Hannes Rakoczy

Kinds of selves: A comparative view on the development of intentionality and self-consciousness

Intentionality and the self are correlative phenomena: All intentionality is someone’s (pertains to some ‘self’, as some philosopher might want to say); and every one (every ‘self’, as the same philosophers might wish to continue) is essentially a bearer of intentional attitudes. Different forms of intentionality should thus constitute different forms of selfhood and self-consciousness.

In this chapter, I will review the ontogeny of different forms and levels of intentionality from the perspective of comparative and developmental psychology and discuss the potential implications of these forms of intentionality for the formation of self-consciousness.

To foreshadow the main arguments: human infants and many other species develop in parallel with regard to simple individual intentionality and even regarding some simple individual intentionality of second order: These forms of intentionality constitute a rudimentary consciousness of oneself as an object among many and even in psychological terms as a subject among others. What is uniquely human, however, is the development of collective (or “we”) intentionality from the second year of life. Such we-intentionality constitutes uniquely human forms of self-consciousness – consciousness of oneself as “one of us”, as a member of a group of rational agents.

Individual intentionality

Intentionality, in the broad philosophical sense of ‘aboutness’, is the mark of the mental (Brentano, 1973; Dennett & Haugeland, 1987; Searle, 1983). To be capable of mentality means to be able to enter into intentional, contentful attitudes towards the world and to be guided by these in reasoning and rational action. Paradigmatic intentional attitudes are believing, perceiving, knowing (that something is the case), desiring something to be the case and intending to do some act. In theoretical reasoning, perceptions and beliefs justify other inferential beliefs (e.g., the belief “that p” and the belief “if p then q” together license the belief “that q”). In practical reasoning, desires rationalize other desires, intentions and acts (e.g., the desire “that p” and the belief “act A brings it about that p” together license the intention to do A).

Simple forms of intentionality develop in parallel ways early in human ontogeny and in many other species. Let me mention just two areas that are highly relevant from a developmental psychological perspective, namely object cognition and planned action. In

---

1 This work was supported by a “Dilthey Fellowship” of the Volkswagen Foundation and the Fritz Thyssen Foundation.

2 The present chapter draws heavily on previous work, in particular on a paper called “Collective Intentionality and the Roots of Human Societal Life” in Learning from animals? (edited by Louise Röskahardy & Eva M. Neumann-Held).
developmental research since Piaget’s (1952) seminal work, thinking about an objective world – in its simplest form: thinking about objects existing as ‘out there’ – and acting intentionally and in planned ways have been stressed as the two major milestones in the transition from purely sensorimotor dealings with the world to intentionality proper. All thinking requires a minimal notion of objectivity: the objects thought about exist independently from the perceiver and enduringly out there in the world. And arguably, all thinking starts from a notion of objects, individuals of certain kinds existing continuously in space and time whether perceived or not (e.g., Strawson, 1959). Regarding human ontogeny, Piaget has described infants’ development from initial undifferentiated sensation without any notion of persisting objects (“out of sight, out of mind”) to what he called “object permanence” – the appreciation that objects continue existing objectively whether perceived or not. In their actions, Piaget and much subsequent research found, infants begin to display object permanence from (at latest) the end of their first year: they begin to search for occluded and hidden objects they previously perceived (with implicit looking time measures, analogous competence can be found even much earlier, see Baillargeon, 1987). Furthermore, infants from around 1 year do not only track objects as chunks of matter continuously existing in space and time; they furthermore individuate objects as objects of certain kinds (e.g., this chair, that table, that rabbit …): Recent studies found that when 1-year-olds see an X going into a box, then see a Y coming out of the box (with no purely spatio-temporal cues to decide how many objects there were, as the two objects were never seen simultaneously) and re-entering the box, and finally find only an X (or only a Y) in the box, they are surprised and continue searching (e.g., van de Walle, Carey & Prevor, 2000; Xu & Carey, 1996; Xu & Baker, 2005). A rich interpretation of these findings is that by one year of age, infants begin to apply our common sense metaphysical framework of objects as enduring substances individuated under sortal (kind) concepts (Xu & Carey, 1996) – and thus share the rudiments of our adult conceptual architecture of objective thought.

And many other animals are on a par with the infants: many primate species, and dogs, for example, reach the highest levels of Piagetian object permanence, levels typically reached by infants in the second year (for an overview, see Tomasello & Call, 1997). Furthermore, recent research suggests that some monkeys and great apes individuate objects qua objects of certain kinds much in the same ways as human 1-year-olds do (Mendes, Rakoczy & Call, 2008; Phillips & Santos, 2007; Santos et al., 2002).

