Social cognition is the capacity to perceive and interpret the behavior of social agents. It has a bright side, which includes the ability to understand others’ intentional actions, infer their mental states, and share experiences; and it has a dark side, which includes misunderstanding, stereotypes, and prejudice. A large portion of social-psychological research over the past 20 years has focused on this dark side; undoubtedly, it has a pressing impact on human social behavior. However, of equal importance is the study of those fundamental capacities that can bridge the gap between two minds and enable social communities to flourish through joint action, close relationships, and an intelligible social structure. This brighter side of social cognition has received increasing attention in recent years and is the focus of this chapter.

Early writers on social perception (e.g., Cantril, 1947; Asch, 1952; Heider, 1958; Tagiuri & Petrullo, 1958; Ostrom, 1984) emphasized the unique features of persons that make them objects of social cognition: that they have motives, thoughts, and feelings; that they have a capacity to act intentionally; and that they perceive the perceiver back, so that humans engage in a process of mutual or reciprocal cognition (Hastorf et al., 1970; Fiske & Taylor, 1991).
This mutual recognition of humans as mental, intentional beings lies at the very core of social cognition (Wellman, 1990; Perner, 1991; Mitchell, 1997; Malle & Hodges, 2005).

In response to the variety of tasks and demands in social life, humans have evolved a suite of interrelated subsystems that together form what may be called the social-cognitive toolbox (Ames, 2004; Cloutier et al., 2005; Malle, 2005; Mitchell, 2006). This toolbox contains abstract concepts (e.g., *agency* and *intentionality*); processes of gaze following, automatic empathy, mimicry, and joint attention; and increasingly complex functions of imaginative simulation and mental state inference. These tools belong together not because they form a module or are implemented in the same brain areas; what unites them is their responsiveness to the social environment with its challenges of ambulant intentional agents – minded, intelligent, and unique individuals.

In what follows I will discuss extant knowledge on each social-cognitive tool and ask whether there are compelling data for or against universality of that tool. However, systematic data on this question are scarce and answers must remain tentative, so I will encourage research more often than review it and rely more on theoretical arguments than empirical demonstrations. However, making theory the starting point of future cross-cultural research may be a valuable complement to the more common documentation of broad cultural group differences.

**THE SOCIAL-COGNITIVE TOOLBOX**

**CONCEPTUAL FRAMEWORK**

Humans perceive people, and interactions among them, through a framework that conceptualizes behavior as fundamentally linked with mental states. This framework, variously called *common-sense psychology*, *folk psychology*, or *theory of mind* consists of two parts: (1) systems that filter, group, and integrate certain stimulus inputs into such concepts or categories as *agent*, *intention*, *belief*, and *reason* (D’Andrade, 1987; Leslie, 1995; Malle & Knobe, 1997; Kashima et al., 1998; Malle, 2004); and (2) assumptions about these categories and their relationships (Fodor, 1983; Morton, 1996). For example, coordinated movements of agents are classified into the category *intentional action* (Wellman & Phillips, 2001; Woodward et al., 2001), and the concept of intentional action relies on the interplay of multiple mental state categories, including *belief* and *desire* (Malle & Knobe, 1997).

The system of folk concepts begins to develop early in childhood, aided by an either innate or very early-developing sensitivity to certain stimulus configurations in streams of behavior (Baird & Baldwin, 2001). Between the second and seventh year of life, this sensitivity grows into a network of sophisticated conceptual assumptions about mind and action (Wellman, 1990; Woodward et al., 2001; Gergely & Csibra, 2003).

Even though this framework is typically labeled a “theory” (as in *theory of mind*; Premack & Woodruff, 1978; Wellman, 1990; Flavell, 1999), there is some
debate over exactly what this theory encompasses (e.g., inferences, experience-based simulations) and whether it is a theory in the first place. I suggest that theory of mind, or its cognate terms, refer to a conceptual framework and be kept separate from the variety of debated psychological processes, such as action parsing, simulation, and inference, that operate within this framework (Malle, 2004, 2005). The entire package of a conceptual framework and the suite of psychological processes then make up the larger phenomenon of human social cognition.

The central categories of the folk theory of mind are arguably agent, intentionality, and mind, and they are closely related to one another (Smedslund, 1997). Agents are entities that can act intentionally, intentional actions require a particular involvement of the mind, and only agents have minds. At first glance seeming circular, this web of concepts is anchored by specific perceptual-cognitive processes. For example, objects perceived as self-propelled and behaving contingently are classified into the category agent (Premack, 1990; Johnson, 2000). Having identified an agent, the human perceiver is sensitive to face, gaze, and motion patterns that reveal whether the agent’s behavior is intentional (Dittrich & Lea, 1994; Phillips et al., 2002), and further analyses of behavior and context lead to inferences of specific goals, beliefs, and emotions. Agents are also assumed to be rational, acting in efficient ways to achieve their goals (Heider, 1958; Malle, 1999; Gergely & Csibra, 2003). Specifically, agents integrate beliefs and desires to form an intention to act, and those beliefs and desires are seen as the agent’s reasons (Davidson, 1963). When explaining intentional behaviors, people indeed offer what they infer to be the agent’s reasons (Malle, 1999, 2004). According to the folk-conceptual framework, only intentional actions emerge from this structure of reason-based causality, whereas unintentional behaviors are produced “mechanically,” through an interplay of numerous causes inside and outside the person (Heider, 1958; Searle, 1983; Mele, 1992; Malle, 1999).

The intentionality concept represents the hub of the folk-conceptual framework because it separates the entire realm of behavior into intentional and unintentional events, for perceptual and cognitive purposes just as much as for evaluative purposes, most notably in the judgments of praise, blame, and moral as well as legal culpability (Shaver, 1985; Weiner, 1995; Malle & Nelson, 2003; Malle, 2006). Developmental change in the folk-conceptual framework also occurs primarily as a differentiation of the intentionality concept. Infants recognize intentionality first as a behavioral category of object-directedness (Wellman & Phillips, 2001; Gergely & Csibra, 2003), and only in subsequent years does the concept becomes richly mentalistic (Malle & Knobe, 1997). The first mental concept appears to be desire, followed by belief (Wellman & Woolley, 1990). Then intentions are differentiated from desires (Astington, 2001; Baird & Moses, 2001; Schult, 2002), a distinction that becomes quite fine-grained among adults (Malle & Knobe, 2001). At least by late adolescence, we also see skill and awareness of fulfilling one’s intention as additional requirements of full-fledged intentional actions (Malle & Knobe, 1997).
Questions of Universality

Some research suggests East–West differences in how much perceived agency is ascribed to groups and individuals (Menon et al., 1999; Kashima et al., 2005), but such studies presuppose existence of a similar concept of intentional agency while delineating possible differences in thresholds of its application (see also Oerter et al., 1996; Edge & Suryani, 2002). Other authors have claimed to observe different “agency conceptions” in Eastern and Western cultures (e.g., Hernandez & Iyengar, 2001), but the evidence speaks only to variation in what motivates people from different cultures to act (e.g., personal effort versus collective endeavors).

