Discerning intentions in dynamic human action

Dare A. Baldwin and Jodie A. Baird

When we observe others in motion, we usually care little about the surface behaviors they exhibit. What matters are their underlying intentions. Judgments about intentions and intentionality dictate how we understand and remember others' actions, how we respond, and what we predict about their future action. A generative knowledge system underlies our skill at discerning intentions, enabling us to comprehend intentions even when action is novel and unfolds in complex ways over time. Recent work spanning many disciplines illuminates some of the processes involved in intention detection. We review these developments and articulate a set of questions cutting across current theoretical dividing lines.

Dare A. Baldwin*
Dept of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227, USA.
*e-mail: baldwin@darkwing.uoregon.edu

Jodie A. Baird
Institute of Child Study, OISE/University of Toronto, 45 Walmer Road, Toronto, ON, Canada M5R 2X2.

A startled mother admonished her four-year-old son for crashing into his younger sibling, at which the young perpetrator cleverly volunteered: 'It was an accident!' When faced with his mother's disgusted disbelief – such blatantly purposeful crashing was clearly no accident – the preschooler asked with genuine perplexity, 'How do *you* know? *I* was the one doing it!' This astute question is the focus of our review: how *do* we discern others' intentions when motions in the world are all that is directly available to us?

Increasingly, the construct of intention is being granted a central position within accounts of human cognitive functioning. We begin our review by defining such interdisciplinary convergence on intention as an important topic of investigation within cognitive science. We then turn to recent efforts to understand just how we might identify others' intentions as we observe them in action. Research investigating intention detection in infants, young children, and other animals provides an important source of ideas in this research enterprise. Not surprisingly, distinct, and in some cases conflicting, theoretical orientations guide researchers' current thinking and experimentation. We attempt to characterize this diversity, but then move beyond theoretical divisions, focusing instead on fundamental questions about intention detection that all accounts will need to address.

Wherefore intentions?

Imagine for a moment failing to grasp the idea that people act in large part to fulfill intentions arising from their beliefs and desires. As you watch people move about the world, you register their motions yet lack a sense of any purpose to these motions. This means that for you, a doctor's diverse actions of offering advice, administering an injection, and performing out-patient surgery are as different as tooth-brushing is from driving a car, because an intention to heal is all that makes these distinctly dissimilar movement patterns cohere conceptually. For this reason, you probably won't see much value in doctors. You also won't see much value in literature, theatre or conversation. After all, what sense can one make of *Oedipus Rex* or *Romeo and Juliet* without a notion of intentions that are offensive versus honorable, and intentions that are thwarted, misguided or misunderstood? What sense can one make of irony, in which a speaker often intends to communicate the opposite of what is said? Indeed, most of what you observe about people and how they engage the world around them will remain utterly opaque. Sadly, some individuals – those with autism and acquired right-hemisphere damage – appear to suffer just these kinds of difficulties, to varying levels of severity[1,2]. Such deficits produce profound disruptions in their daily lives and in the quality of their social relationships.

Our everyday, common-sense ability to interpret and predict others' behavior hinges crucially on judgments about the intentionality of others' actions – whether they act purposefully (with intent) or not – as well as judgments about the specific content of the intentions guiding others' actions – what they intend in undertaking a given action. Until recently, direct investigation of these skills has been rare (but see Refs 3–5 for notable exceptions). One obstacle to such investigation was a deep-seated controversy within the cognitive sciences regarding the usefulness of any account framed with respect to mental states such as intentions. The debate concerned the feasibility and coherence of viewing the mind as possessing states that actually represent content. In question was whether mental states genuinely possess the quality of being *about* things in the world[6–10]. Confusion on this most fundamental of points seriously undermined discussion about intentions and intentionality. Probing the nature of the content of people's intentions and how we might discern such content is obviously difficult if influential figures in the field dismiss the very notion of such content in the first place.

In fact, this representational controversy still rages unresolved, but a growing consensus has emerged in the last two decades that intentions and intentionality – and related mental constructs such as belief and desire – deserve serious investigation

  PII: S1364-6613(00)01615-6

regardless. Our everyday notions of intention and intentionality may turn out to be invalid as characterizations of the genuine contents or processes of the mind/brain[11–13]. Nevertheless, understanding how we operate with these folk notions will be central to accounting for human cognition, because they serve as basic organizing principles guiding how we interpret one another[1,14–20].

Intention and intentionality in action processing

When we read or hear stories about others' actions, our understanding of, and memory for, those actions centers primarily on ideas we've constructed about the actors' motives and goals[21–23]. Similarly, as we witness others in action, it is almost certain that we don't encode the full detail of their motions in space; we probably encode our *interpretation* of those motions in terms of the actors' goals and intentions. In this sense, inferences about intentionality and the content of others' intentions provide the 'gist' or meaning of their actions; it is this constructed understanding that we encode and later retrieve.

