing gesture to a similar but different cue. Specifically, in the key experimental condition they had the experimenter first establish a competitive relationship with the ape, and then subsequently reach unsuccessfully in the direction of the baited bucket (because the hole through which he reached would not enable her arm to go far enough). In this situation, with an extended arm that resembled in many ways a pointing gesture

(but with thwarted effort and without gaze alternation), apes suddenly became successful. One interpretation is that in this situation apes understood the human's simple individual intention to get into the bucket, and from this inferred the presence of food there (and other research has shown their strong skills for making inferences of this type; Call 2004).

But understanding individual instrumental intentions is not the same thing as understanding communicative intentions. Nor is following gaze the same thing as understanding communicative intentions. In simple behavior reading or gaze following, the individual just gathers information from another individual in whatever way it can—by observing behavior and other happenings in the immediate surroundings and making inferences from them. The object choice task, however, is a cooperative and communicative situation in which the subject must understand the joint project, and the experimenter's communicative intentions, in its context. That is, she must understand that the pointing behavior of the human is located in the course of "our project" (see Wittgenstein's construction workers); it is not just an individual instrumental act, but rather something done "for me" and so is relevant in some way for the search task I am facing. Take away the joint project, and there is no communicative meaning left. For example, when in a control condition of the Behne et al. (2005) study the experimenter retreated from the joint project, performed some individual act (e.g. inspected her wristwatch) and in this course distractedly extended her index finger towards one of the boxes as in the pointing condition, children did not interpret this as a communicative act (i.e., did not choose this box).

It is thus the joint attentional frame, or common communicative ground, emanating from the joint project which gives the pointing gesture its meaning in specific contexts (Clark 1996). Thus, if you encounter me on the street and I simply point to the side of a building, the appropriate response would be "Huh?" But if we both know together that you are searching for your new dentist's office, then the point is immediately meaningful. In the object choice task, human infants seem to establish with the experimenter a joint attentional frame that "what we are doing" is playing a game in which I search for the toy and you help me. Upon seeing the point, then, the infant thus asks herself how it is relevant to the joint game (Sperber and Wilson 1986), which leads her to take the point as a proto-assertion as to where the toy is located.

As sophisticated as chimpanzee communication is in its own right, it remains on the level of individualistic intentionality, lacking the structure of we-intentionality in the form of joint attentional frames that characterizes infant communication from the second year on.

1.3. Social Learning

All primates learn from observing others do things. But the human version is distinctively different. Whereas nonhuman primates mostly focus on changes in the environment created by behaving others—so-called emulation learning—humans

focus also on the behavioral means used to create those changes. For example, if a demonstrator opens a nut in some clever way, chimpanzees learn that the nut opens, whereas humans learn in addition the clever strategy used (see Tomasello 1996, for a review). This may be because humans imitate others not just to learn effective ways of dealing with the environment instrumentally, but also to identify with others, to be like them—the so-called social function of imitation (Carpenter 2006). Finally, as noted above, humans also engage in role-reversal imitation in which they do to the other what the other has done to them, or they understand and reproduce the other's role in a collaborative activity merely by observing. The outcome is that human imitative learning enables a much more powerful form of cultural transmission than the weaker, and less social, social learning skills of nonhuman primates, resulting in a kind of cultural ratchet leading to cumulative culture evolution (Tomasello et al. 1993).

With specific reference to shared intentionality, humans in addition employ their skills of imitation in various kinds of interactive contexts involving we-intentionality. Quite often infant imitation takes place in a collective we-context set up by adults and involving informal adult instruction: adults show children what to do with objects, how "we" act with them, supplying feedback and assistance. Tomasello et al. (1993) called this, from the child's point of view, instructed learning, as the adult directions/instructions are learned and internalized along with some instrumental change in the world. No other species on the planet has reliably been observed to do such things in flexible ways across contexts (Kruger and Tomasello 1996). But infants by no means are just passively placed into such collective frameworks and there only perform their individual imitative acts. Rather, from early on they actively participate in collectively intentional forms of imitation of the form "I do what we do" that transcend the purely individualistic "I do what you did".

One such form involves the imitation of conventional objects, particularly artifacts, with *causal usage functions* à la Searle (1995). Not only do infants learn by imitation how to handle objects in the pursuit of idiosyncratic ends, but through imitation they also learn how "we" use scissors, pencils, and chairs, what they are "for". Preschool children reveal rich knowledge about teleological ascription of functions to artifacts when talking about them (e.g., German and Johnson 2002; Kelemen 1999), but how do we know whether younger children really imitate what "we do" with an artifact rather than just "what you did"? One source of indirect evidence comes from "functional fixedness" (Duncker 1945), the phenomenon that we find it hard to use artifacts in ways deviant from their normal use. Functional fixedness can be interpreted as an index of understanding the conventional usage of artifacts, of the ways "we use them". And recent experimental studies have revealed that children from 2 years after imitating novel acts with artifacts show just this kind of fixedness. Another source of evidence comes from naturalistic observations: children in their second year begin to use teleological normative vocabulary in describing artifacts as "broken" (Kagan 1981).