In favor of universality, many linguists count the concepts of agent and intentionality as fundamental to the way humans see the world, and linguistic forms of these concepts have been found across all known languages (Givón, 1975; Bybee, 1994; Wierzbicka, 1996). Developmental evidence similarly finds the concept of intentional agent across all studied cultures (Wellman & Miller, 2006). A few studies have even explored cultural stability of intentionality judgments and their conceptual underpinnings (see Malle & Knobe, 1997), and so far we see striking convergence between Western and Eastern cultures/languages, even across diverse age groups (Zaw, 2006; Ohtsubo, 2007).

Constituents of the intentionality concept (belief, desire, intention, control/skill, and awareness) appear to be themselves good candidates for universality. Wierzbicka (1996) indeed lists all of them, except intention, among the cross-linguistically derived universal conceptual primitives. Bybee (1994) complements this analysis with a convincing case for the universality of the intention concept. One specific concept, that of (false) belief, shows a reliable trajectory of development in 3- to 5-year-old children across many cultures and languages (Wellman et al., 2001). There was an initial indication of slightly different onset times of false-belief development across, for example, African Baka, Peruvian Indian, and US children. But a more recent study suggests that even onset times may be strikingly similar when children from different cultures (Canada, India, Peru, Samoa, and Thailand) complete exactly the same false-belief test (Callaghan et al., 2005). Similarly, Sabbagh et al. (2006) found that a variety of theory of mind tasks as well as executive function showed virtually identical onset and age development in Chinese and US preschool children, though the Chinese children showed significantly higher levels of executive function.

Among other specific mental state concepts, emotions have most extensively been studied as candidates for universality. The evidence on universal recognition of a core set of facially expressed emotions is good (Ekman, 1973; Brown, 1991). Many emotion concepts also appear to be perceived and experienced quite similarly on such dimensions as valence, expectedness, attention direction, and goal-obstructiveness (Wallbott & Scherer, 1988; Mauro et al., 1992). However, linguistic labels (e.g., joy, fear, anger, disgust, sadness in English) do not cut experience at exactly the same boundaries across diverse languages. Instead, the universal ingredient in emotions may be such conceptual primitives as intention, belief, and desire (Wierzbicka, 1996). The adaptive value of emotion expressions
and concepts arguably lies in its communicative functions: to provide the interaction partner with information about mental states that disambiguate an interpersonal situation (Fridlund, 1994; Chovil, 1997).

**FOUNDATIONAL PROCESSES**

A conceptual framework could not by itself meet the fundamental challenges of social cognition; it only categorizes stimuli but does not interpret them. Moreover, concepts are too abstract to provide any explanation, prediction, or guide for a response. Knowing that there is an agent who performs some action with some desire is better than seeing only blobs of different wavelengths, but the perceiver must infer what specific action was performed (e.g., stretching out hand) with what specific motive (e.g., greeting).

The human mind recruits many psychological processes that rely on and interact with the conceptual framework of mind and behavior. Table 12.1 lists processes that have been identified in the literature so far, ranging from relatively basic person and behavior processing to simulation and explicit inferences of mental states and stable attributes. I begin the discussion with the most elementary processes that are presumably available soon after birth.

**Identifying Agents**

For adults, the concept of *agent* is defined by the capacity to act intentionally whereas for infants it begins as a broad distinction that emerges out of a responsiveness to diagnostic features – including having a face, moving in a self-propelled manner, “rationally” bringing about outcomes, and responding contingently to the

<table>
<thead>
<tr>
<th>TABLE 12.1</th>
<th>Layers of Concepts and Processes in Social Cognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conceptual framework</td>
<td>— Agent, intentionality, desire, belief, intention, and so on</td>
</tr>
<tr>
<td>Foundational processes</td>
<td>— Identifying agents</td>
</tr>
<tr>
<td></td>
<td>— Face processing</td>
</tr>
<tr>
<td></td>
<td>— Gaze following</td>
</tr>
<tr>
<td></td>
<td>— Parsing behavior streams</td>
</tr>
<tr>
<td></td>
<td>— Mimicry and imitation</td>
</tr>
<tr>
<td></td>
<td>— Automatic empathy</td>
</tr>
<tr>
<td></td>
<td>— Recognizing goals</td>
</tr>
<tr>
<td></td>
<td>— Joint attention</td>
</tr>
<tr>
<td>Mediating processes</td>
<td>— Self-other discrepancies</td>
</tr>
<tr>
<td></td>
<td>— Inhibitory control</td>
</tr>
<tr>
<td></td>
<td>— Self-awareness</td>
</tr>
<tr>
<td></td>
<td>— Simulating other minds</td>
</tr>
<tr>
<td>Explicit mental state inference</td>
<td></td>
</tr>
<tr>
<td>Inference of stable attributes</td>
<td></td>
</tr>
</tbody>
</table>
infant’s own behavior (Premack, 1990; Bertenthal, 1993; Johnson et al., 1998; Király et al., 2003). Within a few weeks after birth, infants respond favorably to contingently interacting persons but retreat from noninteracting persons (Watson, 1985; Legerstee et al., 1987), and 3-month-old babies discriminate between light point arrays that move randomly and those that are attached to a moving body (Bertenthal, 1993). In adults, there is also suggestive evidence for a region of the visual cortex that responds selectively to images of the human body (Downing et al., 2001). During the first year of life, infants refine their discrimination of self-propelled objects by paying attention to the ensemble of faces, body parts, and articulated motion (Markson & Spelke, 2006). Contingent interaction (the other acts in response to one’s own actions) becomes the central feature for identifying agents (Johnson, 2000), reflecting perhaps the parallel between infants’ own discovery of intentional action control and their recognition of such control in others (Russell, 1996; Hauf & Prinz, 2005).