Recent investigation within psychology begins to reveal an impressive orderliness in the way people conceive of, and operate with, ideas about intentions and intentionality. Malle and Knobe[24], for example, found that adults given action vignettes such as '*Anne is sweating*' versus '*Anne applauded the musicians*' generally agree about whether such actions are intentional or unintentional, and they consistently base these intentionality judgments on five factors: desire, belief, intention, awareness and skill[24]. In particular, adults tend to judge another's action as arising from an intention only if they have reason to believe the individual desires a goal and believes s/he can attain that goal. Given evidence for intention, adults will go on to regard the action as genuinely intentional if there is also reason to believe that the individual has the skill to perform actions relevant to attaining the goal, and is aware of engaging in those actions.

When asked to reflect on the intentionality of others' actions, children as young as 4–5 years of age display a skeletal version of the same notion of intentionality with which adults operate[25,26]. Even three-year-olds' judgments of others' intentionality are often appropriate in such verbal tasks, at least when intentions and desires coincide[27–29].

Basic skills for processing action lie at the heart of skill at discerning intentions. Everyday goal-oriented action sequences are generally continuous, with few if any pauses to mark meaningful boundaries between one action and the next. This means that observers of intentional action face, first and foremost, a challenging parsing problem – they must discover units within the continuous action stream that are relevant to discerning intentions. Adults display consummate skill at such intention-relevant segmentation of the behavior stream. On viewing continuous action sequences, adults agree about where boundaries separating distinct actions lie[30,31], and their judgments about such boundaries coincide with their understanding of the intentions that the actors are undertaking[32]. Moreover, boundaries between intentional actions emerge as psychologically salient in adult's recall of continuous action sequences[32]. For example, adults are better able to recall the locations of tones sounded during an action sequence when those tones coincide with the actor's completion of intentions than when tones occur in the midst of an intentional act. Finally, recent evidence indicates even infants as young as 10–11 months can parse at least some kinds of continuous everyday action along intention boundaries (see Box 1).

Adults also appear to process continuous action streams in terms of hierarchical relations that link smaller-level intentions (e.g. in a kitchen cleaning-up scenario: intending to grasp a dish, turn on the water, pass the dish under the water) with intentions at higher levels (intending to wash a dish or clean a kitchen)[30,33,34]. Furthermore, adults reliably identify certain actions at the more fine-grained level as especially crucial or defining of intentions at the higher level; for instance, the action of scrubbing a dish with a brush is more of a crux for completing the intention to wash a dish than is the equally necessary but less central prior action of turning on the water[35].

Intentions shape action processing

Ideas about others' intentions not only arise as a result of action processing; they frequently shape such processing from the outset. Zadny and Gerard found, for instance, that adults show selective recall for portions of another's action that bear on the intentions they previously ascribed to the actor[36]. In addition, they found that such selectivity operates during adults' processing of the action – at the time action is actually witnessed – rather than merely retrospectively. Similarly, Dodge and his colleagues have shown that school-age children are more inclined to attribute hostile intent when action is ambiguous if they previously have observed hostile intentions in the actor[37].

Our judgments about intentions and intentionality dictate what we do with the behaviors we witness. This is true from as early as the second year of life. Infants of 14–15 months are inclined to imitate behaviors that others perform intentionally, but are not so inclined for apparently unintentional behaviors[38]. Likewise, infants this same age can, in at least some cases, discern the intention motivating someone's failed action, and they will go on to re-enact a successful version of that same action[39]. Clearly, infants in their second year track and encode intentions rather than the surface patterns of motion actors produce.

Box 1. Infants' parsing of dynamic human action

When observing others in action, do infants experience an incomprehensible flow of movement, or are they skilled at finding relevant structure in the behavior stream? To investigate this question, we tested 10–11-month-olds' ability to detect disruptions to the structure inherent in intentional action[a].

Infants in this study viewed short (*circa* 4 s) sequences of continuous everyday action. In one sequence, for example, a woman notices a towel on the floor, reaches for and grasps it, and places it on a nearby rack. In the familiarization phase of the procedure, infants viewed the same action sequence repeatedly across several trials. In the subsequent test phase, infants viewed two different versions of the original action sequence, this time with still-frame pauses introduced at certain points in the course of action. The 'completing' test version highlighted the boundaries between intentions with a pause occurring just as the actor completed an intention (e.g. as she finished grasping the towel). The 'interrupting' test version disrupted the structure of the action sequence with a pause occurring prior to the completion of an intention (e.g. in the midst of her reach for the towel). If infants are sensitive to structure inherent in intentional action, they should look longer at the interrupting test version (which violates this structure) than at the completing version (which preserves this structure).

Infants' looking decreased significantly from the first to the fourth familiarization trial, indicating that they indeed processed the action sequence during this phase. Relative to the last familiarization trial, infants demonstrated renewed interest – that is, longer looking times – only in response to the interrupting test version; they showed no such renewed interest in response to the completing test version. These findings indicate that the interrupting test videos, but not the completing test videos, violated infants' expectations. A second study clarified that, in the absence of familiarization, infants did not have a starting preference for the interrupting test videos over the completing test videos. Together, the findings from this research suggest that, at least by 10–11 months, infants are sensitive to disruptions of action units defined by the completion of intentions.

