In (2002) A. Brook and D. Ross (Eds.), *Daniel Dennett*, Cambridge, UK: Cambridge University Press, pp. 83-116

The Intentional Stance: Developmental and Neurocognitive Perspectives

RICHARD GRIFFIN & SIMON BARON-COHEN

Autism Research Centre
Departments of Experimental Psychology and Psychiatry
University of Cambridge
Downing St, Cambridge, UK, CB2 3EB
Rg238@cam.ac.uk

Acknowledgements: The authors were funded by an MRC Program Grant during the period of this work. The first author was additionally funded by a fellowship from the Cambridge Overseas Trust. We thank Chris Viger, Chris Westbury, and Chris Ashwin for helpful comments on an earlier draft.

Nowhere in the psychological sciences has the philosophy of mind had more influence than on the child development literature generally referred to as children's 'theory of mind.' Developmental journals may seem to be an unlikely place to find Brentano, Frege, and Dennett alongside descriptions of referential opacity and the principle of substitutivity, but it is not at all uncommon in this literature. While the many problems and complexities of the propositional attitude literature are still hotly debated by philosophers, and often ill understood by scientists working in this area, a great deal of empirical progress has already been made. We have Dan Dennett to thank for this extraordinary dialogue between these disciplines.

One of the reasons for Dennett's influence among developmental psychologists and other scientists is his accessible prose. He writes not only for his expert colleagues, but also for those of us working on problems of the mind who haven't grown up with the language of *de re/ de dicto* distinctions, notional worlds, or intensions-with-an-s. Despite his efforts, however, many confusions linger. The ascendance of cognitivism and the computational theory of mind, combined with our strong intuitions about folk psychology, has led many investigators to favourably discuss Dennett's "intentional stance," yet model the competence in ways akin to Fodor's language of thought, despite the apparent incompatibility of these programs. Indeed, scientists are often seen as easy targets by philosophers, who have spent years with their noses deep in these muddy issues. Dennett himself has played the role of 'philosophy police' on many an occasion, but instead of declaring the problems off limits to scientists, he invites participation and promotes informed empirical investigation.

The peer commentary following Premack and Woodruff's (1978) "Does the chimpanzee have a theory of mind?" set the stage and introduced the flavour of the debates surrounding the characterization of systems as 'mindreaders' vs. 'behavior-readers.' Dennett's (1978) commentary laid out some of the difficulties in making this distinction and offered some empirically-friendly suggestions aimed at teasing apart mindreaders from behaviour-readers. A key component absent from Premack and Woodruff's experiments was a measure of false-belief attribution, and moreover, false-belief attribution in a novel situation (to rule out an explanation in terms of experienced regularities and other fodder for the behaviourist's cannon). Dennett illustrated one suggestion, suitable for young children, with a scenario in which Punch had a mistaken belief about the location of Judy. Wimmer and Perner (1983) modified the scenario slightly, and a cottage industry of experiments with young children was born.

The result most investigators found was that, before the ages of four-and-a-half, or so, children do not consistently predict the behavior of someone by taking into account their false-belief. It was a striking finding. Variations on the task abounded, as did replications, and understanding false-belief came to be seen by many as the *sine qua non* of a representational theory of mind. To understand that a mind represents, one must understand that a mind can misrepresent, and that the misrepresentation (false belief) will cause the believer's behaviour. Wellman et al’s (2001) meta-analysis of false-belief tasks underlines just how important this notion has become to the literature, as it includes 571 conditions, even in the face of the strict inclusion criteria imposed by the authors.

The extensive focus on this task is not only due to its status as the mindreading watershed, but also due to the fact the subtlest experimental manipulations will often produce striking differences in young children’s performances. What shouldn’t be missed, however, is that understanding other minds doesn’t suddenly appear to children six months prior to their fifth birthday. Adults have memories of childhood, although it is quite uncommon for those memories to be the momentous discovery that people have minds. It simply doesn’t happen that way as many of our folk psychological mechanisms are already fully operational.