**Questions of Universality**

Currently, no cross-cultural data directly speak to the psychological processes of agency detection. However, the very early-developing responsiveness to particular features (self-propelledness, contingent interaction) and the fact that those features are arguably diagnostic of actual agentic behavior make substantial cultural variation unlikely. Culture can certainly constrain or expand whether particular unobservable entities fall under the agent concept broadly speaking – such as gods, ghosts, and forces of nature – but the feature-sensitive processes of agency detection apply to observable entities only, and for those, universality can be assumed until contradictory data emerge.

**Face and Gaze Processing**

One of the critical features of agents is that they have faces – or so the human brain expects. The human sensitivity to faces has been well documented, from the behavioral to the cognitive to the neural level (Kanwisher, 2000). Sensitivity to faces may be a prerequisite for the newborn’s ability to imitate adult facial behaviors (Meltzoff & Moore, 1977) and to affectively engage with them in mutual smiling. It also allows for the later-developing tendency of emotional contagion (Hatfield et al., 1994), for automatic extraction of categorical information such as sex (Cloutier et al., 2005), and for efficient memory of one’s preferred social partners. The sensitivity to eyes is especially notable, as the remarkable speed at which faces are recognized is lost when eyes are removed (Rousselet et al., 2004).

But we must distinguish between the processing of two aspects of faces: the structural and the dynamic (Posamentier & Abdi, 2003). Structural processing can be easily done with static faces and allows for fast (100–200 ms) identity recognition, relying considerably on the target’s eyes (Mason et al., 2004). Dynamic processing responds to change of gaze or of facial expressions that indicate the target’s altered emotional state. Here too, they eyes play a central role (Bassilli, 1978; Buchan et al., 2007). The distinction between structural and dynamic face processing is also
borne out by data at the neural level. Structural features are heavily computed in the fusiform face area (FFA), and a lesion to this area is associated with impaired recognition of facial identity (Damasio et al., 1982). Eye direction and especially gaze change are heavily computed in a posterior part of the superior temporal sulcus (STS), as both patient studies and single-cell recordings in monkeys indicate (Campbell et al., 1990). On the whole, however, the processing of face and gaze are distributed over a number of brain regions that typically serve other functions as well. In particular, the more dynamic the stimulus processing, the less evidence we find for dedicated modules that are anatomical correlates of function (Farah & Aguirre, 1999). Whereas the FFA is a circumscribed region sensitive to static face perception, the STS also processes language stimuli and is involved in general visual attention. Moreover, other brain areas besides the STS process gaze as well. For example, the amygdala is activated by moving eyes, and amygdala lesions impair gaze perception (Young et al., 1995; Kawashima et al., 1999).

**Gaze Following**

When people follow another person’s gaze, they recognize or extrapolate direction and target of the person’s looking behavior. Gaze following plays important roles in goal inference, joint attention, and social coordination (Frischen et al., 2007), but it typically operates quite reflexively. Adults automatically orient toward the target of a facing partner’s gaze even if there is an incentive to look in the opposite direction (Driver et al., 1999). Gaze following is also present in other animals, such as dogs (Hare & Tomasello, 1999), and chimpanzees can compute the object of a target’s gaze even when that object is not directly visible to them. For example, they look around a barrier to locate the apparent target of the other’s gaze (Brauer et al., 2005).

Gaze following by itself does not constitute a mental state inference, as following another person’s line of sight is a process distinct from constructing the person’s actual mental perspective. We know this because visual perspective taking, but not mere gaze following, allows the perceiver to quickly determine whether an object is to the right or to the left of a target person (Michelon & Zacks, 2006). Gaze following is essentially an eye-behavior monitoring and prediction process, and many other, more mind-sensitive processes have to fall into place before such monitoring delivers mental state inferences. If the perceiver does try to infer mental states, gaze information is readily used (Castiello, 2003).

**Questions of Universality**

The low-level basis of face processing, its availability in neonates, and the absence of any known cultural variation in this sensitivity suggests a universal capacity. Similarly, gaze following is partially reflexive and evolutionarily old, making it a strong candidate for universality as well. The biological function of an agent’s gaze obviously lies in orientation and regulation of locomotive and manipulative behavior in a three-dimensional world. As observers, many organisms take gaze as reliable predictor of action, but humans build on their expanded
conceptual capacities to interpret gaze as a signal of mental states as well (e.g., interest, distraction, discomfort) even in the absence of any anticipated action. Humans’ universal experience of the link between their own gaze and such basic states as attention and intention may facilitate such finer interpretation, but culture also assigns differential meaning to some of these signals, such as looking down as boredom or bashfulness.

Parsing Behavior Streams

One of the most basic (but certainly not easiest) tasks for humans is to properly parse the continuous stream of behavior observed in other people (Baird & Baldwin, 2001). Even though behavior is varied, unpredictable, and interactive, there are plenty of regularities in its production that allow humans to segment the behavior stream into meaningful units. Indeed, by 11 months, babies form clear expectations about the boundaries between action units (Baldwin et al., 2001). But what are those units? An ideal system would parse the perceived stream into the very units that underlie the production of the observed behavior in the first place (Prinz, 1990; Wilson, 2001). From what we know about action control, intentions play a fundamental role in structuring behavior production (Wolpert et al., 2003). And intentions indeed serve as a fundamental unit of behavior perception. This match between production and perception can be found at a very basic level, in what has been termed the “mirror neuron system.” Rizzolatti et al. (1996) discovered that single neurons in the macaque monkey’s brain fire selectively both when the monkey performs a certain action and when it observes that same action being performed by another agent. Among humans, that same matching capacity is implemented in a far more distributed manner, and neuroimaging studies have found extended brain areas that are highly active both in the production and the perception of certain actions (e.g., Decety, 2002; Decety & Grèzes, 2006). With a link between produced and perceived actions, finding the units in the behavior stream becomes easier through practice of performing certain behaviors, and even 3-month olds appear to benefit from such practice transfer (Sommerville et al., 2005).