Reference

a Baldwin, D.A. *et al*. Infants parse dynamic action. *Child Dev.* (in press)

Moreover, by the end of the first year, infants readily distinguish between agents – objects in possession of intentions – and inanimates, using cues such as the presence of facial features and/or contingent behavior to drive the distinction. Infants respond to objects they judge to be agentive versus non-agentive in markedly different ways, even when, on the surface, these objects display identical patterns of motion[40,41].

Intentions guide inferences in children's word learning

Ideas about intentions also guide infants' inferences about the appropriate application of more specialized forms of action, such as linguistic and emotional messages. For example, on hearing a novel word or an emotional message, infants of 12–18 months of age actively check the speaker's face, and rely on cues such as gaze direction, body posture, gestures and the like, to determine the object to which the speaker intends to refer[42–45]. Infants 18 months and older appropriately resist linking a novel word with a novel object they are attending to at the time they hear the word if the speaker fails to show signs that he or she is intending to talk about that object[46]. Furthermore, infants are flexible in their use of intention cues to guide understanding of others' messages. When gaze direction is uninformative, for example, they can use facial expression and/or other action cues to disambiguate the speaker's intended referent[47].

In summary, from infancy onwards we readily process action in intentional terms, despite the complexity of the behavioral stream we actually witness. We spontaneously segment action along intention boundaries, make systematic judgments about actors' intentionality, and use judgments about the nature of the actor's specific intentions to guide attendant observations, inferences, and our own subsequent action. These intention-discernment skills boost our capacity for knowledge acquisition[48,49]. By focusing selectively on portions of motion relevant to identifying actor's intentions, for example, we can quickly learn how to do things we previously could not do. In this way, humans reap special cognitive benefits from the social milieu, making possible a knowledge base and a level of cultural complexity unparalleled in any other species[49,50]. Together, the present findings point to the existence of a powerful and early-emerging cognitive system for discerning intentions. We now turn to ideas about what the nature of this system might be.

Discerning intentions: how?

In answer, two clearly defined and opposing positions have been on offer. One approach contends that we divine intentions and intentionality *directly* as we observe the flow of motion others produce when enacting intentions. Asch favored just such a direct perception account, stemming from his orientation towards Gestalt principles, and others since have echoed these ideas[1,3,51,52]. On this view, intentions and intentionality arise as invariants within the structure of action; invariants our brains are designed to detect.

A quite different approach to the 'how' of discerning intentions highlights the role of inferential mechanisms[32,53–56]. Judgments about intentions and intentionality are thought to be constructed using hard-won knowledge about human behavior and situational context, and specific information about qualities of individual agents, all serving, together with action information, as raw materials recruited in constructing such inferential judgments.

Surprisingly, no serious debate between these contending accounts has yet transpired (but see Ref. 55 for some discussion). In fact, there is little in the way of actual data to enliven such a contest. Moving directly to the issues seems fruitful in such a case. These boil down to four questions: (1) What kind of information about intentions and intentionality is actually available in the surface flow of agents' motions?; (2) Which aspects of this structure can and do people actually detect?; (3) What kinds of additional information, if any, might be needed to account for people's judgments about the intentional content and the purposiveness of others' actions?; (4) How are skills for discerning intentions and intentionality acquired in human development? In constructing answers to such questions, an approach that integrates across the traditional divide gains momentum.

Structure within intentional action

Some years back, McArthur and Baron expressed a plea that those of us concerned with explaining social judgment-making attend to the kinds of structure contained within our social environments that might inform and influence such judgments[57]. Intentional action seems a promising domain within which to fulfill this plea as it exhibits considerable structure. This is in part because biological and social imperatives reliably elicit certain kinds of intentions, and hence actions, from all of us – those related to food, water, shelter, sleep, reproduction, parenting and play, for example. Also, the nature of the physical world necessitates that we undertake organized, structured action if we are to achieve our goals[58,59]. We must move through space to contact objects towards which we have intentions; as well, physical and temporal constraints shape and limit the kinds of motions we can make.

For all these reasons, predictable patterns emerge again and again in the sequencing and temporal dynamics of bodily motion that are probably unique to intentional action, and that correlate with the initiation and completion of intentional acts. For example, to act intentionally on an inanimate object, we first must locate that object with our sensors (inanimates do not do this, as they usually do not have sensors). We then typically launch our bodies in the direction specified by our sensors, extend our arms, shape our hands to grasp the relevant object, manipulate and ultimately release it (inanimates usually do not do any of this either). All of this typically coincides with a characteristic kind of ballistic trajectory that provides a temporal contour or 'envelope' demarcating one intentional act from the next. Interestingly, even the motions of state-of-the-art robots – inanimates designed to approximate the agentive – are easily distinguishable from those of agents: the temporal dynamics are simply wrong (E. Newport, personal communication). Robots display action discontinuities (e.g. arm out, pause, pincers squeeze) that agents do not and, being made of different stuff, physical principles impinge on them differently, yielding rather different patterns of acceleration, deceleration, and the like.