Correlatively with awareness of individuals persisting in space and time even when unperceived, another prerequisite for objective thought is some rudimentary awareness of oneself as an object in space (Strawson, 1959). Again we have very clear criteria in speaking creatures (use of personal pronouns, etc.), but what could count as a pre-/non-verbal indicator of some such rudimentary awareness in non-linguistic animals? The non-linguistic task that came standardly to be used in developmental and comparative psychology is the so-called “mirror rouge” task (Amsterdam, 1972; Gallup, 1970). A mark of rouge is surreptitiously applied to the infant’s/animal’s forehead (infants are distracted, animals often narcotised), and then the subject is placed in front of a mirror. Touching one’s own face to remove the mark is interpreted as an indicator of some rudimentary awareness of oneself as an object in space (the “Me” sensu, James (1890) and Mead (1934)). While younger infants and most other species, including monkeys, consistently fail the task (they treat the mirror
image like a conspecific), infants from around 18 months begin to master this task, and great apes have been shown to succeed (see Tomasello & Call, 1997, for an overview). ¹

The second crucial milestone in the development of intentionality in human ontogeny stressed by Piaget is the emergence of intentional, planned action. While much behaviour may be voluntary right from the start, the first clear instances of intentional instrumental action, that is, actions done purposefully and in a planned way in order to achieve some end held in mind, appears in human ontogeny towards the end of the first year: Infants organize their behaviour in means-ends structures and indicate an awareness of the relations between means and ends. In a classic example, infants remove barriers in order to reach a desired object or pull a cloth towards themselves on which the desires object is placed in order to be able to grasp it. And they will persist until the end is fulfilled, varying their means if necessary (Piaget, 1952; Willats, 1985, 1999). Again, these phenomena are widespread also in the non-human animal kingdom: Many species, notably primates, show instrumental problem-solving of remarkable complexity – Köhler’s apes perhaps being the most famous examples.

In sum, thus, many animals share with us the bare bones of simple individual intentionality. ² Like human infants from around 1, many animals are capable of the most basic form of objective thought: perceiving and cognizing about objects. Great apes (and perhaps some other species) even share with human infants some awareness of themselves as objects among many. And many animals don’t just behave, but perform intentional instrumental acts in planned ways (not to mention the remarkable cognitive abilities recent research has found in many species in such areas as causal reasoning, tools use, memory, simple numerical cognition, etc).

Individual intentionality of second order

Individual intentionality as such thus seems to be common to humans and many other animals. But what about intentionality of second order? Much comparative research in the past three decades has focused on such higher-order intentionality – on the ability to understand others and oneself qua intentional beings (also often called “theory of mind” after Premack’s & Woodruff’s (1978) seminal paper “Does the chimpanzee have a theory of mind?”). After Premack’s and Woodruff’s article, so-called “theory of mind” research became a booming field in developmental and comparative psychology, with joint efforts of philosophers and psychologists to find suitable operationalizations of second-order intentionality (e.g., Bennett, 1978; Dennett, 1978; Harman, 1978; Wimmer & Perner, 1983). What emerged as the agreed upon milestone for full-fledged second-order intentionality was the ability to ascribe to others (and oneself in the past) epistemic subjectivity: to attribute

---

¹ Perhaps dolphins and elephants are capable of mastering the task as well, as some recent (though difficult to interpret) studies suggest (Reiss et al., 2001; Plotnik et al., 2006).

² I am speaking of “simple” intentionality here, as arguably many forms of full-fledged human individual intentionality are essentially dependent on language, a point I will return to later (see also Bermudez, 2003, for a proposed taxonomy of simple non-linguistic intentionality in contrast to linguistically mediated intentionality).
intentional attitudes that represent reality as being a certain way and that aim at truth but potentially fail to do so – paradigmatically, (potentially false) beliefs. Empirically, it turned out that around 4 years of age, in human ontogeny a social-cognitive “revolution” occurs such that children begin to manifest a conglomerate of new behaviours: they ascribe false beliefs to others (and themselves in the past) and explain and predict their actions based thereupon (see Wellman et al., 2001, for a meta-analysis). They distinguish appearances from reality (Flavell et al., 1987), and conflicting perspectives of different viewers on the same situation (Flavell et al., 1981, Perner, 1991). And they begin to intentionally deceive others, i.e., lead them to have false beliefs (Sodian, 1991).

Clearly, second-order intentionality of this kind is crucial to many characteristically human activities and achievements such as reflective thinking, full-fledged communication (according to Gricean analyses), and complex conventional activities (according to David Lewis, 1969). And it seems quite clear and (almost) consensus in the field that no other species, not even chimpanzees, reach these sophisticated levels of second-order intentionality (see, e.g., Call & Tomasello, 1999).

Soon, however, simpler forms of second-order intentionality came into focus: the ability to understand not necessarily full-fledged epistemic subjectivity (in particular, false beliefs), but simpler intentional attitudes such as perception and intention. Recourse to a taxonomy of intentionality by Searle (1983) might help to clarify this issue: Searle, following Anscombe (1957), distinguishes two kinds of intentional attitudes that have close analogues in different kinds of speech acts. First, there are cognitive attitudes with “mind-to-world” direction of fit. Their job, so to speak, is to bring the mind in accordance with the world – they aim at truth (and correspond to assertive speech acts). Beliefs and knowledge are the paradigm cases, but perception falls in this category as well. Second, there are conative or ‘pro’ attitudes (Davidson, 1963) with “world-to-mind” direction of fit, whose job, so to speak, is to bring the world into line with the content of the attitude (and so correspond to directive speech acts). Desires, wishes, hopes, and also intentions are in this category. Now, while beliefs and desires are the paradigm cases on both sides, there are specific attitudes on each side on the ‘periphery’ towards the world that, according to Searle, are the biologically and ontogenetically primary ones: perception (on the mind-to-world side), and intentions (on the world-to-mind side).