Sensitivity to intention-based units in the behavior stream does not entail a full understanding of intentionality. It does, however, present an opportunity – through much repetition – for becoming familiar with prototypical characteristics of intentional action. These features include smooth movement execution, a clear beginning and end, as well as appropriate body, head, and face orientations. Armed with this sensitivity for detecting intentional action, children develop a strong preference for imitating those actions (Carpenter et al., 1998) and a tendency to learn words that refer to intentional, not accidental behaviors (Tomasello & Barton, 1994).

Universality

The process of behavior parsing appears to be neurally quite basic and a requirement for simple forms of imitation found in other animals (Byrne, 2006). Cultures may single out certain actions to be significant that are therefore parsed
more efficiently by members of that culture. However, nothing in the process itself is open to cultural modulation because cultural influences presuppose the very mechanism of parsing (no social learning and communication would be possible without it).

**Imitation and Mimicry**

Humans are the only species in which members instruct each other by demonstrating a novel behavior and inviting the other to perform it in turn. Complementarily, humans are exceptional observational learners, capable of picking up even complex sequences of action after observing them performed only once. This capacity to learn by imitation has great advantages for the helpless human infant who is born with a third of the adult’s brain size and has very few functions hardwired from the start. Imitation is also rewarding because it aligns oneself with valued others in reciprocal interaction and imitative play, amply illustrated by the smile game or that of ceaselessly rolling a ball back and forth (Nadel-Brulfert & Baudonnière, 1982).

Where do those more intentional and sophisticated forms of imitation and instruction come from? In part they build on a simpler copying mechanism called mimicry, which is evolutionary much older (e.g., found in monkeys, birds, even octopi). Reviews by Hatfield et al. (1994) and Chartrand and Bargh (1999) document that social mimicking of gestures, postures, and mannerisms often occurs automatically, without the interactants’ awareness or intention. However, in humans this rather automatic type of imitation has important social functions: it both expresses feelings of social affiliation and can increase affiliation and rapport among mimicking partners (Bernieri, 1988; Chartrand & Bargh, 1999). One way to explain mimicry is with the “mirror system” mentioned above, which transforms perceptions of the other’s behavior into corresponding motor patterns. These patterns have to be preexisting and are primed by their perceptual representations. When these primed patterns are actually executed, we have a case of mimicry (Byrne, 2006). Surprisingly, however, contagious yawning – what one would consider a prototype of automatic mimicking – does not selectively involve activation of the mirror neuron system (Schürmann et al., 2005). We may have to distinguish between two kinds of mimicry. A low-level one relies on reflexive stimulus-response connections and explains phenomena such as contagious yawning and the coordination of flocks of birds. A higher-level one involves action parsing and processing with a sensitivity to context and function (Chartrand & Bargh, 1999). The significant step from such higher-level mimicry to full-blown imitation is the role of novelty: apparently only humans can imitate and teach truly novel behaviors – for which no preexisting neural program exists (Byrne, 2004).

**Universality**

All the current evidence points to mimicry and imitation to exist and function universally across cultures. Forms of imitation have been demonstrated in newborns (Meltzoff & Moore, 1977), and mimicry at various levels of sophistication
exist in many other species. Just like behavior parsing, imitation, and social teaching are prerequisites for human culture itself, ranging from manufacturing tools to establishing social norms to the acquisition of language (e.g., Gergely & Csibra, 2004; Norenzayan et al., 2006).

**Automatic Empathy**

Mimicry is a tool of behavior matching. Closely related is a mechanism of matching mental states, which has been called automatic empathy or emotional contagion (Hatfield et al., 1994; Hodges & Wegner, 1997; Levenson & Ruef, 1997; Stueber, 2006). It involves a target person having a feeling $F$, expressing that feeling in characteristic ways, and the perceiver feeling $\approx F$ as a result. Simple forms of such contagion have been observed in nonhuman primates (Preston & de Waal, 2002), in 3-month-old infants (Haviland & Lelwica, 1987), and even in 34-hour-old newborns who are more likely to cry when other newborns cry (Simner, 1971). How does it work? A plausible model suggests that the perceiver observes the target’s expressive behaviors, something akin to a mirror system activates (but does not execute) the perceiver’s own copy of those behaviors, and feelings associated with this behavior program are activated in the perceiver as well (Hatfield et al., 1994). Thus, mimicry may be at the heart of the loop, but both the target’s behavior and the mimicker’s behavior must be reliably associated with certain emotions, and those are then synchronized (Bavelas et al., 1987).

Automatic empathy would in principle allow perceivers to read off the other person’s emotions from their own emotions. However, perceivers typically do not appreciate that their own emotions may originate in the other’s emotions. As a result, emotional contagion can increase rapport in some cases (La France, 1982) but escalate conflict or instigate crowd violence in others (Levenson & Gottman, 1983; Patten & Arboleda-Flórez, 2004). A skillful social agent could therefore manipulate another person who is in the perceiver role and experiences emotional empathy, just like good actors can cause us to clench our fists, shed tears, and suffer complex emotions (Coplan, 2006).

**Universality**

The capacity to “catch” another person’s emotion is a convincing candidate for universality, especially considering that Levenson and Ruef (1997) identified varieties of contagion that reach deep into the autonomous nervous system and even into the hormonal system. Kokkinaki (2003) provided cross-cultural evidence (though limited to Crete and Scotland) for a reliable interplay between imitation and emotions, involving contagion as well as more systematic affect attunement, and Hatfield et al. (1994) identified cross-cultural evidence for contagion in Malaysia, Africa, and the Yakut territory in northern Asia. Because emotional contagion requires a reliable link between the target’s emotional state and the manifested expressive behavior and another reliable link between the perceiver’s mimicking behavior and this person’s emotional state, there is some room for cultural variation. For one thing, automatic empathy may be less pronounced in cultures that have tighter display rules. For another, expressive
behaviors may not be exactly the same from culture to culture, so even if a perceiver mimics the other’s expressions, he or she may not have an associative link from those expressions to the *same* emotion that generated the target’s expressions in the first place. Even though such a situation would involve behavioral mimicking, the perceiver does not catch the other person’s actual emotion.