In this essay, we review developments in infancy research and cognitive neuroscience. We follow each selective review with a critical analysis, in an attempt to show how thinking in these fields follows or diverges from Dennett’s influential intentional stance. We close by attempting to incorporate some of these findings into Dennett’s larger program of explaining our kind of mind. First, however, we attempt to clarify some of the differences between the intentional stance, folk psychology, and theory of mind, as we see them.

FOLK PSYCHOLOGY, THE INTENTIONAL STANCE, AND THEORY OF MIND

Folk psychology (FP), the intentional stance (IS), and theory of mind (ToM) are often used interchangeably. While this is most often fine for the empirical researcher, there are subtle but important differences.

(a) Theory of mind is a phrase generally limited to the animal or person’s ability to represent themselves or others as having intentional, content-bearing, representational states (e.g., believing that p, or knowing that q, etc.). So we would say that a child or a chimpanzee has a theory of mind when we want to say that the child or the chimpanzee knows that others have beliefs and desires, for example, which play a causal role in behaviour.

(b) Folk psychology (also sometimes called belief-desire psychology, naïve or intuitive psychology, or commonsense psychology) includes theory of mind, but it also includes emotions, qualitative or phenomenal states, traits, dispositions, and empirical generalizations about behavior (e.g., "People who are overtired are generally irritable”).

(c) The intentional stance is Dennett's take both on how we predict behaviour using intentional constructs (the 'craft') and on what intentional states really are. The terms belief and desire are borrowed from folk psychology but they are given a technical meaning. Dennett considers beliefs and desires logical constructs (*abstracta*) rather than theoretical posits (*illata*) which are assumed to have a physical existence. This latter notion is more akin to folk and even scientific thinking about intentional states. Thus, according to the intentional stance, beliefs and desires are not reducible to brain-states. The theory assumes that the belief-desire profile of the system is holistic, so beliefs/desires can not be attributed in isolation. Instead, new beliefs and desires are predicted from the previous belief-desire profile of the system. Moreover, the theory assumes that the system under analysis functions optimally and rationally, and it is a black box theory, since the physical instantiation of these intentional states is irrelevant to the theory's predictive efficacy. In this way the theory is normative, and shares much with game theory or decision theory. Any system whose behavior can be predicted by the intentional stance is considered an *intentional system*.

In practice, in psychology, all 3 of these terms (ToM, FP, and IS) have been used interchangeably. The difference between attributing a propositional attitude to another and representing (in the brain) that attribution in the *form* of a propositional attitude is sometimes overlooked in the ToM literature, though this difference is precisely what Dennett is trying to illuminate. He calls what we do actually do in folk psychology *the craft* and contrasts this with how we talk about what we do, which he calls *the ideology* (Dennett 1991). For Dennett, the IS as he has laid it out *is* the craft, and our intuitive notion of beliefs and desires as in the head somewhere is no more than false ideology. His technical spin on these terms serves to separate the intentional stance from commonsense psychology. In the two literatures we review in this essay (infancy, and the brain basis of mentalistic ascription) we hope to show how advances in cognitive science are illuminating additional aspects of the craft of folk psychology.

THE DEVELOPMENT OF THE INTENTIONAL STANCE

Although it isn't until children are around 4-years-old that they can predict behavior based on a false belief (in a novel situation) (Wimmer & Perner 1983; Wellman et al. 2001), many researchers are prepared to grant much younger children a theory of mind despite this obvious shortcoming. One reason is that children are using so-called simpler mental state terms such as *want, pretend, know*, and *think*, in quite sophisticated ways, soon after they learn to speak (Barstch and Wellman 1995).