What came into focus thus was the following possibility: just as the first intentional attitudes that develop in human infants and other animals are perception (of an objective world) and intentional action, the first form of second-order intentionality to develop should be understanding of perception and action. And with this came into focus the further possibility that the divide between humans and other animals might go, ontogenetically speaking, even deeper: whereas the development of simple individual intentionality runs in parallel in humans and others, humans leave all other animals behind in developing even simple forms of second-order intentionality.

The development of such simple forms of second-order intentionality in human infancy has been described in analogy to the social-cognitive revolution at 4 years as the “9-month-revolution” (Tomasello, 1995, 1999). This cognitive revolution of social understanding manifests itself in several distinct – but cognitively related – behaviours that first emerge from around this time: After having been engaged in ‘dyadic’ behaviours with either persons...
or objects for some time in the first year, children now for the first time begin to engage in ‘triadic’ behaviours that involve a referential triangle between child, other person and outside object/event which is jointly perceived/attended to or acted upon. Thus, infants at this age begin to flexibly and reliably look where adults are looking (gaze following), understand what others do and don’t see (perceptual perspective taking), use adults as social reference points to disambiguate novel events (social referencing), and act on objects in the way they have seen adults act on them (imitative learning) – revealing an understanding of the adults’ attitude/directedness towards the outside events (for an overview, see Carpendale & Lewis, 2006; Carpenter, Nagell & Tomasello, 1998; Tomasello et al., 2005). At this same age, infants also begin to use communicative gestures such as the pointing gesture to direct adult attention and behaviour to outside entities and make proto-comments on them (Liszkowski et al., 2004) – trying to influence the adult’s attitude/directedness towards the outside events (Tomasello, 1995). The fact that all these skills emerge in developmental synchrony and correlated fashion (Carpenter et al., 1998) suggests a common underlying cognitive basis – an emerging understanding of oneself and others as intentional agents.

Comparatively, until quite recently, it was widely believed in the field that even such simple intentionality of second order was a uniquely human phenomenon. New experimental findings, however, show that at least chimpanzees develop quite analogous cognitive abilities: First, a series of studies by Hare et al. (2000, 2001) found that chimps understand something about others’ perception. In a food competition situation, a subordinate and a dominant chimpanzee were placed into separate rooms on opposite sides of a third room. In the crucial conditions, food was placed in the third room such that the subordinate could see two pieces of food hidden while the dominant only saw one (his line of sight to the second one being blocked by a barrier). The basic finding was that the subordinates did indeed take into account what the dominants could and could not see: knowing that the dominants would take all the food they could see, the subordinates went for the food that only they themselves could see much more often than they went for the food that both they and the dominant could see. Several control procedures and conditions (one using a transparent barrier that the subordinate apparently understood did not block the dominant’s visual access to the food) effectively ruled out simpler explanation in terms of mere behaviour-reading.

Second, a study by Call, Hare, Carpenter and Tomasello (2004) suggests that chimpanzees understand something about intentional action. Chimpanzees were presented with a human who had food in his hands and then behaved in different ways, marked as either unwilling or unable to give them the food. There were three conditions in which the experimenter was unwilling in different ways (e.g., just staring at the ape, eating the food, teasing the ape with the food). These conditions were each paired with two unable conditions (e.g., trying to get the food out of a jar, and dropping it accidentally). In each group of matched conditions, the surface topography of the experimenter’s behaviour (body movements and gaze direction) was kept as similar as possible. The main finding was that chimpanzees were more impatient – banged on the cage more, left the area sooner – when the human was being mean (unwilling) than when the human was trying but failing (unable), even though in neither case did they get the food. The chimps in this study behaved in
analogous ways as did human infants in a comparison study from 9 months of age (Behne et al., 2005).

The upshot of these lines of research on higher-order intentionality is thus the following: Complex higher-order intentionality in the form of a full-fledged folk psychology invoking beliefs and related subjective epistemic attitudes clearly seems to be a uniquely human, heavily language-dependent achievement developing from around 4 years of age. Simpler forms of second-order intentionality in contrast – an understanding of others and oneself as intentional and perceiving actors – develop in quite parallel ways in human ontogeny from around 1 year and in at least some other primate species.

Collective intentionality

Against this background, the following possibility recently came into focus: What is at bottom uniquely human and a likely foundation of specifically human forms of life, is not so much individual intentionality, but the ability, developing from the second year in human ontogeny, to enter into collective (or “We”) intentionality (Tomasello & Rakoczy, 2003; Tomasello et al., 2005).