This is all to say that on a purely structural level – the level of statistical regularities – there is considerable information correlated with intentions that is inherent in the flow of goal-directed action. Differing statistical patterns may even arise for intentional as opposed to unintentional action; for example, in intentional action we first locate with our sensors, and then display directed bodily motion. In unintentional actions, such as stubbing one's toe or slipping on a banana peel, motion comes first and location with sensors after the fact. The details here – exactly what the relevant statistical patterns of motion are, how predictive of intentions they really are, and so forth – is a truly important topic for future investigation.

Structure we can detect

Recently, Saffran and her colleagues demonstrated that 7–8-month-old infants as well as adults are much more skilled at detecting statistical regularities within complex stimuli (auditory/linguistic stimuli in their case) than was previously suspected[60,61]. Infants, for instance, are able to segment a continuous string of syllables into word-like units based solely on non-invariant statistical information and after only thirty seconds of exposure to the regularities in question. Other work demonstrates that infants' structure-detection skills are not restricted to linguistic stimuli; infants can use comparable structural information to segment non-linguistic sequences of tones[62]. We propose that infants as well as adults might be able to recruit analogous statistical capabilities to assist in analyzing dynamic human action, and these abilities might greatly assist their processing of intentional action. Research investigating these possibilities is currently underway in our laboratory.

An ingenious technique created by Blythe *et al.* provides additional plausibility to the idea that structure detection plays a role in our processing of intentional action[63]. They asked adults to move a bug-shaped computer cursor on a screen to enact a variety of different intentions (e.g. intentions to pursue versus evade) towards a second bug-like shape over which they had no control. The patterns of motion induced in the participants' bug were recorded digitally, and then analyzed by several computational models designed to detect and categorize distinct motion structures. The models – using purely structural information – succeeded at categorizing the motions at accuracy levels ranging between 62–82%, on average. These findings are striking because they suggest that judgments about the content of at least some kinds of intentions is strongly supported by structural information – a suggestion that Premack and Premack voiced several years previously[52,64]. Interestingly, however, adult human observers given the same information as the computational models displayed an average accuracy level of only 49%, suggesting some important discrepancies between the way humans and such models utilize the available structural information.

Clearly, we are only at the outset of discovering what structure people can detect in intentional action, but this looks to be a promising direction for investigation. In particular, Blythe *et al*.'s research, as well as related work by Siskind and Thibadeau, points to the value of computational modeling techniques for probing such questions[65,66].

Requirements beyond structure detection

Searle tells us why structural information alone will never suffice to account for our judgments about the content of others' intentions[67]:

> 'If I am going for a walk to Hyde Park, there are any number of things that are happening in the course of my walk, but their descriptions do not describe my intentional actions, because in acting, what I am doing depends in large part on what I think I am doing. So for example, I am also moving in the general direction of Patagonia, shaking the hair on my head up and down, wearing out my shoes and moving a lot of air molecules. However, none of these other descriptions seems to get at what is essential about this action, as the action it is.' (Ref. 67, p. 58.)

In other words, the surface flow of motion people produce in most, if not all, cases is consistent with a multitude of different intentions. Thus when observing others in action, we rely on other sources of information – knowledge about human behavior in general, specific knowledge about the particular individual involved, knowledge about the situation – to help disambiguate which among the many candidate intentions is relevant in any given case. Some of the research reported earlier demonstrates such top-down, or inferential, effects in operation; for example, the work by Zadny and Gerard showing that people's expectations about an agent's intentions lead them to encode and remember portions of the action consistent with the expected intentions[36].

Social psychological research indirectly provides a wealth of potentially relevant information about the kinds of processes operative in guiding judgments of intention and intentionality, although as yet there has been little attempt to address these issues directly (but see Malle[68]). Adopting a developmental stance may be especially enlightening by providing information about the kinds of errors in judgment about intentions and intentionality children are prone to in the absence of top-down knowledge about the individual, the situation, or humans in general. One of us can recall such an error: her toddler's furious reaction when she agreed to help extricate him from yet another tight spot, but then first moved in the opposite direction to put down a pan of boiling soup. This child seemed not yet to possess the relevant knowledge that mothers typically protect their children from scalding fluids and thus interpreted a movement in the wrong direction as a straightforward failure to enact the desired intention. This anecdote also highlights some of the complexities of intentional action that children must learn to cope with on inferential grounds. For one, intentional actions are frequently interrupted (others interrupt, needed instruments are missing or broken, etc.); one must somehow recognize coherence across disparate segments. In addition, children must come to recognize the fact noted earlier that intentions are not isomorphic with action: one can carry out a variety of actions to fulfill a given intention, and a given action is consistent with a variety of possible intentions. That children have difficulty with this is apparent to anyone who attempts to convince their preschooler that nurses who carry out immunization injections do not actually intend to hurt them. In the first direct investigation of these issues, Baird recently found that even four-year-olds sometimes have difficulty appreciating that one and the same action can arise from distinct intentions (unpublished doctoral dissertation).