**Recognizing Goals and Intentions**

When parsing the stream of behavior into meaningful units, humans do not attend to behavior by itself but to another person’s acting upon the world. This process requires an understanding of self-propelled agents being directed to objects in the world (Wellman & Phillips, 2001). Directedness reveals itself in simple cues (head turn, gaze orientation) and predicts characteristic actions (e.g., grasp, show) as well as emotions (e.g., satisfaction of reaching the goal). In this way, infants who are 6- to 12-months old begin to appreciate that behaviors can be directed at something (Woodward, 1998; Gergely & Csibra, 2003), and this is their first concept of intentional action. The direction is typically toward a physical object, and the actions toward it are basic, such as grasping, placing, or moving (Woodward et al., 2001). Such behaviors trigger the infant’s notion of goal-directedness most reliably when performed by a human arm and hand, not by a mechanical device (Woodward, 1998). However, even a gloved hand is interpreted as acting in a goal-directed manner if the hand has been previously associated with the whole person (Guajardo & Woodward, 2004). This pattern illustrates quite powerfully the relationships among conceptual components of the human theory of mind (here, between *agent* and *action*).

**Intentions as Action Contents**

It is one thing to recognize an action as intentional; it takes additional processing to recognize the content of the action – *what* the person is trying to do. Adults, who have accumulated experience with many actions, show a remarkable capacity to recognize both intention and details of the planning process revealed in the kinematics of human behavior (Runeson & Frykholm, 1983). For example, if a target person knows from previous lifts how heavy a given object is and prepares to lift that object again, perceivers can judge the object’s weight merely from the (point-light) outlines of the target’s body as it prepares the movement. Remarkably, if the target intends to pretend lifting a heavy object that is actually quite light, people recognize the deceptive intention as well as the object’s true weight. Inferring intentions from minimal behavioral clues is well illustrated by athletes who are sensitive to the other person’s tiniest indication of what they are trying to do – defensive players in football trying to read the offensive players’ intentions from looks or body twitches or soccer keepers trying to kill a penalty kick.

The origins of this remarkable capacity can be seen in the 9- to 12-month old’s understanding of simple goal-directed actions such as grasping or handing something over. However, these actions are transparent – they form a perceivable unit with their observable goal. Other behaviors, such as pulling a rectangular object out
of one’s pocket and putting green paper on the counter, require considerable world knowledge to understand. This world knowledge grows rapidly during the second year of life, and it allows children to infer an agent’s intention even if the corresponding goal object is not visible. In a study by Meltzoff (1995), 18-month olds who watched an adult try but fail to perform certain acts on an unfamiliar object imitated what they saw the actor do but also continued on with actions that the adult had never performed but apparently had intended to perform. Thus, children were able to infer the actor’s goal or intention from the behavior. This, however, occurred only when the failed attempts were performed by an agent, not when the “other” was a mechanical device (Meltzoff, 1995). In the same vein, Russell and Danish (2007) found that children imitate an incomplete action (and achieve completion) only when the action was performed by a human hand, not when it was performed by an arrow.

In each of these studies, children may infer the actor’s intentions without necessarily conceptualizing them as genuinely mental states. However, they do go beyond the observable behavior as they infer a complex goal object – an action sequence or an action outcome – that is itself not observable. Perhaps the child’s own experiences of attention, planning, and voluntary action control, as they get increasingly sophisticated, provide the template for eventually inferring a mental state (Russell, 1996; Rochat, 2002).

**Universality**

The world knowledge involved in goal and intention inference brings culture strongly into play. Cultures provide and demand practice of certain scripts and schemas that are immediately understandable for members of the given culture but mysterious to those outside. Just consider the variation across cultures and languages in the bodily signals that represent “yes,” “no,” or certain insults. Despite this variability of specific scripts, the general capacity to infer goals and intentions does not appear to be culturally variable. No cross-cultural data speak to 18- to 24-month-old children performing the abovementioned goal and inference tasks. However, evidence on older children shows no relevant cultural variation. Tardif and Wellman (2000) replicated in Chinese children the pattern found in American children (Bartsch & Wellman, 1995) that the onset of using desire verbs (during the third year) precedes the onset of using belief verbs (during the fourth year). Likewise, using experimental theory of mind tasks, Wellman et al. (2006) showed that both mainland Chinese and American children ages 3 to 5 understood concepts of desire before they understood concepts of belief. Simpson (2003) compared European American and Mexican American children 4- to 9-years old and found no differences in their use of desires to explain behavior. Finally, Barrett et al. (2005) showed that German adults and 4- to 5-year-old children as well as Shuar adults from Amazonian Ecuador performed similarly in using abstract motion cues to correctly to infer such intentions as chasing, fighting, courting, and playing.

**Joint Attention**

Humans enter situations of joint attention when both participants attend to an object and are also aware of each other’s attending to that object. This is a
deeply social process because the two individuals form a higher unit, a “we,” that delivers affective, cognitive, and behavioral benefits. Like no other animal, people seek out and enjoy sharing thoughts, emotions, attitudes, perceptions, and knowledge (Higgins & Pittman, in press). Having the same thought or emotion is deeply rewarding, validating, and bonding. Human language is the best vehicle to achieve such acts of sharing, as in disclosure, gossip, and story telling. Communication itself is an act of sharing as well: Speakers establish a common ground of what is mutually known, what each speaker’s symbols refer to, and what the overarching purpose of the communicative exchange is (Clark, 1996). And language acquisition already depends on the capacity to share attention to objects, because learning the name of an object requires that the new speaker and the established speaker both know to what object a linguistic term refers (Baldwin, 1995).

Joint attention develops in the 9- to 18-month-old child, originating from two important patterns. The first pattern is the child’s affective tuning to the adult’s attitude toward objects (Hobson, 2005). For example, when the child registers the adult’s directedness to an object, the child tends to take on the adult’s attitude toward that object, exhibiting emotional contagion. The second pattern is the adult’s inclination to attend to whatever the child happens to be attending to at the moment, thereby facilitating alignment between attention foci (Franco, 2005). For example, the child grasps a ball, and the adult comments: “Yeah, look at the ball.”

Together, these two tendencies cover a large number of situations: when the child already attends to an object, adults align their attention with the child’s; when the child is not yet attending to something, the adult’s expression of interest can draw the child into aligned attention to a particular object. During the second year of life, the child begins to point to objects that are out of reach, thus inviting the rewarding state of joint attention and attitude sharing. The child also checks the adult’s attention before pointing, showing clear signs of appreciation for the fact that people can attend to different things. Children also look to the other’s facial and bodily cues to determine the appropriate attitude toward an object or action (what is called social referencing), and they use attentional cues to determine toward which object the other is showing a certain attitude (Repacholi, 1998; Moses et al., 2001).