With collective intentionality we deal when two or more subjects share an intentional “we” attitude which is not straightforwardly reducible to individual intentional attitudes. When you and I meet and agree to take a walk together, to use an example from Margaret Gilbert (1990), we form and then pursue the joint We-intention “We walk together”, which is not reducible to the sum of my individual intention “I walk” plus your analogous one. When I pursue my individual intention to walk and you pursue yours, we might end up walking beside each other, but not together. When we pursue our We-intention, in contrast, each individual does walk, of course, but acts as part of a joint action.

As in the case of individual intentionality, different kinds of collective intentional attitudes can be distinguished: collective beliefs, collective desires, etc. The central cases of collective intentionality for the present purposes are the arguably basic ones, namely collective intentions and actions – which constitute the class of cooperative acts and lie at the heart of societal and institutional life.

Clearly, collective intentionality presupposes individual intentionality of first order (for us to walk together, I have to be able to walk intentionally myself) and second order (for us to walk together, I have to have a grasp on your walking intentionally and how I can adapt to it): “The biologically primitive sense of the other person as a candidate for shared intentionality is a necessary condition of all collective behavior” (Searle, 1990, p.415). But clearly, individual intentionality, while being necessary, is not sufficient for collective intentionality. And so the present proposal is that while humans share with other animals simple forms of individual intentionality of first and second order, only humans have the ability to build on these to enter into collective intentionality.

The relation between individual and collective intentionality is a dialectical one: On the one hand, human infants are cognitively equipped to understand each other as persons, as

---

Kinds of selves

potential cooperators. Based on this equipment, they enter into collective intentionality and culture. But on the other hand, once children enter into collective intentionality and culture, acquire conventional practices and above all a language, this in turn shapes and transforms their individual cognitive development by supplying them with new means for thinking, much as Vygotsky and Mead have stressed (Tomasello & Rakoczy, 2003).

Before we turn to the empirical phenomena, some further taxonomic distinctions within the class of collective intentional affairs are relevant. Walking together is an example of a cooperative activity that does not essentially involve the conventional use of objects and any assignment of functions. Though such cooperative activities constitute the most basic form of collective intentionality, their cognitive structure is already quite complex: the individual participants have to understand each other as intentional actors, have to form and pursue a joint intention; and in the course of the joint act they have to be mutually responsive to each others’ intentions and acts, often involving division of labour and complementary roles (e.g., Bratman, 1992). Crucially, even simple joint activities involve a normative dimension of commitment: when we embark on a joint project, each of the participants is committed to contributing her or his part to the pursuit of the common goal, supplying support to the other when needed, etc. (e.g., Bratman, 1992; Gilbert, 1990).

An important sub-class of collective intentionality involves the conventional use of objects and the collective ascription of functions to these objects. Using tools to build something together, or using pieces of wood to play chess together, are examples. Two kinds of functions can be distinguished here, with two corresponding degrees of conventionality: causal usage functions are functions we ascribe to objects when we collectively use them instrumentally, i.e., as tools, and when we design and create objects as tools. The objects fulfil the function partly due to their physical causal makeup – the knife due to its sharpness, the hammer due to its hardness. Such causal usage functions are thus conventional in a weak sense: nothing in itself makes a certain object a tool, but we can assign the function to the object simply by making use of its intrinsic physical makeup for our instrumental purposes.

Status functions, in contrast, are conventional in a stronger sense. They are assigned to objects merely as a matter of collective practice, where the objects cannot fulfil the function due to their intrinsic properties. A slip of paper is money, for example, and a piece of wood is a queen in chess, but one could have decided to pay with wood and play with paper. An

---

4 I will here hardly touch upon the development of language and its relation to collective intentionality – as this would easily go beyond the scope of the present chapter. On the one hand, language as a conventional practice is itself an instance of collective intentionality and thus in some sense secondary to collective intentionality. On the other hand, of course, language is in some sense the fundamental collective activity without which many other collective practices would not be conceivable.

5 Strictly speaking, functions are not only assigned to objects, but to actions as well (and, in fact, actions are logically the primary case – the status of objects is dependent on relevant actions one can do with the objects). Language is the paradigmatic example: Emitting such and such sounds in the right context according to the right rules counts as speaking. But I will here focus on the case of object functions, first because regarding objects the general forms of collective intentionality can best be illustrated. And second, because ontogenetically, it is plausible that children come to understand function assignment to objects before they understand it in the case of actions.
object has a certain *status function* only in virtue of the collective intentional treatment of it as having this *status function* – the *status function* is brought into existence, constituted merely by collective intentionality. “X counts as a Y in context C” is the formula that expresses status function creation: “This piece of metal counts as money in our currency area”, for example, or “This piece of wood counts as a king in chess”.

Collective intentionality with the creation of status functions is what lies at the heart of institutional reality. Status functions create institutional facts (e.g., “This is a queen”, “This is money”, “This is a University”), that is, observer dependent facts that only hold in the eyes of a beholder collective creating them – in contrast to brute facts ‘out there’ (“This is a piece of wood”). Institutional reality as a system of status functions pervades our normal adult social life to the degree that we live as much in an institutional as in a natural world – we go to work or school, earn money to pay our rent, own property, are citizens, husbands or wives, and all day long we utter sounds with semantic status functions (meaning), i.e., speak a language.