The normative dimension and the collective intentionality of imitation can be seen even more clearly in the case of early imitation of actions that involve some simple *status function* assignment à la Searle (1995). Of course children from 1 year learn a language, a system of status functions. And 2-year-olds have been observed to engage in seemingly normative behavior in response to other's linguistic 'mistakes' by correcting false statements (Pea 1980, 1982). The case of language is difficult to interpret in this context, however, because it is not clear whether children—before they acquire metalinguistic awareness from age 4–5—have to see an utterance in a dual perspective as sound event X which counts as speech act Y in context C, or whether they rather just see through language, so to speak (a point we will return to below). Clearer cases are games involving objects which children know and which get additional status in the context of a game. Simple rule games fall under this heading. A particularly interesting case is pretend play. Children imitatively learn to engage in pretend play (Rakoczy et al. 2005a), and from their second year engage in simple joint pretense. For example, they do such things as pretend with others that a stone is an apple. In the context of this pretense, the stone counts as an 'apple'. In recent studies we found that children from 2 years not only imitatively learn such game actions and engage in joint games, but indicate an awareness of the normative dimensions of games with status: when others do not respect the status function of the objects in the relevant game context, they protest (Rakoczy et al. in press. Rakoczy, submitted). Furthermore, after this protest, children often go on to supplement the critique with some explanation as to how to do the act correctly. That is, these children not only participate in collective instructed learning as pupils, they themselves have begun their career as instructors.

So, human infants become imitative learners from 1 year of age, their imitative abilities revealing remarkable systematicity. In contrast, we do not have good evidence for comparable imitative capacities in great apes. Imitation itself might be a cradle of cooperation in infants, and its absence in apes partly explanatory for their lack of cooperation proper. From the second year on, human imitation not only applies to individual acts, but involves a dimension of collective intentionality, in the form of instructed learning, and even normativity: Children learn how "we" treat objects and assign them functions, and what normative implications this carries. They even—in simple ways—instruct others to do so. Great apes, in contrast, do not reveal any such collective intentionality in their social learning, either in the form of instructed learning or in the form of normativity.

1.4. Summary

Nascent collective intentionality, involving the corresponding normative aspects, develops in the second year of human ontogeny in the areas of collaboration, communication, and social learning. In contrast, closer empirical investigations of animal social life do not reveal any convincing evidence for

collective we-intentionality in nonhuman species. Searle is too promiscuous in granting collective intentionality widely to many animals. This overattribution, we think, rests on two mistakes: First, Searle's notion of collective intentionality overemphasizes its primitiveness and neglects social-cognitive requirements on the part of the individuals participating in cooperation. Many species, such as ants, do not fulfil these requirements—they do not understand each other as subjects—and so cannot be proper cooperators. Second, as just argued at length, even for apes, who do fulfil these requirements, stricter criteria for ascribing collective intentionality proper are needed.

As to the first point, in some sense we agree with Searle that we-intentionality is conceptually not reducible to simpler forms of individual intentionality.[4] We agree that for collective "we"-intentionality there are no necessary and sufficient conditions in terms of individual "I"-intentionality. However, in his argument against reductive individualistic accounts, Searle might throw out the baby with the bathwater, might make we-intentions too primitive in the sense that he does not sufficiently consider individual cognitive prerequisites that are necessary though not sufficient for collective intentionality. Specifically, we-intentionality plausibly builds on, but is not reducible to individual social cognitive abilities to interpret each other. Only if I have some grasp on your acting intentionally, can I enter into a shared we-intentional action with you (see Tomasello and Rakoczy 2003).

At one point Searle does consider individual prerequisites for collective intentionality quite along the lines we do here: "Collective intentionality presupposes a Background sense of the other as a candidate for cooperative agency; that is, it presupposes a sense of others as more than mere conscious agents, indeed as actual or potential members of a cooperative activity" (1990: 414). But then he goes on to stress that these prerequisites are not cognitive ones, but rather belong to the non-intentional skills of the Background. If that simply means that we do not abstractly believe that others are persons (as we do not abstractly believe that there are objects in the world) Searle's proposal sounds fine. However, isn't Searle's general "sense", at least partly, a disposition to have certain beliefs about specific others in specific situations? For example, that Peter here and now can (consciously) see X, but not Y, that Mary now would like to do Z, etc.—and that in collectively intentional activities with Peter or Mary their perceptual and intentional situation would have to be coordinated with one's own?

While these might be to some degree just terminological matters, there is a more substantial issue. Social insect behavior is an instance of collective intentionality, says Searle, and collective intentionality presupposes a sense of others

[4] When Searle talks about individual and collective intentionality, he mostly talks about intentional attitudes of an individual in both cases, but in the former case the attitudes have the form "I intend", "I desire", etc., whereas in the latter they have an irreducible "we"–form ("We intend to take a walk together".). To avoid confusion, we will follow this terminology here.

as intentional agents and potential cooperators. So social insects understand each other as intentional subjects. But what kind of independent evidence do we have for such a claim—independent, that is, from the social coordination?

None, we think. In fact, whereas human infants develop this kind of understanding from 1 year, almost no other species (apart from great apes to some degree) does. What are the indicators? As intentional action (on the world-to-mind fit side) and perception (on the mind-to-world fit side) are the biologically and ontogenetically most basic forms of intentionality (Searle 1983), the most basic cases of understanding one another as intentional subjects should be the cases of perception and of intentional action. Roughly, infants from 1 year of age begin to manifest a bundle of new social behaviors such as following and directing the attention of others (on the perceptual side) and imitating others' acts (on the side of intentional action) that are triadic in structure (there is a triangle child–other–object/situation in the world) and that are best interpreted as reflecting a nascent understanding of other and self as equivalent perceiving and acting intentional subjects (Carpenter et al. 1998b). While many species engage in dyadic interactions with objects on the one hand, and with conspecifics on the other hand, triadic phenomena are markedly absent from most species except humans and to some degree other primates.

So Searle claims that collective intentionality is widespread in the animal kingdom and, by implication, that many species have a sense of each other as intentional agents. In contrast, we think there is no good evidence for collective intentionality proper in nonhuman species. And this is not surprising for most non-primate species given that we do not have good independent evidence for the requisite social cognitive abilities in these species.