There is another, deeper, sense in which judgments about the content of others' intentions are fundamentally bound up with world knowledge. Consider kitchen clean-up intentions, once again. Without doubt, infants observing such activity are unable to grasp the content of the tidier's intentions in anything like the way we as adult observers might understand these actions (intentions to wash dishes, stow ice cream in the freezer, hang towels to dry, for example). Infants lack knowledge of all sorts of minutia about the world (the need to sanitize dirty dishes and the role that water plays in such a process, the function of refrigerators, and the evaporative properties of water) that infuse the intentions we impute. Even as adults, our different experiential histories inevitably yield differences in the knowledge

Questions for future research

- What implications does a propensity to construe others' behavior with respect to intentions and intentionality have for other aspects of our cognitive functioning?
- What kinds of disruptions occur in detecting others' intentions, and what implications do such disruptions have for other aspects of cognitive functioning?
- In what ways are skills for discerning intentions and intentionality universal, and in what ways might these skills be shaped by culture and experience? Is it useful to speak of a 'native intentional system' in the same sense that we speak of a native language?
- What are the neurophysiological systems that support intention detection, and how specialized are they?
- Do maturational changes in brain plasticity limit the ability to acquire, later in life, culturally-specific skills for discerning intentions?

we possess. These differences lend unique nuances to our judgments of others' intentions, even when we witness identical action information with the same starting expectations.

A generative knowledge system

The upshot is that discerning intentions is a complex enterprise; it is knowledge driven as well as rooted in structure detection. In fact, our skills in this arena point to the operation of a generative knowledge system, one that is probably just as rich and complex as the generative system underlying language.

We choose the phrase 'generative knowledge system' advisedly. Adults, and even infants, readily discern intentions within novel streams of action. This capacity to deal with novel action sequences is one sense in which intentional understanding is clearly generative. And intentional understanding is functionally a knowledge *system* as well. As noted earlier, we can discern intentions in action at multiple levels of analysis – the level of small actions (e.g. grasp a dish, place a dish on the counter), higher-level tasks (e.g. wash a dish, hang a towel), global intentions (e.g. clean the kitchen), daily goals (e.g. prepare for a mother-in-law's visit), and so forth. At each of these levels, judgments of intentions and intentionality are structured according to a unique set of principles. That is, what makes a small action well-formed (e.g. appropriate hand shape for grasping the relevant goal object, rate of motion, etc.) is very different from what underlies well-formedness at any of the higher levels. Adult-like skill at discerning intentions and intentionality requires fluent action analysis at all these levels, just as adult-like linguistic fluency requires principled knowledge at many levels of analysis within language (e.g. phonology, syntax, semantics, pragmatics).

Ontogeny

If discerning intentions occurs by the operation of a generative knowledge system, then how is this system acquired in the course of human development? And is it a uniquely human development (see Box 2)? As in the language domain, some have suggested that human infants possess neural structures innately equipped with dedicated micro-circuitry enabling them to detect intentionality[1,64] and even some kinds of intentional content, such as intentions to approach versus escape[52]. The notion here is that witnessing certain kinds of movement patterns – self-initiated motion, for example – simply triggers a percept of intentionality in the infant. We urge restraint with regard to such hypotheses – restraint from dismissing these suggestions out of hand, and comparable restraint from rushing to embrace them. Not enough is known yet about the domain of intention detection to support intelligent guesswork about the form innate knowledge or processes might take.

At the same time, it seems safe to conclude in general terms that acquisition in the domain of intentional understanding must be heavily constrained. The rapid emergence of this knowledge system in children's development, and the fact that we achieve agreement in our judgments about intentions and intentionality despite the radical underdetermination of the information provided in the flow of action itself both point to this conclusion. Committing to constraints is not, however, committing to a radical nativist agenda. It is crucial to remember that constraints on learning can take many different forms, the bulk of which impact development only via complex organism–environment interactions[69].

As we have suggested here and elsewhere, structure-detection skills infants possess may serve as one kind of constraint facilitating the acquisition of intentional understanding[32]. If infants can deploy structure-detection skills in the domain of motion processing that are at all comparable to those demonstrated in the language domain, such skills may enable them to get started at making relevant discriminations (e.g. distinguishing at least some intentional and unintentional acts) and to segment continuous action along intention boundaries. Learning algorithms, such as those proposed by Gentner and Medina[70] as well as Holyoak and Thagard[71], that induce progressively richer, higher levels of structure from such initial perceptual distinctions might then help to explain the blossoming of genuine intentional understanding. Research into the precise path such conceptual development takes is only beginning, but ingenious studies by Woodward and colleagues[72,73], Wellman and Phillips[74], and Gergely and colleagues[75,76] clarify that foundational skills are set in place during the first year of life, including skills for identifying appropriate actions as (1) directed towards a specific object (a 'goal') and (2) predictive of other actions and emotional displays.

Box 2. Construing others' intentions: a uniquely human propensity?

In 1978, Premack and Woodruff sparked intense interest in this question with an ingenious series of studies investigating chimpanzees' ability to discern the intention of someone struggling to solve a problem[a]. After viewing such action sequences, the chimpanzee in their studies displayed skill in selecting a photograph that depicted the solution to the problem from among a group of distracter photographs. However, a host of important questions remained as to whether the chimpanzee's success in the Premack and Woodruff task stemmed from genuine appreciation of mental states such as intention. For example, such success might have traded on an understanding of the causal structure of the problem without any recognition of the actor's intentions.