Specific normative dimensions are involved in the different forms of collective “we” intentionality. In cooperation, as seen above, we commit ourselves to pursuing the joint action and are therefore responsible for trying our best in this pursuit. The assignment of causal usage functions brings with it the notions of good functioning and malfunctioning and the notions of appropriate and inappropriate uses of tools. Status functions, finally, involve a specific kind of rules, namely constitutive rules. Whereas regulative rules regulate an already existing activity (e.g., rules regarding on which side to drive regulate driving, which already exists before the rule), constitutive rules bring into existence the very activity they apply to (Rawls, 1955; Searle, 1969, 1995). For example, the rule of chess “The king can be moved one field in all directions” does not regulate an activity that already exists, but together with the other rules of chess it *constitutes* the very game. Formally, “X counts as a Y in context C” specifies a constitutive rule: that X is a Y in the relevant context; and that it is a Y in the relevant context, confers normative powers to the objects and carries normative implications (that it ought to be treated as a Y). A piece of wood is a queen in the context of chess; and that means it has the power to move in certain ways, ought to be used accordingly, and ought not to be used as firewood in this context, for example.

In sum, collective intentionality involves two or more subjects who share an irreducible “we” attitude, paradigmatically a “we” intention. Some forms of collective intentionality involve the collective assignment of functions to objects. The strongest of such functions, status functions, are those that get collectively assigned to objects merely by virtue of convention, when objects are collectively treated as having that function (“counting as something”). Constitutive rules underlie status functions, create institutional reality and bring with them normative implications – that the objects ought to be treated according to the rules in the relevant context. With this taxonomy at hand, let us now turn to the development of the different forms of collective intentionality in human ontogeny from the second year on and, from a comparative point of view, to the question how this development contrasts with that of other species.
Kinds of selves

In human ontogeny, simple collective intentionality develops from the second year in the domains of cooperative actions and pre-linguistic communication. Children from one and a half begin to engage in collaborative games with complementary roles and turn-taking structure and in collaborative instrumental activities with clearly differentiated roles (Brownell & Carriger, 1990; Eckerman & Didow, 1996; Warneken, Chen & Tomasello, 2006). In the course of such collaborative acts, they communicate pre-linguistically in appropriate ways (e.g., pointing to the required place for the partner). When the collaboration threatens to break down, they re-engage the partner and assign him his role (again by pointing; Warneken et al., 2006). Children at this age, but not chimpanzees, also seem to have a simple understanding of complementary roles in joint activities, as indicated in their spontaneous role-reversal imitation (children: Carpenter et al., 2005; chimpanzees: Tomasello & Carpenter, 2005).

And communication itself, of course, is a cooperative activity characterized by collective intentionality. Even pre-linguistically, using pointing and other gestures, infants make proto-declarative communicative acts that are not just instrumental for attaining some individual end (like in proto-imperative acts of the form “gimme…”; Rivas, 2005): They point out information, for example, that others need (e.g., about the location of a lost object; Liszkowski et al., 2006). Chimpanzees, in contrast, do not spontaneously point; and the ones who learn to do so in human environments only ever use it proto-imperatively for instrumental purposes (Rivas, 2005; Tomasello et al., 2005). Infants’ rudimentary “sense of the other as candidate for shared intentionality” enables participation in these forms of joint cooperative and communicative activities which in turn function as a foundation and scaffold for the acquisition of language (Bruner, 1983; Tomasello, 2003).

Taken together, these studies thus suggest that during the second year of human ontogeny, children develop a nascent ability to engage in cooperative activities as the basic form of collective intentionality: On a simple level, they form and pursue shared “we” intentions with others, with a rudimentary awareness of the commitments and role structures characteristic of cooperative enterprises. The behaviour of chimpanzees, in contrast, does not necessarily warrant the ascription of collective intentionality proper, but might plausibly be characterized as complex social coordination only.

Collective status assignment and proto-institutional activities

Let us now turn to collective intentionality with the assignment of status functions. This form of collective intentionality lies at the heart of institutional reality without which human
society would be virtually inconceivable. And it is here that the dividing line between human sociality and that of other species can be seen most clearly:

Human beings have a capacity which, as far as I can tell, is not possessed by any other animal species, to assign functions to objects where the objects cannot perform the function in virtue of their physical structure alone, but only in virtue of the collective assignment or acceptance of the object as having a certain status and with that status a function. Obvious examples are money, private property and positions of political leadership. (Searle, 2005, p. 7-8)

Money and political leadership are obvious examples of status functions, but from an ontogenetic point of view, it is equally obvious that young children early in development do not have much interesting grasp on such phenomena. What I would like to suggest as a potential cradle for children’s entry into collective intentionality with status function creation, though, is playing games (see Rakoczy, 2006, 2007, in press-a; Rakoczy & Tomasello, 2007). In fact, adult rule games such as chess are also among the paradigmatic examples for practices involving status functions: “This piece of wood counts as a king in the context of chess”, for example, and “In chess, the king moves one field in any direction”. Of course, 2-year-olds don’t play chess. But what children begin to, is to play simple rule games, and in particular, games of pretence.