Turning to the second point, we should note that the case of the great apes (and perhaps other primates) is quite different in that they do reveal some basic abilities to interpret each other. In their case we do have independent evidence that they understand each other as perceiving subjects (Hare et al. 2000, 2001) and as intentional agents (Call et al. 2004). Both infants and chimps thus go some way together from being intentional agents to understanding one another as such in simple ways. But whereas human infants move further from there to sharing into collective we-intentionality, we do not have good evidence for such a move in the case of the apes when we look at the phenomena more closely and with stricter criteria (as argued in Section 1.1–1.3.). And here it is plausible to see this as rooted in the fact that only humans develop the requisite "biological primitive sense of the other as a candidate for *shared* intentionality" (Searle 1990: 415; our emphasis).

Avoiding these two mistakes of overattributing collective intentionality not only helps us to appreciate the uniquely human skills of creating social reality, but also helps to isolate the even more sophisticated skills required for understanding constitutive rules and status functions, which lead into full-blown human institutional reality. This seems to happen in human ontogeny one or more years after infants' initial skills of shared intentionality have emerged.

Understanding the transition to these more sophisticated skills would thus provide a fuller and more differentiated account of collective intentionality. It is to that task that we now turn.

2. THE TRANSITION TO INSTITUTIONAL REALITY

Collective intentionality constitutes social facts. Many species, according to Searle, live in a world of social facts. Beyond this, however, institutional reality is constituted by special forms of collective intentionality involving the assignment of status functions and constitutive rules. And institutional reality is the heart of uniquely human forms of life, says Searle:

> The institutional ontology of human civilization, the special way in which human institutional reality differs from the social structures and behavior of other animals, is a matter of status functions imposed according to constitutive rules and procedures. Status functions are the glue that hold human societies together. (2005: 9)

While Searle is surely right about institutional ontology being a uniquely human phenomenon, we think overattributing collective intentionality in its simpler forms to many animals prevents one plausible explanation of why no other species on the planet has money, marriages, or property: namely, because no other species has the requisite social ontology (i.e., collective intentionality) on which to build an institutional ontology.

"X counts as Y in context C" is the formula for assigning status functions through constitutive rules, and therefore the logical basis of institutional life. For example, "This piece of wood counts as a queen in the context of the game of chess". Now, many species are individual intentional beings, capable of perceiving and classifying objects, as pieces of wood for example. So clearly the "X" is not the crucial part for explaining why only humans are institutional beings. In a recent paper (Searle, in press) Searle explores the idea that what is crucial is the "X counts as Y" part. This, so Searle claims, requires the ability to think about one object on two levels at once, and in the paradigmatic cases this in turn requires the ability to symbolically represent the X as being a Y. And nonhuman animals lack these abilities. We agree that no other species counts an object as having a collectively defined status beyond its brute identity, and that this is an important part of the story. But it might be that great apes, and perhaps other species are capable of privately perceiving objects on different levels. For example, apes seem capable of using objects such as replicas as scale models for states of affairs in the world in the way human children do (Kuhlmeier and Boysen 2002). True, this probably only involves appreciating natural meaning (Grice 1957), and falls short of understanding collective assignment of anything. But it can be seen as a simple form of treating objects on different levels at once. Thinking about objects on two levels is surely necessary for participating in status assignment, but not sufficient. So what is missing in the apes?

One possibility is that the "C" term already carries a lot of baggage, that the context C presents the background which has to be in place before the "X counts as Y" can be grounded in the first place. The context in question already has to be one of collectively intentional practices, and apes do not have such practices, and therefore their rudimentary capacities (if there are such) to think about objects at different levels privately (involving natural meaning) cannot be made good for creating status (involving nonnatural meaning). Searle seems to miss this possibility due to his overattribution of collective intentionality to large parts of the animal kingdom.

Taking this possibility seriously, the following picture of the ontogeny from individual intentionality to collective intentionality to institutional reality in humans emerges. Children in their first year and chimps go a long way together: they share simple forms of individual intentionality, and some basic forms of understanding intentionality in others and themselves. From the second year of life, however, the ways depart: infants enter into collective we-intentionality in the form of collaboration, preverbal communication and social learning. Early collaboration is confined to instrumental and simple play actions, early communication to gestures (such as pointing), and early social learning to the imitation of nonsymbolic acts. In other words, there is collective intentionality, there are shared practices, in some cases (imitation of acts with artifacts) even involving the collective assignment of causal usage functions—but there is no institutional structure yet. Context C is in place but still awaits status creation. How does this develop? Searle's (1995) fairy-tale about the historical development of borders probably is a good analogy. Imagine two tribes building a stone wall to separate their territories. In the course of the collaborative project of building it, the wall is assigned the causal usage function of separating the territories in virtue of its causal properties. Now over the years the wall decays, and after a while only some stones remain which could not physically fulfill the function of separating the territories. However, in the course of time the practice of not crossing the border has become so entrenched that the stones, even while not functioning as a wall anymore, have come to count as a line not to be transgressed: they have acquired the status of a border. And, we should like to add, they have acquired this status against the background of the context C, which is constituted by the tribes' collective practice of trying to get along with each other etc.

This is a good analogy for ontogeny in one respect: what today is a collective practice for children, yet without status, tomorrow becomes an institutional affair. In another respect, however, ontogeny is obviously different from history: while the historical story is supposed to show how status can arise in a whole community, children grow into an already existing institutional reality, and the question is how they develop the ability to participate in it. As status arises out of the tribes' collectively intentional practice, their collaboration and communication, so children's participation arises out of their engagement in collective intentionality (though with a central role for scaffolding by adult members of the

community). In contrast, as apes do not engage in collective intentionality proper, there is no context C rich enough to ground the emergence of institutional life.