An argument more in the style of philosophy provided by Leslie (1987) is based on the two-year-old's abilities to understand the propositional attitude of pretence. Because having a theory of mind is generally thought of as being able to represent representations (qua representations), Leslie posits a meta- or M-representation system with a 'decoupler' mechanism to serve this function. In order for the child to pretend, her or she must be able to hold simultaneous representations on-line without confusing the two (e.g., the banana is a telephone and the banana is a banana). This ability, according to Leslie, involves the same distinction as that between propositional attitudes and propositional content, as it is the computation of the relation:

> agent (e.g. the child) + an attitude (e.g., pretending) + a primary representation (the object being manipulated – e.g., the phone) + secondary representation (decoupled in nature – e.g., the phone as a banana).

Two-year-old children appear to be capable of drawing pretend consequences from pretend assumptions. For instance, Harris (1993) found that two-and-a-half-year-old children can distinguish between a pretend and a real outcome from a pretend or real assumption (e.g., that chocolate would be wet/dry after having pretend/real tea poured on it). This kind of reasoning, if laid out in as rules consisting of embedded conditionals (e.g., If A, then if B, then C) appears to be the same kind of reasoning necessary to pass the false-belief task (Frye, Zelazo, and Palfai 1995), which the same children won't pass for almost another two years. Moreover, these data show that young children can overcome an apparent bias to reason only about what (they think) is real in the world.

Additional evidence of an early theory of mind comes from a series of studies where eighteen-month-old children are claimed to infer the intentions and goals of an actor who fails in their attempt to carry out an action (e.g., pulling apart miniature dumbbells) (Meltzoff 1995). In these studies, children are more likely to complete the actor's unfulfilled goal than to spontaneously perform those actions; nor do the children complete the failed actions when they are performed by a machine, presumably because the children do not attribute intentions to the machine. Moreover, fourteen- to eighteen-month-old children will imitate an unfulfilled goal if the action is marked linguistically as purposeful (e.g., "Let's put this on here. There we go!") but not if it is marked as accidental (e.g., "Let's put this on here. Whoops!") (Carpenter et al. 1998).

Several recent theories of language development suggest that understanding the intentions of the speaker is a key component in word learning (Tomasello 1999; Bloom 2000). Eighteen-month-olds will not map novel words onto an object they are looking at when they hear a word, but instead map the word onto the object the *speaker* is looking at (Baldwin 1995). Additionally, eighteen-month-olds will map novel verbs ("Can you fep the ball?") onto intentional but not accidental actions, even if they have never seen the completed action. That is, if they witness an experimenter attempting to ‘fep’ something but failing, they infer that ’fep’ refers to the action required to fulfill the actor’s intended goal (Tomasello and Barton 1994).

A number of important behaviors emerge around the child’s first birthday which similarly invite an interpretation of infants as having a simple theory of mind. These include the onset of communicative gesturing (such as declarative pointing), gaze following or social referencing. All of these come under the heading of joint or shared attention. In these situations, infants alternate their gaze between the adult’s eyes and facial emotional expression and an object or event to help determine their course of action or to share information. For instance, a parent’s facial expression can influence whether a twelve-month-old will cross a visual cliff (Sorce et al. 1985) or whether a ten-month old will interact with a stranger (Feinman and Lewis 1983). In all these cases, the argument is that the infant is coding the adult’s mental state of attention to, or emotion about, a state of affairs.

Likewise, Repacholi and Gopnik (1997) found that when fourteen-month-old children are asked, “Can I have some food?”, they will give an experimenter crackers rather than raw broccoli, despite the fact that the experimenter had just expressed interest in the broccoli (e.g., by smiling and saying “Mmm”) and disgust at the crackers (e.g., by looking at the crackers, putting on a ‘disgust’ face, and saying “Yuk!”). Eighteen-month-olds, by constrast, will give the experimenter the broccoli. From this it seems that, by eighteen months , infants can set aside their own desire (for crackers) and recognize the adult’s different desire (for broccoli).