The affective alignment we see in the earliest forms of joint attention continues far into adulthood in the process of social convergence, famously demonstrated in Sherif’s (1935) autokinetic studies. There, people jointly perceived an ambiguous stimulus – the apparent movement of a light dot in a pitch-dark room – and spontaneously arrived at a consensual judgment (the width of the movement) within just a few trials of the experiment. Their convergence manifested not only cognitively but in an affective alignment as well. They jointly agreed on what was “right,” and they defended that stance in subsequent rounds of making judgments in the presence of other people. Thus, adult norm formation may derive from the fundamental human tendency to align one’s attitudes with those of others.
**Universality**

Current knowledge of joint attention nurtures confidence in the assumption of its universality. The joint attention capacity appears to be a necessary or at least highly facilitative ingredient in language learning (Tomasello & Farrar, 1986; Baldwin, 1995), and if it is compromised, as in autistic children, linguistic and social behaviors are severely compromised as well (Mundy, 1995). The universal adaptive value of joint attention for social living is also compelling in light of its connection to norm formation and moral judgment, two universals in their own right. So until we discover new data, joint attention should be considered a fundamental human (perhaps primate) process of social and cognitive relatedness without which neither language nor culture could exist.

**MEDIATING PROCESSES: EN ROUTE TO MENTAL STATE INFERENCES**

Out of all the discussed foundational concepts and processes, humans weave an explicit understanding of mental states that enables them to read emotions, desires, and interests, detect deceiving intentions, and actively take even a stranger’s perspective. One of the greatest puzzles for social-cognitive science is to delineate exactly how humans develop, at both the process and conceptual level, from sensitive behavior readers to perceivers of the mind, forming explicit and verbalizable representations of the large variety of human mental states (Coricelli, 2005). The puzzle has not yet been solved, but existing evidence allows us to identify three contributors that will have to be part of the final analysis: the experience of discrepancies between self and other; the capacity for inhibitory control; and the maturing of self-awareness. Finally, the process of simulating other minds encompasses all three of these capacities (as well as earlier ones) and is perhaps the central bridging mechanism between foundational processes and full-blown mental state inferences.

**Self-Other Discrepancies**

By the beginning of the second year, toddlers move independently and begin to experience conflict between their own desires of touching, moving, and swallowing objects and caretakers’ desires (Biringen et al., 1995). The child learns, sometimes through clashes of will, that people differ in interest and desire, and these experiences of mismatch may well force the development of concepts that distinguish different mental states in different people. In one study (Repacholi & Gopnik, 1997), 18-month olds understood that an adult apparently disliked (showed a disgust face toward) an object that the child liked (goldfish cracker), and they understood that the adult liked (showed a happy face toward) an object that the child disliked (raw broccoli). When the adult said, “Give me some,” the 18-month-old children handed the adult the object that the adult liked (raw broccoli), whereas a comparison group of 14-month olds handed the adult the object that they themselves liked.
Considerable evidence shows that the desire concept is the first mental concept that children master reliably (Wellman, 1990; Perner, 1991), and mental state verbs of desire indeed appear in children’s speech just around 18 months (Bartsch & Wellman, 1995). Increasingly during the preschool years, children use mental state talk to negotiate conflict (Howe et al., 1998), and it might be in part these conflicts – with parents, siblings, and peers – that demand the understanding of others’ mental states. Indirect evidence also comes from studies in which the presence of siblings predicts accelerated theory of mind development (Lewis et al., 1996; Peterson, 2001). Even though these studies typically measure theory of mind as false-belief understanding, the importance of conflict remains, because explicitly recognizing another person’s distinct beliefs represents the experience of epistemic conflict.

The experience and regulation of conflicts may thus help shift the early tendency toward self-other convergence (e.g., in emotional contagion) to a more flexible process in which the child’s own internal state may be a starting point, but differences between self and other are simulated and predicted as well (Gordon, 1992; Goldman, 2006). Of course, the self-other distinction is not perfect: the stronger the child’s own desire, for example, the less likely the other person’s distinct desire will be appreciated (Moore et al., 1995). The same holds in adults, for whom the salience and intensity of an internal state can distort, slow, or prevent inferences about the other person’s internal state (e.g., Malle & Pearce, 2001; Van Boven & Loewenstein, 2003; Barr & Keysar, 2005; Birch, 2005). Thus, it takes a corrective mechanism to bring the other person’s mind into the foreground; and that mechanism is inhibitory control.

**Inhibitory Control**

The child’s experience of conflicting desires (and later beliefs) challenges, I argued, the earlier processes of self-other attunement and matching (e.g., in imitation and joint attention). That earlier stage is characterized by a “like-me” assumption (Meltzoff & Brooks, 2001), a form of naïve realism in which the child’s own emotional, motivational, and epistemic states are assumed to be shared by others. To break up this equation and appreciate conflicting states the child must inhibit its own mental state just enough to recognize the alternative – somebody wanting, liking, feeling, or thinking something else. The facilitative role of inhibitory control in developing theory of mind performance is well documented, though primarily for the classic false-belief test (Leslie & Polizzi, 1998; Carlson et al., 2002; Lang & Perner, 2002). Already earlier, from about 16 months on, inhibitory control may enable children’s capacity for pretense (Bosco et al., 2006), which can be considered an inhibition of a default interpretation (e.g., of actions and objects) in favor of new interpretation, while keeping both interpretations to some extent active. Also, inhibitory control supports emotion regulation, such as suppressing a momentary urge to receive a larger reward later on (Metcalfe & Mischel, 1999), and it is needed for deception (Hughes, 1998) and irony (Lucariello & Mindolovich, 1995). Not surprisingly, then, inhibitory control skills correlate with mental state inference.
skills even in adulthood (Chasiotis & Kiessling, 2004). Inhibition strengthens both the emergence of explicit mental state inferences in the child and the instigation of such inferences in the adult (Epley et al., 2004).

Self-Awareness

A third process that supports the emergence of mental state inferences is the child’s increasingly sophisticated self-awareness. For if the mind can discriminate among its own states (which are computationally and experientially easily accessible) it can also learn to discriminate those of others. Inhibitory control is already a type of self-awareness in that it operates on an intention (later on other representations) and is thus a mental state regulating another mental state.