Institutional reality, as Searle and many others have stressed, is a holistic affair: most status is constituted in a network of other status (think of the interrelated matters of money, exchange, private property, etc.), most constitutive rules are not understandable without understanding many others. But how can children then ever enter into this net? Two entering gates, two cradles for becoming an institutional being, we would like to suggest, are language (obviously), and playing games, particularly games of make-believe.

Pre-linguistic communication, as we have argued above, already is a collectively intentional activity. Based on that activity, infants in their second year acquire language as a public, conventional system of constitutive rules and status functions. Language is thus in one sense the continuation of communication with other means. But on the other hand, the means are so special in that they involve conventional meaning, constitutive rules, and thereby normative dimensions that they constitute more than just a continuation.

So children in their second year acquire the ability to correctly perform simple speech acts, at first often in conjunction with deictic gestures (the gesture typically handling reference, plus a linguistic predication), subsequently more independent from gesturing (Tomasello 2003). But not only do they speak correctly, *act in accordance* with the rules themselves—they show an awareness of *following the rules*. First, as mentioned above, 2-year-olds actively correct others' false statements and false labeling (Pea 1980, 1982). Second, soon after children have mastered some basic word use, they begin to play with language, mislabeling objects or making false statements in an obviously intentional and amused way (Johnson and Mervis 1997)—which can be interpreted as an index of their awareness of the normative structure that is violated.

Language as an activity governed by constitutive rules and the corresponding status functions has a double role, as Searle has stressed: being a system of status functions itself, it is also the logical foundation for most other status functions. That some X counts as a Y usually requires some way of symbolizing it as being a Y, and the default way to do this is using language to declare it a Y (on pain of well-known rule regresses this obviously does not apply, however, to status assignment in language itself).

However, although children do acquire language in the second year of life, and indicate some awareness of its normative structure, it is an interesting question whether they understand the status structure of this activity: whether they understand the dual nature of utterances as sound events that on a second level have the status of speech acts (see Kalish 2005). Probably young children do not understand this dual structure, but rather just implicitly master the linguistic practice until they develop a metalinguistic awareness some years later at 4–5 (e.g., Doherty and Perner 1998). They can use language, but see through it, as we use our eyes to see without seeing them (see Wittgenstein 1961, § 5.633). Language is thus the first status involving activity in which children learn to

participate. But their participation is such that they do not have to appreciate the dual structure of status function assignment, do not have to see something as an X while counting it as a Y.

This is different in the case of games, particularly games of pretend play with objects. Games are among the standard textbook examples of institutional activities ("This piece of wood counts as a queen in chess"). Games of pretend play are a special and developmentally very interesting case: there is a similar dual structure of status assignment, such that, e.g., a wooden block counts as something else in the course of the game. But interestingly, this something else, the "Y", is not, as in the case of rule games, only understandable through the game. Imagine an adult and a child 'cutting', 'peeling', and 'eating' a wooden block. "This wooden block counts as an 'apple' in our pretense game" is the form of the status function assignment (which is practically established in the standard case by saying things like "This is our apple now, okay?", see Walton 1990). This status assignment brings with it the usual normative structure of institutional practices: in the context of the game, the wooden block, counting as an 'apple', has to be treated appropriately, can be peeled, eaten, found disgusting, etc., but cannot be drunk or driven.

If pretend play was confined to simple pretense acts like pretending to sleep or pretending to fly, it might be objected that the same concerns as in the case of language would apply: that children do not necessarily have to see an act of pretending as X (flapping one's arms) which additionally counts as Y ('flying'). Perhaps this concern would hold, but in fact early pretense does involve treating objects ("props" as Walton 1990 calls them) as if they were something else, and the concern does not hold for these cases. Particularly, when there are familiar objects used in joint pretending, having a pretense status of being something else, the child has to, in some sense, be aware of the dual structure—otherwise we would not talk of pretense but of delusions (the child losing a tooth in trying to take a bite out of the wooden block does not count as pretending). In contrast to the case of language, the child thus has to see the object as an X that in the context of the game counts as a 'Y'.

Due to this logical structure of pretend play, we think games of pretense are one plausible cradle for the child to enter more broadly and actively into the world of status functions and institutional facts, more actively than in the case of language which does not require the dual view on Xs as Ys (Rakoczy 2006, 2007).

Children begin to engage in pretend play from 1.5 years, imitatively learning it from others and mainly practicing it in joint games of pretending (Rakoczy et al. 2005a,b). By 2 years, their pretending with others has become systematic: they respect the stipulations of the pretense scenario ("this wooden block is our apple, and this pen is our knife")—the status function assignment, that is—and produce normatively appropriate inferential acts. For example, they pretend to cut the wooden block with the pen, handle the pen 'carefully' because it is 'sharp', etc. (Harris and Kavanaugh 1993; Rakoczy et al. 2004; Rakoczy and Tomasello, 2006). And not only do they act appropriately themselves, but also

indicate an awareness of the normative structure of such practice more actively: when a third person joins the game, but makes a "mistake", i.e., does not respect the pretense status of an object, they protest and criticize her (Rakoczy, submitted). So by the third year of life, children have entered into the basics of this remarkable practice of games of pretending, collectively treat objects they know to be Xs as Ys, follow and respect the implications of the proto-constitutive rules of the game, and normatively criticize deviations from the rules.