Premack and Woodruff's seminal work inspired many researchers to investigate further this and related questions about non-human primates' ability to reason about others' mental states, including their intentions, beliefs, knowledge, emotions and visual experiences. The answer to whether non-human primates possess such skills is currently in doubt. On the one hand, chimpanzees display behaviors that hint at an appreciation of mental states. For instance, recent findings from Call and Tomasello suggest that orangutans and chimpanzees can distinguish intentional from accidental actions in at least some cases[b]. Another such potentially sophisticated behavior – chimpanzees' ability to follow and utilize gaze-direction information – has received considerable attention from researchers[c–g]. Such gaze following suggests, at least on the face of it, that chimpanzees recognize gaze direction as a cue to others' visual experience and intentions. At the same time, many ingenious and rigorously controlled experiments strongly suggest that they lack a deeper, conceptual appreciation of gaze as a window on others' visual experience[d,e,h–k].

Based on the pattern of findings across a broad range of studies, Povinelli and colleagues have proposed the 'reinterpretation hypothesis'[l]: that non-human primates as well as human infants possess skills for analyzing others' *behavior* and using aspects of behavior to predict others' actions, but that only humans come to possess the ability to interpret behavior in genuinely mentalistic terms. That is, perhaps humans have evolved a more advanced mechanism for reasoning about ('reinterpreting') behavior (e.g. gaze direction) as indicative of mental experience (e.g. visual experience, intentions, knowledge, beliefs and the like) – a mechanism that chimpanzees and other non-human primates do not possess. This hypothesis resolves the apparent tension between (a) the striking similarities in behavior across primate species (including humans), and (b) the failures of understanding shown by all primate species other than our own when mentalistic reasoning is carefully probed.

All in all, there seems as yet to be no clear basis on which to attribute mentalistic reasoning, including genuine inferences about visual experiences or intentions, to even our closest phylogenetic relatives. At the same time, there seems every reason to suspect that the great apes, and most likely many other species, share some of the action-processing skills that humans possess. Teasing out the specifics remains a fascinating challenge to future research.

References

a Premack, D. and Woodruff, G. (1978) Does the chimpanzee have a theory of mind? *Beh. Brain Sci.* 4, 515–526

b Call, J. and Tomasello, M. (1998) Distinguishing intentional from accidental actions in orangutans (*Pongo pygmaeus*), chimpanzees (*Pan troglodytes*), and human children (*Homo sapiens*). *J. Comp. Psychol.* 112, 192–206

c Emery, N. *et al.* (1997) Gaze following and joint attention in rhesus monkeys (*Macaca mulatta*), *J. Comp. Psychol.* 111, 286–293

d Reaux, J.E. *et al.* (1999) A longitudinal investigation of chimpanzee's understanding of visual perception. *Child Dev.* 70, 275–290

e Povinelli, D.J. and Eddy, T.J. (1996) Chimpanzees: joint visual attention. *Psychol. Sci.* 7, 129–135

f Tomasello, M. *et al.* (1998) Five primate species follow the visual gaze of conspecifics. *Anim. Behav.* 55, 1063–1069

g Hare, B. *et al.* (2000) Chimpanzees know what conspecifics do and do not see. *Anim. Behav.* 59, 1–15

h Call, J. *et al.* (1998) Chimpanzee gaze following in an object choice task. *Anim. Cognit.* 1, 89–100

i Povinelli, D.J. and Eddy, T.J. (1996) What young chimpanzees know about seeing. *Monogr. Soc. Res. Child Dev.* 61, 1–152

j Povinelli, D.J. *et al.* (1999) Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *Br. J. Dev. Psychol.* 17, 37–60

k Tomasello, M. *et al.* (1997) The comprehension of novel communicative signs by apes and human children. *Child Dev.* 68, 1067–1081

l Povinelli, D.J. and Giambrone, S. Reasoning about beliefs: a human specialization? *Child Dev.* (in press)

Conclusion

Many times each day we make judgments about others' intentions as we witness them in action; this is the basis on which we understand, predict, and attempt to influence one another's actions. We discern intentions in others' action via a complex combination of structure-detection skills and relevant knowledge. Our knowledge in this domain, although largely implicit, is systematic, multi-faceted and generative. Beyond these broad outlines much remains unknown about our intention-discernment skills, including the mechanisms that enable their rapid acquisition in human development. We view these issues as among the most challenging and exciting frontiers of investigation in the field of cognitive science.

Acknowledgements

This manuscript was prepared while the first author was a Fellow at the Center for Advanced Study in the Behavioral Sciences; she is grateful for the financial support provided by the William T. Grant Foundation, award #95167795. Writing of this manuscript was also supported by a John Merck Scholars Award and a National Science Foundation New Young Investigator grant to the first author, as well as by a NICHD National Research Service Award to the second author. Our sincere thanks to Diego Fernandez-Duque, Dedre Gentner, Alison Gopnik, Bertram Malle, Ellen Markman, Elijah Milgram, Lou Moses, Elissa Newport, Daniel Povinelli, Amanda Woodward, and the anonymous reviewers for thought-provoking and helpful discussion.