The more general capacity for self-awareness is often indexed by the mirror self-recognition test (Gallup, 1970). Children pass the test if, while watching their mirror image, they detect and remove a color dot from their face that the experimenter had unobtrusively placed there. This bodily self-recognition emerges at around 14 months (Brooks-Gunn & Lewis, 1984), and with it comes a more reliable discrimination between one’s own and other people’s behaviors. Synchronous imitation of another person’s behavior, for example, emerges only among children who pass the self-recognition test (Asendorpf et al., 1996). Visual self-recognition also predicts the child’s joint attention at 18 months (Nichols et al., 2005), which itself predicts later social-cognitive functioning (Charman et al., 2001). However, 2-year-old children with autism pass the self-recognition test but fail most joint attention tests (Mundy, 1995), so joint attention capacities rely on more than self-recognition.

Simulating Other Minds

According to the well-known simulation theory of human mentalizing, social perceivers use their own faculties of perceiving, feeling, and thinking “off-line” to interpret other people’s perceptions, feelings, and thoughts. They pretend, as it were, to be in the target’s particular situation and recruit their own situation-appropriate mental states to attribute to the other person (Gordon, 1992). However, many phenomena have been related to or subsumed under the term simulation (Goldman, 2006). In the simplest case, a perceiver’s motor system mirrors another person’s action but does not perform the relevant action (Wilson, 2001); rather, the goal state that normally goes along with that type of action is activated and the perceiver understands the target person’s action in terms of that goal state. A related second phenomenon is projection (Ross et al., 1977; Ames, 2004). The social perceiver assumes that “other = like me” and therefore ascribes his or her own (simulated or experienced) mental states to the other person (Meltzoff & Brooks, 2001). When failing to take into account the other’s distinctness, perceivers who merely project display a naïve realism (Ross & Ward, 1996) or egocentrism (Barr & Keysar, 2005). The third variant of simulation is the sophisticated attempt to construct a model of the other person’s situation and of his or her most likely mental states from all available data: observed
actions, empathic responses, general knowledge, inference, and so on. Perceivers may still assume some degree of similarity with the target person but patch mere projection and actively try to identify the target’s mental states that differ from their own (Gordon, 1992; Epley et al., 2004).

Even though many precursors to mentalizing, such as motor mirroring, automatic empathy, and imitation, have been adduced as evidence for the fundamental nature of simulation (e.g., Gallese & Goldman, 1998; Goldman, 2006), I consider it most prudent to use the term simulation only for the process of more or less explicitly constructing a model of the other person’s mental states. Some of the precursors share with this more sophisticated simulation process the involvement of the perceiver’s own internal states – action programs, associated goals, contagious emotions, and the like. But as the earlier discussion showed, the precursors that somehow involve “the self” are manifestly distinct, and subsuming, say, motor mimicry, projection, and model construction under one broad notion of simulation is not fruitful. A more specific conception of simulation also connects well to simulation processes that have been described in the decision-making literature, in which the agent simulates possible future states and makes a decision according to the value of those simulated outcomes (Klein & Crandall, 1995). The specific conception of simulation also goes beyond the classically warring factions of simulation theory. When constructing another person’s mental states, the perceiver makes use of all available knowledge, rules, theories, and his or her own mental states, real or imagined (Nichols & Stich, 2003). Simulation is thus one important process in the social-cognitive toolbox – alongside knowledge retrieval, inference, and even some foundational processes that may resemble modular systems.

Universality

Of the mediating mechanisms that facilitate the young human’s transition from intelligent behavior reading to genuine mind reading, all are good candidates for universality, though, as usual, the evidence is scarce. Self-other discrepancies, defined as clashes between two people’s goals (or preferences or attitudes) are bound to exist across all cultures, even though there will be variations in the specific goals over which, say, parents and children most readily collide. The culturally preferred resolution of such discrepancies may vary as well, with possible consequences for the speed with which children appreciate the distinctness of mental states. Inhibitory control has been identified cross-culturally both in studies of individual differences (Rothbart et al., 2001) and as a predictor of theory of mind development (Chasiotis et al., 2006). Self-awareness as a principled capacity is unlikely to vary cross-culturally (in fact, it has been shown in other primates and in dolphins), but culturally variable parenting styles appear to influence the time of onset for self-recognition and self-regulation (Keller et al., 2004). Simulation, finally, is hardly in question as a universal human capacity (Goldman, 2006), though the frequency and value of simulating other minds and possible worlds may well vary across cultures.
EXPLICIT MENTAL STATE INFERENCES

Perhaps the hallmark of the social-cognitive toolbox is the capacity to make explicit mental state inferences. This tool allows the social perceiver to consciously represent the other person’s mental state as a distinct fact of the world—but one that has a unique point of view and a subjective experiential quality. Sometimes called perspective taking, this capacity yields an understanding of the multiplicity of visual, affective, and interpretive views vis-à-vis the same reality. Typically developing children and adults not only appreciate the variability of desires, beliefs, intentions, and emotions but are also motivated to find out what specific mental state another person might have.

The famous indicator for whether a child has mastered explicit perspective taking is the false-belief test. In a classic paradigm (Wimmer & Perner, 1983), the child participant and one other person see the experimenter hide an object in container 1, next to an empty container 2. The other person then leaves the scene and the experimenter switches, in the child’s plain view, the object from container 1 to container 2. The other person returns and the child is asked where the other person will look for the object (or, more difficult, where the other person thinks the object is). The correct answer is of course “In container 1,” but the child’s own knowledge of reality is a powerful default response (Birch, 2005; Nickerson, 1999). Most children master this difficult inference at age 4, though some changes in the task or question formulation can lower this threshold (e.g., Siegal & Beattie, 1991; Freeman, & Lacohée, 1995). By contrast, genuinely mentalistic desire inferences emerge 1–2 years earlier.

The greater difficulty of belief over desire inferences can be seen in adults as well. When people explain intentional actions, they primarily use reasons, which are typically the beliefs and desires the agent had when deciding to act (Malle, 1999). However, when we compare people’s explanations for their own actions (the “actor” perspective) and their explanations for other people’s actions (the “observer” perspective), we reliably find that actors use more belief reasons, and fewer desire reasons, than observers do (Malle et al., 2007). Moreover, intimate observers use more belief reasons than stranger observers do (Malle et al., 2007). In recent research we also found that when people watch short video clips of agents performing a variety of everyday actions, they more readily infer the agents’ goals (desires) than their beliefs, and they are also faster in making these goal inferences (Holbrook, 2006).