In embryonic and isolated form we thus have here the basic structure of institutional reality in the games of 2-year-olds. Of course this is a long way from money, marriage, and universities, but the seeds are there, and so joint pretending quite plausibly can be considered the central cradle for, and the entering gate into institutional life. This proposal, we think, gains plausibility from the following lines of reasoning: First, all status function assignment is symbolic, or linguistic in a broad sense: "The sense in which symbolization in [a] broad linguistic sense is essential to all institutional facts is that the move from X to Y in the formula 'X counts as Y in C' is already a symbolic move" (Searle 1999: 155). This, we have already seen, is why language is in some sense the basis for all institutional reality. And, true, in the case of pretending, the standard way to set the scene is to declare things like "This is an apple now". But in many cases one can do without such linguistic (in the narrow sense) declarations. Rather, the introductory acts of the pretense can set the scene (think of putting on a Batman mask to make clear who is going to be in the game; or in our previous example, one could pretend to take the block from a tree, pretend to peel it, etc.). They are then symbolizing or linguistic acts in the broad sense Searle alludes to. The pretense acts make the required symbolizing move and the object itself is in some sense the symbol. This fits nicely with a remark by Searle: "In the limiting case, we can use the object itself to represent the Y status function" (1999: 155). Pretend play, we think, is just such a limiting case, and probably the ontogenetically primary one.

Second, let us come back to the question how, given that institutional reality is holistically structured, one can ever enter into this net. This, of course, is reminiscent of the question how one can ever learn a language. And the related question how one can ever come to interpret others as intentional beings. Regarding the latter question, we have already seen that Searle's claim that perception on the one side (mind-to-world fit) and intentional action on the other (world-to-mind fit) are the most basic forms of intentionality is shared by young children: they begin, at 1 year, to interpret others and themselves as perceiving and intentional subjects first, and then work their way inwards into the intentional net, so to speak, until they acquire a concept of belief, for example, at age 4 (e.g., Perner 1991). Now, in the case of institutional reality, are there "outer" areas of the net which can supply an entry analogous to observation sentences (or something close in case there are no such things) in the case of language, and to intentional action and perception in the case of understanding intentionality? One interesting possibility here is that some institutional activities are less bound up

in the net than others, and therefore more accessible. Games come to mind as a plausible candidate. In fact, Searle at one point explores such a possibility but then, we think, dismisses it too quickly. After having noted the principle of the holism of institutional reality he says:

> It might seem that games are a counterexample to this general principle because, of course, games are designed precisely with the idea that they should be forms of activity that do not connect with the rest of our lives in a way that social facts generally do. Today's baseball game need have no consequences for tomorrow in the way that today's wars, revolutions, buyings, and sellings are intended precisely to have consequences for tomorrow and into the indefinite future. Nonetheless, even in the case of games, there are systematic dependencies on other forms of social facts. (1991: 343)

True, there are still dependencies. But we think, generally not as intricately as in other areas of institutional life. For ontogenetic purposes, the interesting limiting cases are now those where there is little dependency, or at least those which can be practiced by a child without having to be aware of much dependency, and which therefore supply the possibility to enter into the institutional net. And pretend play, it seems, is just such a case. One aspect of the relative independence of pretense games from other institutional affairs is how the context C is specified. Usually, joint pretense is created on the spot in dyads or bigger groups playing together. That is, the context is paradigmatically defined as "our game here and now". Thus, both regarding the persons involved and temporally, this context is much more restricted, and therefore probably much more manageable as an entry than contexts which refer to whole societies or quantify over types of events, etc.

In sum, joint games of pretending are one important cradle for children's entering into institutional reality, indeed, the key cradle beside language which is of course in some sense the fundamental one. But it is one step closer to the bed of institutional reality than language in that pretending (in contrast to early discourse) requires children to take the double-view of status assignment on objects. And once they have mastered this double-view at the very outer edge of institutional reality, in "nonserious", short-lived, relatively isolated and personally restricted contexts, they can begin to work their way inward the institutional network towards ever more general, serious, and temporarily extended practices.

One possible way into the institutional net is via the development of performative speech acts. These speech acts, when uttered in the right context (such as "Your name is Peter" at a baptism) create a state of affairs just by being uttered (Searle 1969). And many of these states of affairs are institutional facts: "Your name is Peter"; "You are married"; "I hereby declare you guilty", etc. And many institutional facts would be unthinkable without the corresponding performative speech acts (Searle 1995). Now, as we have seen earlier, the typical way to set the scene in joint games of make-believe is to say things like "This is our apple now, and I'm now Adam and you're Eve". The striking structural analogy between such declarations and performative speech acts is by no means a coincidence, we think. Rather, such stage-setting in joint pretending should be seen as proto-performative speech acts. And these lay the basis for participation in the performative (in the

more narrow sense) making of institutional facts in a wider range of domains developing subsequently. What is "I'm Adam now, you're Eve" today in a game, tomorrow becomes "I hereby declare you husband and wife" in earnest.

Pretend and other games are probably the first areas where children participate in status assignment, involving a dual view on objects. But of course this is still some way from mature institutional life. Empirically, research on children's understanding of more "serious" status in such areas as money, private property or the nature of naming has typically found that children do not show much understanding in these areas before the school years (for an overview, see Kalish 2005). In a recent study by Kalish et al. (2000), for example, children's status understanding in pretense ("This is now our apple") was directly compared to their understanding of status in serious collective decisions about naming ("This horse is now called Peter") and property issues ("This is now yours"). While 3-year-old children understood "nonserious" status in the pretense context, only much older children (at 7 years) showed competence regarding serious collective decisions.