An interesting experiment from Johnson, Slaughter and Carey (1998) teased apart some of the potential cues which will elicit gaze monitoring in children and found that of central importance was contingent interaction. Johnson et al. found that twelve-month-old infants will reliably follow the gaze of a faceless animal-like object provided that the object reacted contingently with them. These children would also follow the gaze if the object had a face alone, but they would not follow the object’s gaze if the object was faceless and did not interact contingently, even if it produced the same self-generated behaviors as in the contingent condition. Infants only a few months old interact contingently with caregivers, and become distressed when this interaction is interrupted (Field et al. 1986). Moreover, if shown a contingently interactive adult on a video monitor, infants will respond with greater positive affect than toward a noncontingent video (Hains et al. 1992). Contingent interaction seems to be one important cue the infant searches for as a sign that the object opposite them is both animate and an agent capable of seeing/attending[1].

Fodor (1987) once quipped that young children get smarter and smarter as experimental techniques improve. Indeed. We now know that infants in the first few months of life have considerable knowledge about the properties of physical objects.

---

[1] Hood (1995) has shown that infants as young as five-months-old use another’s eye direction alone to direct their attention.

For instance, they know that objects remain cohesive and bounded as they move (principle of cohesion), that their motion is continuous and that they continue to exist and move even if occluded from view (principle of continuity), and that objects effect each other only upon contact (principal of contact) (Spelke & Van de Walle 1993; Baillargeon 1995). These core principles constrain reasoning about physical objects and infants show surprise when they are violated. However, the constraints of the third principle -- the principle of contact -- do not hold for *animate* objects. Animate objects move in the absence of contact, as in self-initiated movement, and such movements may indeed be affected by distal events, through communication or perception for instance.

Schlottman and Surian (1999) used a launching event paradigm (Michotte 1963) to test whether infants perceive causation-at-a-distance with animate objects. They followed a method used by Leslie and Keeble (1987), who showed that six-month-olds were capable of perceiving ‘contact’ causality, but added the variable of *nonrigid movement* to the simple shapes which served as stimulus items in these studies. Instead of rigidly moving along a stable trajectory, these shapes elongated from squares into rectangles and back again, resulting in a ‘caterpillar –like’ motion. Results confirmed that, if given these cues of animacy, nine-month-olds do indeed perceive causation-at-a-distance. Other experiments have used different paradigms to test if younger infants are capable of appreciating this principle though the results are less clear (Spelke et al. 1995).

Studies with infants between 3 and 6-months-old show that by this age, children are able to distinguish between different kinds of motion, such as ‘biological’ motion from random motion (Bertenthal 1993; Rochat et al. 1997). Although an appreciation of goal-directedness wasn’t looked at in these studies, Woodward (1998) claims to have shown *goal-attribution* in five-month-olds. By habituating five-month-old infants to a hand reaching for one of two objects, Woodward found that the babies looked longer when the hand reached for the object not previously obtained, regardless of its position. She concluded that the infants were not encoding the structural elements of the display (e.g., movement to the left or to the right), but were instead encoding the goal of the actor’s reach. The claim is strengthened by a condition where the infants did not look longer when the hand was replaced by a metal rod. The rod condition helps to rule out an explanation in terms of a conditioned response (or at least one formed during the habituation phase). It also suggests, like the Meltzoff study, that these children will not attribute goal-directedness to objects that lack cues of animacy.

Perhaps the best evidence of goal-detection in infancy comes from a set of habituation/dishabituation studies by Gergely and colleagues (Gergely et al. 1995; Csibra et al. 1999). In these studies, computer animated circles are shown moving along various trajectories and overcoming obstacles in order to achieve a goal (e.g., to contact the other circle). In one study (Gergely et al. 1995), infants were habituated to either a ball jumping over a barrier and contacting the other ball (experimental group), or to a ball jumping along the same trajectory with the barrier to the right of both circles (control group). The barriers were removed for the test conditions and both groups were shown two conditions: one in which a ball moved in a straight line to contact the other ball (direct), and another in which the ball jumped along the same trajectory as in the habituation phase (indirect).