From a comparative point of view, pretend play is quite clearly a uniquely human phenomenon. Though there are a few anecdotes of pretence-like behaviour in some human-raised animals (for an overview, see Mitchell, 2002), these are difficult to interpret, and generally, it is quite clear that no other species reliably engages in pretend play as we know it (for excellent reviews of precursors to pretend play in great apes, see Gomez & Martin-Andrade, 2002, 2005). Ontogenetically, children usually start to engage in simple pretend play in their second year.

Let’s take as an example two siblings pretending that their parents’ mobile phones are bananas. Child 1 takes a phone, puts it to her mouth, saying to her brother, “Hm, how delicious this banana is. Want some?” The brother then takes the phone, pretends to peel it and to take a bit, etc. Though this is not an instance of playing an established game with fixed rules, it is an instance of collectively playing a game with the assignment of transient status functions, making up ad hoc constitutive rules on the spot. “This phone counts as a ‘banana’ in our pretence context” is the central status function assignment. As the scenario unfolds, “It counts as peeled now” and then “It counts as eaten up now” enter the scene. These assignments bring with them a normative structure of the joint activity. “X counts as Y in context C” means that in C, X ought to be treated accordingly as a Y. In the siblings’ pretence game: once declared a banana, the phone ought to be treated accordingly in the game. Some pretence acts are inferentially licensed in the game, others are not. Pretending to peel the phone/banana, pretending to eat it or to bake a cake with it are licensed, pretending to drive it or pretending to fax it are not (see Walton, 1990).

Children from 2 years do in fact seem to grasp this normative structure created through joint pretence stipulations – as indicated in their inferentially appropriate responses to others’ pretence acts. When an experimenter pretended to pour tea into a cup, for example, children pretended to drink from the cup. When the experimenter pretended to spill tea on
the table, in contrast, children pretended to clean the table (Harris & Kavanaugh, 1993; Rakoczy & Tomasello, 2006; Rakoczy et al., 2004). And they systematically distinguish such pretence acts from superficially analogous behaviours with different intentional structure: when an experimenter pretended to pour from a (full but closed) container into a cup, they themselves – inferentially appropriately – pretended to drink from the cup. However, when the experimenter made the same pouring movements with the same kind of container, but marked them as frustrated attempt, they – again inferentially appropriately – completed the failed attempt by opening the container and really pouring (Rakoczy & Tomasello, 2006; Rakoczy et al., 2004).

That is, young children respect the inferential structure that comes along with collective intentionality and status function assignment, as indicated in their own actions. They even appreciate the basic context-relativity of status assignment: They understand, for example, that one and the same object can have different fictional status in different contexts and can act inferentially appropriately in each of the contexts and flexibly switch between them (Wyman, Rakoczy & Tomasello, in press). But what do they understand about the normativity that status functions introduce? Are they really following a rule, or are they just acting in accordance with a rule, so to speak? Do they indicate an awareness of the normative structure more directly and explicitly than in their own acts? Would they not only act correctly themselves, but criticise others for incorrect acts? This is crucial as critique, beyond mere surprise, in response to incorrect acts is the hallmark of appreciating normative structure. (Mere surprise is the appropriate response when there are acts deviant from purely statistical regularities.)

In recent studies, we addressed this issue (Rakoczy, in press-b; Wyman, Rakoczy & Tomasello, submitted). 2- and 3-year-olds were engaged in games of pretence with status functions assignment to objects. For example, with a pile of building blocks, one block was pretended to be a piece of soap, all other blocks were pretended to be sandwiches. The child and the adult pretended to wash their hands with the ‘soap’ and to eat the ‘sandwiches’. Then at some point came a third character (a puppet), joined into the pretence (“Oh, may I join your game?”) and in the target condition performed pretence acts which were normatively inappropriate in the light of the status functions of the objects. For example, she pretended to eat the block that was the ‘soap’. In the control condition, the puppet pretended appropriately. 3-year-olds (and to a lesser degree 2-year-olds) frequently protested explicitly against such violations of the constitutive rules of the pretence game (e.g., “No, that’s not a sandwich, that’s our soap.”) in the experimental condition, but were content in the control condition.7

In sum, in joint games of make-believe young children from two actively and knowingly participate in collective intentionality with status function creation – as indicated both in their own competent inferential actions and in their normative responses to other’s mistakes. Similar patterns were found also regarding young children’s playing of simple rule games. In a recent set of studies (Rakoczy, Warneken & Tomasello, 2008), we applied a