Such findings suggest that early participation in status assignment in pretense still lacks important building blocks of institutional life. In particular, though young children see a wooden block as such and as counting as 'apple' in a pretend game in some sense, they probably do not yet have much explicit understanding of the logical structure of this: of the fact that "this is a wooden block" is a brute fact, observer-independently true, whereas "This is an 'apple'" is an observer-dependent fact, true in the context of the game because people act according to it in some way. Grasping these logical structures requires mastery of basic epistemological notions. In particular, what is required is a grasp of the logical peculiarities of much status creation: such acts of creation—typically through performatives—have both a mind-to-world direction of fit (they aim at truth), and a world-to-mind direction of fit (they change the world).

The basic epistemological prerequisites are gradually acquired by children from the preschool years, with a major transition at age 4 when children master the basics of the concept of belief (for a review, see Wellman et al. 2001). Speaking a language and participating in specific forms of perspective-shifting discourse, and in discourse with propositional attitude constructions, has been found to be crucial for the development of the concept of belief and related epistemological notions (e.g., Lohmann and Tomasello 2003; see Astington and Baird 2005 for an overview). So language, itself of course a system of status functions, in a dialectical way, helps lay the foundations—via the development of epistemological notions—for developing more sophisticated forms of status assignment and institutional reality.

In sum, the general picture that emerges is this: Language and playing games, particularly pretense games, are the two central cradles for children's growing into institutional life which emerge in the second year of life, based on children's participation in simpler we-intentionality. Language is primary in the sense that most other status depends on language. But young children when they acquire language probably do not yet view speech acts in a dual way as sound that carries

meaning, but rather they see through language. In games of pretense, in contrast, young children not only participate in status assigning practices, they have to implicitly see the normative status structure of the enterprise, have to take a dual perspective on the objects involved. Being an activity rather isolated from the rest of societal life, being a concrete activity where the context C in which status holds is here and now, and involving proto-performatives ("This is our 'apple' now") structurally analogous to real, status-creating performatives, playing pretense games is an important cradle and foundation for children's entry into institutional reality more generally. Developing language and discourse skills enables the acquisition of central epistemological notions, and armed with such notions children can move on from their nascent skills in the area of games to a more general and explicit understanding of status and institutional reality.

3. CONCLUSION

Searle's program over the last decades is a heroic attempt to supply a general ontological framework for describing everything. Particularly his work in the ontology of intentional, social, institutional, and linguistic reality is philosophy at its best: conceptually enlightening and empirically useful for researchers in the cognitive and social sciences.

While applauding Searle's general approach of nonreductive naturalism, we have focused here on some shortcomings in his theory of collective intentionality.[5] Searle starts from the insight that reductive individualist accounts are fatally flawed in their attempt to find necessary and sufficient conditions for we-intentionality in terms of I-intentionality. But Searle, we think, overreacts by declaring we-intentionality too primitive. In particular, the necessary but not in themselves sufficient preconditions in terms of individual social cognitive abilities are neglected. On Searle's account we-intentionality is so primitive that reasonable criteria and restrictions for its ascription are missing, promiscuous attribution being the consequence. True, much socially coordinated behavior of many species prima facie looks like human cooperation. But on a closer look and with more stringent criteria it turns out we do not have good evidence to consider this behavior more than just coordination. And it turns out that for most species we do not have good evidence that they might even come close to fulfilling the individual social cognitive requirements for being cooperators in the first place.

[5] As empirical researchers we have not touched on the more purely philosophical controversies (though these also have implications for describing the relevant empirical phenomena, of course). For example, controversies regarding Searle's radical internalism (brains in vats can have we-intentions), the problem of "free-floating Y terms" (Smith 2003), the question how sharp the distinctions are between regulative and constitutive rules (how many rules of chess can you change without making it a different game?), and between usage and status functions (chairs could be—causally—used for many purposes, so isn't it like a status that they are for sitting?).

By not appreciating the complexities of collective intentionality proper (as empirical research documents it in early human ontogeny), by promiscuously granting it too widely, Searle neglects important and fundamental differences between the social ontology of humans and the social worlds of other animals. Consequently, the emergence of human institutional reality on top of social reality and the question why this is uniquely human is left mysterious: If apes have collective intentionality and social ontology (as Searle thinks), and if they can treat objects on two levels at once (as recent research might suggest), why don't they get their institutions going? In contrast to Searle, we think there is no collective intentionality proper in other animals, in most species because they do not even have the requisite social cognitive abilities, in apes because they do not make it from "I" and "You" to "We", from a simple understanding of each other to sharing with each other. And so there is no foundation for institutional life in the nonhuman animal kingdom. Human uniqueness begins with collective intentionality and social ontology, not with institutional reality.

True, in the current account we still have plenty of mysteries. But we hope we have located them better in the phenomena of the empirical world. Future empirical research into these phenomena would profit from philosophy supplying more refined analyses of collective intentionality, analyses filling the right middle ground between reductive individualism and just-so-primitivism, analyses that also make room for describing different forms of collective intentionality growing in complexity from child pretense to adult theatre.

Acknowledgements. This work was supported by a "Dilthey Fellowship" of the Volkswagen Foundation and the Fritz Thyssen Foundation to Hannes Rakoczy.

REFERENCES

Astington, J.W. and Baird, J.A. (2005). Introduction: why language matters. In J.W. Astington and J.A. Baird (eds.), *Why Language Matters for Theory of Mind*. Oxford: Oxford University Press.

Bates, E., Camaioni, L., and Volterra, V. (1975). Acquisition of performatives prior to speech. *Merrill-Palmer Quarterly,* 21(3), 205–226.

Behne, T., Carpenter, M., and Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science,* 8(6): 492–499.

Boesch, C. and Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Tai-National-Park. *American Journal of Physical Anthropology,* 78(4): 547–573.

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review,* 101(2): 327–341.

Brownell, C. and Carriger, M.S. (1990). Changes in cooperation and self-other differentiation during the second year. *Child Development,* 61: 1164–1174.