The experimental group looked longer at the indirect condition, despite the fact that they were habituated to the same trajectory, while the control group looked equally long at both. The authors interpret these results as showing that these infants attributed to the ball the goal of contacting the other ball. They also argue that the infants were not surprised by the direct condition because it was in complete accordance with achieving that goal, that is, it was expected. The indirect condition was surprising to the experimental group because it was not the best way of achieving the goal, indeed, it was irrational, and infants assume that goal-directed agents are rational. Those familiar with Dennett's position will note the familiar rationality assumptions introduced here. The authors see this as well, as is evident by their title "Taking the intentional stance at 12 months of age."

PROBLEMS OF INTERPRETATION

Can a twelve-month-old infant really take the intentional stance? Before we evaluate the interpretations of some of the above experiments, we should look a bit more closely at what Dennett's stance entails. Dennett provides some "rough and ready" principles for intentional attribution which are very difficult to fit into a developmental picture as successful intentional attribution relies on many interdependent and interconnected abilities. For instance, he tells us that:

(1) A system's beliefs are those it ought to have, given its perceptual capacities, its epistemic needs, and its biography.

(2) A system's desires are those it ought to have, given its biological needs and the most practicable means of satisfying them.

(3) A system's behaviour will consist of those acts that it would be rational for an agent with those beliefs and desires to perform. (1987, p.49)

On this picture, intentional interpretation is a package deal, with desire attribution relying on belief attribution and vice versa, all bound up with assumptions of rationality, which, again, are dependent on belief and desire attribution. Psychologists are interested in explaining all of the abilities, but the data suggest that they do not come to the child as a package, nor are they all dependant on each other. How can an infant take the intentional stance while knowing next to nothing about an agent's/system's perceptual capacities, biological/epistemic needs, and biography? In terms of understanding perceptual capacities, children have a quite uneven developmental profile, and even 2 and 3-year-olds are unclear about how information is obtained through various sensory modalities, even seeing (Gopnik and Graf 1988; Pratt and Bryant 1990). Knowledge of various biological functions comes later still (Keil 1992: Carey 1995), and biographies, insofar as knowledge of them is necessary for successful and unique intentional interpretation, may take a good deal of life experience and learning, and depending on the problem, may not be available even to adults.

Yet when Dennett lays out these principles he is setting the guidelines for success with various and possibly unique intentional interpretations, that is, for flexible interpretation and prediction of intentional systems with different belief-desire profiles. While the vast majority of six-year-olds can not take the intentional stance on the Republican party or the Roman Catholic church, they do pretty well with people, so we can expect the core mechanisms to be in place, with success on these other systems dependant largely on experience. But developmentalists are interested in explaining how children *develop* a belief-desire psychology. The assumption of rationality and the implementation of normativity may be central to an older child’s or adult’s intentional stance but psychologists are generally wary of granting these reasoning and/or conceptual abilities to preschoolers or infants. Nonetheless, as reviewed above, many psychologists would like to grant young children an understanding of ‘simple mental states’ such as desires, goals, intentions, attention, and perception (for a review see Johnson 2000).

Desires, goals, and intentions

Philosophy tells us that desires, goals, and intentions, unlike beliefs, are neither true nor false, they are either fulfilled or unfulfilled. One cannot have a false desire, goal, or intention. Psychologists have borrowed this notion to support the claim that these states may be conceptually simpler for young children in that they don’t require positing an attitude toward the truth of a proposition (or state of affairs). But philosophical accounts do not permit desires, goals, and intentions to guide actions alone because they are insufficient, in and of themselves, to carry out actions, and hence must be mediated by beliefs about the world (Bennett, 1978). Thus, if we are to grant the young child a notion a desires, we’ll need to supply a concept importantly analogous to belief to support the child’s predictive capabilities, at least until the concept of belief becomes more adult-like, as it gradually does between 2 and 5 years of age.