References

1 Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and Theory of Mind*, MIT Press
2 Sabbagh, M.A. (1999) Communicative intentions and language: evidence from right-hemisphere damage and autism. *Brain Lang.* 70, 29–69
3 Asch, S.E. (1952) *Social Psychology*, Prentice-Hall
4 Grice, P. (1957) Meaning. *Philos. Rev.*, 66, 377–388
5 Heider, F. (1958) *The Psychology of Interpersonal Relations*, John Wiley & Sons
6 Brentano, F. (1973/1874) *Psychology from an Empirical Standpoint* (Kraus, O. *et al.*, eds), Routledge
7 Dennett, D.C. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*, Harvester Press
8 Dretske, F. (1981) *Knowledge and the Flow of Information*, MIT Press
9 Fodor, J.A. (1981) *Representations: Philosophical Essays on the Foundations of Cognitive Science*, Harvard University Press
10 Searle, J.R. (1983) *Intentionality: An Essay in Philosophy of Mind*, Cambridge University Press
11 Churchland, P.S. (1986) *Neurophilosophy: Toward a Unified Science of the Mind–Brain*, MIT Press
12 Dennett, D.C. (1987) *The Intentional Stance*, MIT Press
13 Stich, S.P. (1983) *From Folk Psychology to Cognitive Science: The Case Against Belief*. MIT Press
14 Anscombe, G.E.M. (1958) *Intention*, Cornell University Press
15 Bogdan, R.J. (1994) *Grounds for Cognition: How Goal-Guided Behavior Shapes the Mind*, Erlbaum
16 Bratman, M.E. (1987) *Intention, Plans, and Practical Reason*, Harvard University Press
17 Goldman, A. (1970) *A Theory of Human Action*, Prentice-Hall
18 Perner, J. (1991) *Understanding the Representational Mind*, MIT Press
19 Wellman, H. (1990) *Children's Theories of Mind*, MIT Press
20 Zelazo, P. *et al.* (1999) *Theories of mind in action: Development and evolution of social understanding and self-control*, Erlbaum
21 Bower, G.H. and Rinck, M. (1999) Goals as generators of activation in narrative understanding. In *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* (Goldman, S.R. *et al.*, eds), pp. 111–134, Erlbaum
22 Schank, R.C. and Abelson, R.P. (1977) *Scripts, Plans, Goals, Understanding*, Erlbaum
23 Trabasso, T. and Nickels, M. (1992) The development of goal plans of action in the narration of a picture story. *Discuss. Proc.* 15, 249–275
24 Malle, B.F. and Knobe, J. (1997) The folk concept of intentionality. *J. Exp. Soc. Psychol.* 33, 101–121
25 Feinfield, K. *et al.* (1999) Young children's understanding of intention. *Cog. Dev.* 14, 463–486
26 Shultz, T.R. *et al.* (1980) Development of the ability to distinguish intended actions from mistakes, reflexes, and passive movements. *Br. J. Soc. Clin. Psychol.* 49, 736–743
27 Astington, J.W. (1991) Intention in the child's theory of mind. In *Children's Theories of Mind* (Frye, D. and Moore, C., eds), pp. 157–172, Erlbaum
28 Shultz, T.R. and Wells, D. (1985) Judging the intentionality of action-outcomes. *Devel. Psychol.* 21, 83–89
29 Yuill, N. (1984) Young children's coordination of motive and outcome in judgements of satisfaction and morality. *Br. J. Dev. Psychol.* 2, 73–81
30 Newtson, D. (1973) Attribution and the unit of perception of ongoing behavior. *J. Pers. Soc. Psychol.* 28, 28–38
31 Newtson, D. and Engquist, G. (1976) The perceptual organization of ongoing behavior. *J. Exp. Soc. Psychol.* 12, 436–450
32 Baird, J.A. and Baldwin, D.A. Making sense of human behavior: action parsing and intentional inference. In *Understanding Intentions and Intentionality: Foundations of Social Cognition* (Malle, B.F. *et al.*, eds), MIT Press (in press)
33 Zacks, J. and Tversky, B., Event structure in perception and cognition. *Psychol. Bull.* (in press)
34 Newtson, D. (1976) Foundations of attribution: the perception of ongoing behavior. In *New Directions in Attribution Research* (Harvey, J. *et al.*, eds), pp. 223–247, Erlbaum
35 Reed, E.S. *et al.* (1992) Visually based descriptions of an everyday action, *Ecological Psychol.* 4, 129–152
36 Zadny, J. and Gerard, H.B. (1974) Attributed intentions and informational selectivity. *J. Exp. Soc. Psychol.* 10, 34–52
37 Dodge, K.A. *et al.* (1984) The assessment of intention-cue detection skills in children: implications for developmental psychopathology. *Child Dev.* 55, 163–173
38 Carpenter, M. *et al.* (1998) Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behav. Dev.* 21, 315–330
39 Meltzoff, A. and Brooks, R. 'Like Me' as a building block for understanding other minds: Bodily acts, attention, and intention. Malle, B.F. *et al.* eds, *Intentions and Intentionality: Foundations of Social Cognition*, MIT Press (in press)