Bruner, J. (1983). *Child's Talk*. Oxford: Oxford University Press.

Call, J. (2004). Inferences about the location of food in the great apes (*Pan paniscus, Pan troglodytes, Gorilla gorilla,* and *Pongo pygmaeus*). *Journal of Comparative Psychology*, 118: 232–241.

Call, J., Hare, B., Carpenter, M., and Tomasello, M. (2004). 'Unwilling' versus 'unable': chimpanzees' understanding of human intentional action. *Developmental Science,* 7(4): 488–498.

Call, J. and Tomasello, M. (2005). What do chimpanzees know about seeing revisited: An explanation of the third kind. In N. Eilan, C. Hoerl, T. McCormack, and J. Roessler (eds.), *Joint Attention: Communication and Other Minds*. Oxford: Clarendon Press.

Carpenter, M. (2006). Instrumental, social, and shared goals and intentions in imitation. In S. J. Rogers and J. Williams (eds.), *Imitation and the Development of the Social Mind: Lessons from Typical Development and Autism*. New York: Guilford.

Carpenter, M., Nagell, K., and Tomasello, M. (1998b). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63(4): 176.

Carpenter, M., Tomasello, M., and Striano, T. (2005). Role reversal imitation and language in typically developing infants and children with autism. *Infancy*, 8(3): 253–278.

Chalmeau, R. (1994). Do chimpanzees cooperate in a learning task? *Primates*, 35: 385–392.

Cheney, D.L. and Seyfarth, R.M. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. Chicago, IL: University of Chicago Press.

Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Crawford, M.P. (1937). The cooperative solving of problems by young chimpanzees. *Comparative Psychology Monographs*, 14: 1–88.

Doherty, M. and Perner, J. (1998). Metalinguistic awareness and theory of mind: Just two words for the same thing? *Cognitive Development*, 13: 279–305.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58(5).

Eckerman, C.O. and Didow, S.M. (1996). Nonverbal imitation and toddlers' mastery of verbal means of achieving coordinated action. *Developmental Psychology*, 32: 141–152.

German, T.P. and Johnson, S.C. (2002). Function and the origins of the design stance. *Journal of Cognition & Development,* 3(3): 279–300.

Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge, MA: Harvard University Press.

Grice, H.P. (1957). Meaning. *The Philosophical Review*, 64: 377–388.

Hare, B., Call, J., Agnetta, B., and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour*, 59(4): 771–785.

Hare, B., Call, J., and Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61(1): 139–151.

Hare, B. and Tomasello, M. (2004). Chimpanzees are more skillful in competitive than cooperative cognitive tasks. *Animal Behaviour*, 68(3): 571–581.

Harris, P. L. and Kavanaugh, R. D. (1993). Young children's understanding of pretense. *Monographs of the Society for Research in Child Development*, 58(1)[231]: v–92.

Hurley, S. (2001). Overintellectualizing the mind. *Philosophy and Phenomenological Research*, 63(2): 423–431.

Itakura, S., Agnetta, B., Hare, B., and Tomasello, M. (1999). Chimpanzee use of human and conspecific social cues to locate hidden food. *Developmental Science*, 2(4): 448–456.

Johnson, K.E. and Mervis, C.B. (1997). First steps in the emergence of verbal humor: a case study. *Infant Behavior & Development,* 20(2): 187–196.

Kagan, J. (1981). *The Second Year: The Emergence of Self-Awareness*. Cambridge, MA: Harvard University Press.

Kalish, C. (2005). Becoming status conscious: children's appreciation of social reality. *Philosophical Explorations*, 8(3): 245–263.

Kalish, C., Weissman, M., and Bernstein, D. (2000). Taking decisions seriously: young children's understanding of conventional truth. *Child Development*, 71(5): 1289–1308.

Kaminski, J., Call, J., and Tomasello, M. (2004). Body orientation and face orientation: two factors controlling apes' begging behavior from humans. *Animal Cognition*, 7(4): 216–223.

Kelemen, D. (1999). The scope of teleological thinking in preschool children. *Cognition,* 70: 241–272.

Kim, J. (1995). Mental causation in Searle's 'biological naturalism'. *Philosophy and Phenomenological Research,* 55: 189–194.

Köhler, W. (1925). *The Mentality of Apes*. New York: Harcourt Brace.

Kruger, A.C. and Tomasello, M. (1996). Cultural learning and learning culture. In D.R. Olson and N. Torrance (eds.), *The Handbook of Education and Human Development. New Models of Learning, Teaching and Schooling*. Oxford: Blackwell, pp. 369–387.

Kuhlmeier, V.A. and Boysen, S.T. (2002). Chimpanzees (*Pan troglodytes*) recognize spatial and object correspondences between a scale model and its referent. *Psychological Science,* 13(1): 60–63.

Leavens, D. and Hopkins, J. (1998). Intentional communication by chimpanzees: a cross-sectional study of the use of referential gestures. *Developmental Psychology,* 34(5): 813–822.

Liszkowski, U., Carpenter, M., Henning, A., Striano, T., and Tomasello, M. (2004). Twelve-month-olds point to share attention and interest. *Developmental Science,* 7(3): 297–307.

Liszkowski, U., Carpenter, M., Striano, T., and Tomasello, M. (2006). Twelve- and 18-month-olds point to provide information for others. *Journal of Cognition and Development,* 7(2): 173–187.

Lohmann, H. and Tomasello, M. (2003). The role of language in the development of false belief understanding: a training study. *Child Development,* 74: 1130–1144.

Pea, R.D. (1980). The development of negation in early child language. In D.R. Olson (ed.), *The Social Foundation of Language and Insight: Essays in Honor of Jerome Bruner*. New York: W.W. Norton.