---

7 See, e.g. one article with supplementary online material for some video examples of such protest. Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: Young children’s awareness of the normative structure of games. *Developmental Psychology, 44*(3), 875-881, supplementary material: [http://dx.doi.org/10.1037/0012-1649.44.3.875.supp](http://dx.doi.org/10.1037/0012-1649.44.3.875.supp).
similar logic as in the pretence study mentioned above: Children were engaged in a joint
game with a partner, when at some point a third character came, wanted to join the game,
but violated the constitutive game rules. In one study, for example, 2- and 3-year-olds were
shown novel actions with novel objects which were marked as conventional games in the
experimental condition (in the control condition, the exact same actions were shown to the
child, but marked merely as just-so behaviour). In the experimental condition, an
Experimenter (E) showed the child the novel objects, declared, “I’ll show you a game, it’s
called ‘daxing’” (novel verb), and presented the target act (e.g., pushing a wooden block to a
target location with a special tool). She also made accidental mistakes (e.g., moving the
block to the same target location, but without the tool), marked as such (“Oops! That’s not
daxing.”). In the control condition, E performed the same behaviours with the objects, but
they were all marked neutrally (“Look, one can do this, and this”). After the demonstration
and after the child acted with the objects, a puppet announced, “I will dax too”
(experimental condition) / ”My turn” (control condition) and performed some act different
from the target act. The children showed clear verbal protest to the puppet (e.g., “No, that’s
wrong”) frequently in the experimental condition (but rarely in the control condition).

In a second study, young children’s (2 and 3 years old) understanding of the context-
relative normativity of constitutive rules was investigated. “X counts as a Y in context C”
means that in C, X ought to be treated as a Y, but it does not carry any normative
implications how to treat X outside of C. An experimenter and the child played with some
known objects, e.g., building blocks. First, they used them in the normal way (i.e., building a
tower). Then E started the game (“I’ll show you a game, it’s called ‘daxing’”) in which the
object got a status function (e.g., one block was used as a dice); and E and the child played
the game together. Then the puppet came, announcing that she would join the game in the
experimental condition. In the control condition, in contrast, she announced that she would
not join, but rather do something different. Formally speaking, in the experimental condition
the puppet entered into context C (and thus subjected herself to the rule in question),
whereas in the control condition she stepped out of C (and therefore outside the scope of the
rule). Again, children showed protest in response to the puppet’s acts frequently in the
experimental condition (but hardly any in the control condition). That is, children not only
understood the normative structure of the game rules, but also appreciated the context-
relativity of this normative structure.

In embryonic and isolated form we thus have here the basic structure of institutional
reality in the games of 2- and 3-year-olds. Of course this is a long way from money,
marriage and universities, but the seeds are there, and so joint pretending and playing other
games quite plausibly can be considered the central cradle for and the entering gate into
institutional life.8

8 Of course, language is the first instance of collective intentionality involving status functions into which
young children enter in rudimentary form from 1 year on. However, and this is one of the reasons why I
haven’t touched upon language in this context, arguably, young language learners do not have to have any
understanding whatsoever of the logical status of constitutive rules and the creation of status functions.
Probably young children just implicitly master the linguistic practice until they develop a meta-linguistic
awareness some years later at 4-5 (e.g., Doherty & 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). The situation is different in
Developing intentionality and self-consciousness

That the self is culturally or linguistically constituted, as many like to say nowadays, is surely a too strong claim if stated in such unqualified form. There are forms of self-awareness that seem to be widespread in the animal kingdom and not essentially tied to rich sociality.

Human infants and many other species develop rudimentary abilities of individual intentionality in analogous ways: perception of an objective world (in the rudimentary sense of positing objects existing independently of being perceived) and the ability to act purposively and in planned ways in that world. Such simple intentionality already implies some at least implicit self-awareness: all perception is already co-perception of oneself, as Gibsonians have been stressing for a long time (above all Neisser, 1988). But infants, great apes and perhaps other species develop individual intentionality that includes self-awareness going beyond such just implicit forms: infants and apes recognize that they are one among the objects persisting in the world.

Even more than that: Infants and apes go beyond individual intentionality and develop simple forms of intentionality of higher order. Not only do they perceive and act intentionally – they understand that others and they themselves perceive and act. True, this is very far from a full-fledged ‘theory of mind’ that seems quite clearly uniquely human, dependent on language and developing much later. But still it probably constitutes basic forms of awareness of oneself not only as an object among others, but as a subject among others with some psychological life.

There are forms of self-awareness, however, that quite clearly seem to be uniquely human and that in fact seem to be culturally constituted in some sense: the forms of awareness linked directly and indirectly to collective intentionality.

Let me first turn to self-awareness directly linked to collective intentionality. In a seminal paper on collective intentionality, Searle claimed that what makes human cooperation possible cognitively was a “biologically primitive sense of the other person as a candidate for shared intentionality” (1990, p. 415, my italics). The research reviewed here suggests that such a sense develops in humans and only humans in the second year and in fact lays the basis for uniquely human social life. True, other primates do develop a simple sense of others as perceiving and acting agents, as we have seen. But this sense seems to remain confined to a sense of the other as an individual agent, as a potential competitor perhaps. The “Machiavellian Intelligence Hypothesis” (Byrne and Whiten, 1988) with its emphasis on the social cognition underlying competition and individual manipulation of conspecifics (i.e., a form of individual second-order intentionality) might thus be well taken for non-human primates. It falls short, however, of adequately describing the inherently
collective dimension of human social cognition. Already human infants from the second year begin to see others as “candidates for shared intentionality”.

Now, this development of consciousness of others has an obvious flipside regarding the development of self-consciousness: a sense of oneself as a candidate for shared intentionality with others – a sense of oneself as “one of us”.