Leslie’s (1994) critique of Wellman’s (1990) ‘drive’ theory of desires exemplifies the difficulty in characterizing this early competence. Wellman would like to give the infant a theory of desires without giving him or her proposition-like knowledge. He contends that an infant represents another person’s internal drive toward an object without the ability to embed the object into a proposition. For instance, the infant can only represent the other person as 'wanting' an apple, say. But as Leslie points out, Wellman's attempts to subvert the full propositional attitude notion of desire has a major flaw: It is almost useless in terms of predicting behaviour. On this formulation, Wellman's infant may be able to predict that someone 'wants' an apple, but it does not allow the infant to predict what that person will do with apple (i.e., she cannot represent the notion that the person wants to eat the apple, for instance). Leslie suggests that instead of dropping the propositional content (or state of affairs), Wellman should drop the attitude, formulating the representation as "ACTING to bring about [a state of affairs]" (p.139), thus avoiding the referentially opaque nature of propositional attitudes. Leslie prudently leaves it open as to whether these representations are instantiated in propositional (language of thought) form.

It is unclear whether substituting WANTS with ACTS, ATTEMPTS, TRIES, or some other apparently nonmentalistic term will serve our purposes; this typically behaviourist move has seldom worked before. Leslie's amendment to Wellman's account, while it replaces desire and (putatively) its philosophical baggage, requires that the infant is capable of representing a more complex future state of affairs, or if you will, the goal of the actor. Leslie sees no problem with this and notes that a representation of a future state of affairs is also assumed in the violation-of-expectancy procedures in the physical knowledge tasks describe above. However, no one argues that the infant is attributing a goal or desire to these inanimate objects (to emerge from the occluder for instance, or to come to rest on the ground after a fall). Very young children appear to have expectations about event outcomes but having an expectation is not the same thing as attributing a desire, goal, or intention to another. The important step from expectations of event outcomes into the domain of psychological reasoning may lie with the attribution of causes to those events, as characterized by the difference between expecting that an agent will do such and such and attributing as causal the agent's *intention* to do such and such, for instance. But this is no small step, and while a careful use of language may help to clarify what sort of competence we are looking for, it may draw neat lines that bear little relation to the actual psychological mechanisms supporting that competence. This is always the danger when employing folk concepts in a scientific psychology.

In our everyday language the terms desire, intention, and goal are often used interchangeably. If we are told that "someone intends to do *x*," it seems fair to paraphrase the statement as "someone has the goal of doing *x*," and vice versa. Here the terms refer to something like a plan, which we naturally take to be in the head. So if a researcher claims that an infant detects another's goals, it is natural to think that they likewise detect intentions. Put another way, it seems strained to say that someone has the goal of doing *x* but does not intend to do *x*. Nevertheless, most infancy researchers would like to avoid this conflation, and use goal-detection to refer to the infants ability to detect that an agent's behaviour is directed at or about an object (including other agents) or a state of affairs. Thus, because the mechanisms underlying this competence are claimed to detect aboutness relations, they are claimed to detect a basic intentionality. The same may be claimed for the cluster of joint or shared attention behaviours, and well as the instances of social referencing mentioned above.

It could certainly be argued that these attributions are not mentalistic at all, but a logical argument alone is not much use at this stage of the game. What is important is to show, empirically, what a proposed mechanism *does*. If a researcher or theorist claims that an infant represents another's goals, for instance, then it needs to be shown what the child can *do* with this representation. If, on the other hand, the young child's competence is best characterized by lower-level learning mechanisms, such as correlation or contingency detection, then there may be no work left for a representation of another's goals to do. In such a case, it would be the researcher who is taking the intentional stance rather than the child. (We discuss the point in more detail in the section on cognitive neuroscience.)

Bringing Rationality into the Picture

The studies by Gergely et al. (1995) and Csibra et al. (1999) provide the best evidence of the kind of goal-attribution mentioned above, but it should be clear by now that this competence is not the intentional stance proper. To be fair, the authors know quite well that infants are not taking the intentional stance, and suggest instead that infants take a *teleological stance* (Gergely & Csibra 1997; Keil, 1994), a nonmentalistic, noncausal, precursor to the intentional stance that interprets actions as goal-directed. The authors are not prepared to grant infants knowledge of beliefs, desires, or intentions, but claim that they capable of interpreting actions as occurring *in order to* achieve something.