Pea, R.D. (1982). Origins of verbal logic: spontaneous denials by two- and three-year-olds. *Journal of Child Language,* 9(3): 597–626.

Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Piaget, J. (1936). *The Origins of Intelligence in Children.* New York: W.W. Norton.

Povinelli, D. J. and O'Neill, D. K. (2000). Do chimpanzees use their gestures to instruct each other? In S. Baron-Cohen, F. Tager, H. Cohen, and D.J. Cohen (eds.), *Understanding Other Minds*, 2nd edn. Oxford: Oxford University Press.

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

Rakoczy, H. (2007). Play, games, and the development of collective intentionality. In C. Kalish & M. Sabbagh (eds.), *Conventionality in cognitive development: How children acquire representations in language, thought and action. New Directions in Child and Adolescent Development. No. 115* (pp. 53–67). San Francisco: Jossey-Bass.

Rakoczy, H. (submitted). Young children understand the normative structure of joint pretend games.

Rakoczy, H. and Tomasello, M. (2006). Two-year-olds grasp the intentional structure of pretense acts. *Developmental Science*, 9(6): 558–565.

Rakoczy, H., Tomasello, M., and Striano, T. (2004). Young children know that trying is not pretending—a test of the "behaving-as-if" contrual of children's early concept of "pretense". *Developmental Psychology,* 40(3): 388–399.

Rakoczy, H., Tomasello, M., and Striano, T. (2005a). On tools and toys: How children learn to act on and pretend with 'virigin' objects. *Developmental Science,* 8(1): 57–73.

Rakoczy, H., Tomasello, M., and Striano, T. (2005b). How children turn objects into symbols: A cultural learning account. In L. Namy (ed.), *Symbol Use and Symbol Representation*. New York: Erlbaum.

Rakoczy, H., Warneken, F., and Tomasello, M. (in press). The sources of normativity: Young children's awareness of the normative structure of games. *Developmental Psychology.*

Ratner, N. and Bruner, J. (1978). Games, social exchange and the acquisition of language. *Journal of Child Language,* 5: 391–401.

Ross, H.S. and Lollis, S.P. (1987). Communication within infant social games. *Developmental Psychology,* 23(2): 241–248.

Savage-Rumbaugh, S. (1990). Language as a cause-effect communication system. *Philosophical Psychology,* 3: 55–76.

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge, MA: Cambridge University Press.
Searle, J.R. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
Searle, J.R. (1990). Collective intentions and actions. In P. Cohen, J. Morgan, and M. Pollack (eds.), *Intentions in Communication*. Cambridge, MA: MIT Press.
Searle, J.R. (1991). Intentionalistic explanation in the social sciences. *Philosophy of Social Sciences,* 21(3): 332–344.
Searle, J.R. (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.
Searle, J. (1994). Animal minds. *Midwest Studies in Philosophy,* 14: 206–219.
Searle, J.R. (1995). *The Construction of Social Reality*. New York: Free Press.
Searle, J.R. (1999). *Mind, Language, and Society. Philosophy in the Real World*. London: Weidenfeld & Nicolson.
Searle, J.R. (2001). *Rationality in Action*. Cambridge, MA: MIT Press.
Searle, J.R. (2005). What is an institution. *Journal of Institutional Economics,* 1(1): 1–22.
Searle, J.R. (in press). What is language: some preliminary remarks. In S. L. Tsohatzidis (ed.), *John Searle's Philosophy of Language: Force, Meaning, and Mind*. Cambridge: Cambridge University Press.
Smith, B. (2003). The ontology of social reality. *American Journal of Economics and Sociology,* 62(1): 285–299.
Sperber, D. and Wilson, D. (1986). *Relevance*. Cambride, MA: Harvard University Press.
Tomasello, M. (1996). *Do apes ape?* In C. M. Heyes and B. G. Galef (eds.), *Social Learning and the Roots of Culture*. New York: Academic Press.
Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.
Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, M. and Call, J. (1997). *Primate Cognition*. New York: Oxford University Press.
Tomasello, M., Gust, D., and Frost, G.T. (1989). A longitudinal investigation of gestural communication in young chimpanzees. *Primates*, 30(1): 35–50.
Tomasello, M., Call, J., and Gluckman, A. (1997). Comprehension of novel communicative signs by apes and human children. *Child Development,* 68(6): 1067–1080.
Tomasello, M., Call, J., and Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour,* 55(4): 1063–1069.
Tomasello, M. and Carpenter, M. (2005). The emergence of social cognition in three young chimpanzees. *Monographs of the Society for Research in Child Development,* 70(1): 1–136.
Tomasello, M. and Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language,* 18(2): 121–147.
Tomasello, M., Kruger, A.C., and Ratner, H.H. (1993). Cultural learning. *Behavioral and Brain Sciences,* 16: 495–552.
Tomasello, M., Call, J., Nagell, K., Olguin, R., et al. (1994). The learning and use of gestural signals by young chimpanzees: a trans-generational study. *Primates,* 35(2): 137–154.
Tomasello, M., Hare, B., and Agnetta, B. (1999). Chimpanzees, *Pan troglodytes,* follow gaze direction geometrically. *Animal Behaviour,* 58(4): 769–777.
Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences,* 28: 675–691.
Tomasello, M., George, B., Kruger, A. C., Farrar, M. J., and Evans, A. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution,* 14: 175–186.
Walton, K. L. (1990). *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press.

Warneken, F., Chen, F., and Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development,* 77(3): 640–663.

Wellman, H.M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development,* 72(3): 655–684.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. London: Routledge & Kegan Paul.