Again, Gibsonians have stressed that most social interaction of even primitive kinds implies some implicit sense of oneself as behaving socially (Neisser’s (1988) “interpersonal self”). And such implicit awareness of oneself as interacting seems widespread in the animal kingdom. But such an implicit awareness that there is contingent interaction going on between oneself and conspecifics in the sense of the “interpersonal self” in Neisser’s taxonomy is a far cry from the awareness of oneself as “one of us”, as member of a group of intentional, potentially cooperating rational agents.

The latter essentially includes awareness of the equivalence between oneself and other potential cooperators: Others are “like me” in important respects; and I am “like them”. This is why we can join forces and assign interchangeable complementary roles to you and me – and children from around one and a half years seem to understand this structure as indicated in their spontaneous role reversal imitation: when in the context of a new collaborative activity they first learn to play one of two complementary roles, they then spontaneously reverse roles and act out the other part when appropriate (Carpenter, Tomasello & Striano, 2005).

Thinking about oneself as bound up with others in collective intentionality not only involves thinking of others and oneself as equivalent in certain respects (e.g., as interchangeable role-fulfillers in complementary activities), but essentially involves normative elements. The notion of a person is to some degree a forensic notion, as has often been stressed since Locke: Persons are the bearers of rights, duties and commitments among each other. In one of the first articles that introduced the notion of we-intentions, Sellars (1962) links these insights regarding the concept of a person to collective intentionality, normativity and self-consciousness:

>T> to think of a featherless biped as a person is to construe its behaviour in terms of actual or potential membership in an embracing group each member of which thinks of itself as a member of the group. Let us call such a group a ‘community’. […] Now, the fundamental principles of a community, which define what is ‘correct’ or ‘incorrect’, ‘right’ or ‘wrong’, ‘done’ or ‘not done’, are the most general common intentions of that community with respect to the behaviour of members of the group. It follows that to recognize a featherless biped or dolphin or Martian as a person requires that one thinks thoughts of the form, ‘We (one) shall do (or abstain from doing) actions of kind A in circumstances of kind C’. […] Thus the conceptual framework of persons is the framework in which we think of one another as sharing the community intentions which provide the ambience of principles and standards (above all, those which make meaningful discourse and rationality itself possible) within which we live our own individual lives. (p. 76-77).

Thinking of oneself as “one of us” thus is essentially normatively structured, is locating oneself in a shared “space of reasons”. There is no evidence to date that any animals apart
from humans places themselves in such a shared space and think of themselves as bound up with each other in such ways, as “one of us”. Human infants from their second year [on?], in contrast, though still a far cry away from adult communal and institutional reasoning, begin to engage in collective intentionality with others, setting up normatively structured spaces of joint activities, e.g., in the domain of games. This nascent collective intentionality goes along, one could say if one wanted to add to the ever growing catalogue of forms of self-awareness, with a nascent ‘collective self’ – consciousness of oneself as member of an embracing group of cooperators.

This is one way in which growing self-consciousness is directly linked to collective intentionality. If we zoom out now again on the broader question of the relations of forms of intentionality and self-consciousness from a comparative perspective, some more indirect links come into focus. In particular, collective intentionality seems to lay the ground for more complex forms of individual higher-order intentionality which in turn is constitutive of very central and morally relevant forms of mature self-consciousness. Let me explain: Humans from around 4 years, as we have seen, develop more complex forms of second-order intentionality – in fact, what they develop is the basic structure of our mature folk psychology: they ascribe to others and themselves different kinds of propositional attitudes, beliefs and desires as central cases, and make use of these in rational action explanation. No other species, not even chimpanzees (who are capable of simpler second-order intentionality) show any indications of such complex folk psychology (see Call & Tomasello, 2008). Why is that the case? Do humans have, and other species lack, special adaptations for such complex second-order intentionality? Possible, but unlikely, we have argued (Tomasello & Rakoczy, 2003). More plausible is the following: we know from much recent research that complex folk psychology (in particular, ascribing beliefs) is highly dependent on quite extended experience with specific language and discourse (see Astington & Baird, 2005, for a review of the empirical findings; for recent more principled philosophical arguments that complex folk psychology is necessarily linguistically mediated or constituted, see Bermudez, 2003; Garfield et al., 2001). And engaging in discourse is a cooperative enterprise that human children grow into from the second year, but that other species do not develop (though they have all other kinds of cognitive prerequisites for language) because of their lacking abilities of shared or collective intentionality. So humans and other species share simple second-order individual intentionality (understanding perception and intentional action), but only humans have we-intentionality, develop discourse and, based on discourse experience, complex second-order intentionality.

And quite clearly, such forms of complex higher-order intentionality constitute central and morally relevant aspects of distinctively human self-consciousness: The ability to entertain higher-order beliefs, for example, turns us into epistemically self-conscious beings, open-minded towards questions of epistemic virtues and vices. And the ability to take a higher-order perspective on our desires, bringing our first-order desires into line with our more general ideas about who we would like to be, is – if Frankfurt (1971) is right – a constitutive element in acquiring a free will and becoming a person.
References