40 Johnson, S. *et al.* (1998) Whose gaze will infants follow? The elicitation of gaze-following in 12-month-old infants. *Devel. Sci.*, 1, 233–238
41 Johnson, S.C. (2000) The recognition of mentalistic agents in infancy. *Trends in Cog. Sci.*, 4, 22–28
42 Baldwin, D.A. (1991) Infants' contribution to the achievement of joint reference. *Child Dev.* 62, 875–890
43 Baldwin, D.A. (1993) Infants' ability to consult the speaker for clues to word reference. *J. Child Lang.* 20, 395–418
44 Hollich, G.J. *et al.* Breaking the language barrier: an emergentist coalition model for the origins of word learning, *Monogr. Soc. Res. Child Dev.* (in press)
45 Moses, L.J. *et al.* Evidence for referential understanding in the emotions domain at 12 and 18 months. *Child Dev.* (in press)
46 Baldwin, D.A. *et al.* (1997) Infants' reliance on a social criterion for establishing word-object relations. *Child Dev.* 67, 3135–3153
47 Tomasello, M. *et al.* (1996) Eighteen-month-old children learn words in non-ostensive contexts. *J. Child Lang.*, 23, 157–176
48 Baldwin, D.A. (2000) Interpersonal understanding fuels knowledge acquisition. *Curr. Dir. Psychol. Sci.* 9, 40–45
49 Tomasello, M. (1999) *Cultural Origins of Human Cognition*, Harvard University Press
50 Baldwin, D.A. and Moses, L.J. (1996) The ontogeny of social information-gathering. *Child Dev*. 67, 1915–1939
51 Newtson, D. (1993) The dynamics of action and interaction. In *A Dynamic Systems Approach to Development: Applications* (Smith, L.B. and Thelen, E., eds), pp. 241–264, MIT Press
52 Premack, D. and Premack, A.J. (1995) Intention as psychological cause. In *Causal Cognition: A Multidisciplinary Debate* (Sperber, D. *et al.*, eds), pp. 185–199, Clarendon Press
53 Baldwin, D.A. and Baird, J.A. (1999) Action analysis: a gateway to intentional inference. In *Early Social Cognition* (Rochat, P., ed.), pp. 215–240, Erlbaum
54 Dittrich, W.H. and Lea, S.E.G. (1994) Visual perception of intentional motion. *Perception* 23, 253–268
55 Malle, B. F. *et al. Intentions and Intentionality: Foundations of Social Cognition*, MIT Press (in press)
56 Piaget, J. (1932) *The Moral Judgment of the Child*, Kegan Paul
57 McArthur, L.Z. and Baron, R.M. (1983) Toward an ecological theory of social perception. *Psychol. Rev.* 90, 215–238
58 Gallistel, C.R. (1980) *The Organization of Action: A New Synthesis*, Erlbaum
59 Miller, G.A. *et al.* (1960) *Plans, and the Structure of Behavior*, Henry Holt
60 Aslin, R.N. *et al.* (1998) Computation of conditional probability statistics by 8-month-old infants. *Psychol. Sci.* 9, 321–324
61 Saffran, J.R. *et al.* (1996) Statistical learning by 8-month-old infants. *Science* 274, 1926–1928
62 Saffran, J.R. *et al.* (1999) Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52
63 Blythe, P.W. *et al.* (1999) How motion reveals intention: categorizing social interactions. In *Simple Heuristics That Make Us Smart* (Gigerenzer, G. and Todd, P.M., eds), pp. 257–285, Oxford University Press
64 Premack, D. (1990) The infant's theory of self-propelled objects. *Cognition* 36, 1–16
65 Siskind, J.M. (1994) Grounding language in perception. *Artif. Intell. Rev.* 8, 371–391
66 Thibadeau, R. (1986) Artificial perception of actions. *Cognit. Sci.* 10, 117–149
67 Searle, J. (1984) *Minds, Brains, and Science*, Harvard University Press
68 Malle, B.F. (1999) How people explain behavior: a new theoretical framework. *Pers. Soc. Psychol. Rev.* 3, 23–48
69 Elman, J.L. *et al.* (1996) *Rethinking Innateness: A Connectionist Perspective on Development*, MIT Press
70 Gentner, D. and Medina, J. (1998) Similarity and the development of rules, *Cognition* 65, 263–297
71 Holyoak, K.J. and Thagard, P. (1995) *Mental Leaps: Analogy in Creative Thought*, MIT Press
72 Woodward, A.L. (1998) Infants selectively encode the goal object of an actor's reach. *Cognition* 69,1–34
73 Woodward, A.L. and Somerville, J.A. Twelve-month-olds interpret action in context. *Psychol. Sci.* (in press)
74 Wellman, H.M. and Phillips, A.T. Developing intentional understandings. In B. Malle, *et al.* (eds), *Intentions and Intentionality: Foundations of Social Cognition*, MIT Press (in press)
75 Csibra, G. *et al.* (1999) Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition* 72, 253–284
76 Gergely, G. *et al.* (1995) Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193