Research on perspective taking in adults has shown that efforts to understand the other person’s mental states lead to a number of positive outcomes, such as greater perceived similarity with the person, reduced stereotyping, greater sympathy, forgiveness, and helping, though it can also lead to personal distress (for a review, see Davis, 2005). Accuracy of explicit mental state inferences has been described as moderate, though only one paradigm has been used so far to provide such accuracy estimates (Ickes, 1993). Very little research has been devoted to the actual psychological processes underlying explicit mental state inferences—either the integration of conceptual connections, simulations, and knowledge
retrieval, or the relationship between more implicit, lower-level tools (e.g., emotional empathy, joint attention) and the explicit form of mental state inferences. The dearth of work on these issues is a result of the emphasis in developmental research on the performance of such inferences (and the methodological limitations to studying processes in children) and the emphasis in social-psychological research on people’s inferences of personality traits. This type of inference is discussed below.

Universality

The cross-cultural evidence speaks strongly for the existence of false-belief inferences and their developmental order as succeeding desire inferences. Both analyses of speech patterns (Tardif & Wellman, 2000) and numerous experiments support the existence and natural order of desire and belief inferences (Wellman et al., 2001; Callaghan et al., 2005; Wellman et al., 2006). More generally, however, cultures may vary in the effort and readiness with which their members explicitly infer other people’s mental states (Lillard, 1998; Wu & Keysar, 2007). Unfortunately, the patterns of results are by no means clear (and sometimes contradictory), as cultural differences can occur at many different levels. It may be more or less appropriate to publicly utter mental state ascriptions; it may be more or less difficult to gather behavior cues as indicators of mental states (e.g., because of display rules); and there may be more or less impact of status, power, and practice in performing certain inferences. For example, in hierarchical societies, those lower in power are likely to exert more effort and gain more practice in making such inferences. Despite the interesting dynamics revealed by such differences, the universality of mental state inferences in general should not be in question.

INFERENCEs OF STABLE ATTRIBUTES

Human social perceivers ascribe not only momentary mental states to others; they also infer abstract and stable attributes, including status, personality traits, attitudes, and abilities. Social-psychological research has uncovered a considerable amount of knowledge on the inference of attitudes and personality traits, and many reviews of this research are available (Gilbert & Malone, 1995; Uleman et al., in press). Here I will briefly relate these stable attribute inferences to mental state inferences.

Function

Whereas the primary function of the social-cognitive tools described above is to make sense of immediate events (such as actions and mental states), the primary function of attribute inferences is to accumulate more general information about the person. This information is useful for predicting future behavior tendencies, so attribute inferences will be frequent when encountering strangers – in a job interview, at a party, on a first date. In interactions with familiar others, by contrast, attribute inferences will be rare whereas mental state inferences are
still continuously needed. To illustrate, in folk explanations of behavior, mental states are cited about 65% of the time, whereas stable person attributes are cited no more than 5% of the time (Malle, 2004; Malle et al., 2007; see also Lewis, 1995).

Accuracy

Research has shown that people are capable of inferring stable attributes from behavior even without any awareness of making those inferences (Uleman et al., in press). Other work suggests that people may be all too willing to make such inferences (Ross, 1977; Gilbert & Malone, 1995). At the same time, evidence exists that people’s tendency to infer traits from minimal information comes with noteworthy, though certainly not perfect, accuracy (Funder, 1999), and having such an inclination to infer traits appears to be associated with positive, not negative social adaptation (Block & Funder, 1986). Overall, questions of accuracy and adaptive value are difficult to assess as long as no ecologically sensitive data sets are available that would document the frequency of spontaneous attribute inferences, the relative frequency of correct and incorrect ones, and the costs and benefits of these inferences in social life.

Inference Hierarchy

Inferences about specific mental states can be informed by knowledge about the target person’s stable attributes, such as when a person’s desire is inferred from her presumed temperament. More typically, the dependence goes in the opposite direction, such that people build stable attribute inferences from more specific mental state inferences. For example, when people observe a behavior and infer that the agent’s motive is to help another person, they are more likely to infer that the agent is generous and helpful than when they either infer a different motive or cannot figure out the motive at all (Read et al., 1990; Fein, 1996; Reeder et al., 2004). Hoffman et al. (1981) also showed that people who process behavioral information in goal-based categories remember that information better than those who form impressions of the person in trait-based categories. We also know from the developmental literature that children learn to infer traits and use them to explain behavior after they learn to ascribe and use mental states (Yuill & Pearson, 1998). The primacy of mental state inferences over trait inferences is even more directly reflected in recent data from my lab. Holbrook (2006) found that people who watch short video clips of a variety of behaviors are far more likely to make goal inferences than trait inferences, and their goal inferences are considerably faster.

Universality

Folk conceptions of personality have been most thoroughly studied within the framework of lexical studies, with recent work beginning to explore the stability of personality constructs across many languages and cultures (Saucier & Goldberg, 2001). A few dimensions emerge as remarkably stable (e.g., dynamism and social propriety), but variation increases once abstract factors break into more specific
facets. In addition, some cultures or languages put more emphasis on some dimensions than on others. However, cultural differences in the frequency of using certain trait dimensions do not cast doubt on the universal human proclivity to infer stable attributes in the first place. Previously reported claims of a weaker proclivity in Eastern cultures have not stood up to empirical tests (Choi et al., 1999; Krull et al., 1999). And even though individual differences appear to exist within and across cultures in the conviction of how stable certain traits, abilities and skills really are (Dweck et al., 1993), no extant research suggests that some cultures have no conception of personality – are radical situationists, for example.

CONCLUSION

To survive in complex social communities humans have evolved, and develop from birth on, a remarkable set of psychological tools that range from fundamental concepts to simple processes and complex inferences. Social living demands the use of all of these tools, and they in turn enable communication, relationships, and culture. The current evidence base suggests considerable universality of these social-cognitive tools, though research has also documented variations in their relative use. No less universal than this bright side of social cognition is the dark side that includes stereotyping, prejudice, and hostile intergroup perception. However, to triumph over this darker side, humans can only nurture their capacities to be cognitively engaged and emotionally moved by others, to appreciate and simulate their distinct mental states, and infer as best they can their values and character. That, and nothing less, makes us human.

REFERENCES