The authors see rationality assumptions as a set of constraints on the evaluation of multiple alternatives.[2] These constraints are suggested to play *the very same functional role* in the teleological stance as does rationality in the intentional stance. So what role does rationality play in the intentional stance? It is a background assumption that constrains hypotheses as to how an agent will act given its belief-desire profile. Irrational agents are unpredictable from the intentional stance so the rationality assumption is, in a sense, forced upon us. The assumption is quite implicit, however, and it is unlikely that there is any explicit representation of this sort when we predict behaviour.

Thus, when Gergely and Csibra consider rationality a 'property' attributed to actions (not agents) they may be forcing the notion to do more work than necessary. Constraints on reasoning in any domain are important but there are cheaper and easier

---

[2] It is not entirely clear what role rationality plays for the Gergely and Csibra model. Rationality, to them, is a "core inferential principle (the principle of rationality)," a "property" that is attributed to actions, and a constraint on evaluation of multiple alternatives (1997, pp. 232-233).

ways to get them. A free constraint on multiple hypotheses (expectations may be more accurate) is the child's limited ability to represent multiple ways for A to get to B, for instance. Additional constraints on goal-directed action can also come from core principles in the domain of folk physics, e.g., that agents can not pass through solid objects, that agents' motion is continuous, that agents remain bounded when moving, etc. Even a principle akin to 'agents move in the shortest path toward their targets' need not bring in rationality and may even be learned by example. Gergely & Csibra (see also Csibra and Gergely 1998) introduce an important issue here by looking for and proposing constraints in this domain, which will no doubt differ in important ways from constraints on physical and mechanical reasoning. Though rationality assumptions may be too much too soon, the principles that the infant is using to constrain expectations of goal-directed behaviour will indeed be similar to the rationality assumptions inherent in the intentional stance, implicit as they may be.

An interesting analog between the transparency of belief and the transparency of rationality may serve to highlight the similarities and differences between these rationality assumptions. Just as beliefs only come into focus when there is a conflict or discrepancy of some sort, so too with rationality. Young children do quite well predicting behaviour without the concept of belief or rationality because a) most beliefs are true, and children's knowledge of what (they think) is real in the world is an adequate substitute for the concept of belief – that is, most beliefs are shared between the interpreter and interpretee, and b) natural selection builds (relatively) rational agents, and hence most behaviour is rational. In this way, rationality only comes into focus when a behaviour is found to be inexplicable. The belief-desire profile of the system in question is crucial, however, in that the interpreter may update this profile to make sense of the behaviour (e.g., in the case of false belief or impaired perception). If no amount of revising renders the behaviour intelligible then the system is deemed irrational, and hence unpredictable from the intentional stance.

Developmental Summary

Dennett has been criticized both for setting the 'mindreading' bar too high (Johnson 2000) and too low (Premack and Premack 1997). While it is true that Dennett suggested the false belief task as a measure of teasing apart first-order intentional systems from second-order intentional systems, he never suggested that it was the only evidence. Indeed, in 1983, just as the child's ToM literature was beginning, Dennett made the following prediction:

> It will turn out on further exploration that [young children] will exhibit mixed and confusing symptoms of higher-order intentionality. They will pass some higher-order tests and fail others; they will in some regards reveal themselves to be alert to third-order sophistications, while disappointing us with their failure to grasp some apparently even simpler second-order points. No crisp,

> 'rigorous' set of intentional hypotheses of any order will clearly be confirmed. (1983/1987, p.255)

We took the liberty of inserting 'young children' into that passage -- Dennett was actually referring to higher nonhuman animals -- but the change is perfectly consistent with his program. Immediately following this passage he writes "I expect the results of intentional interpretations…of small children, to be riddled with the sorts of gaps and foggy places that are inevitable in the interpretation of systems that are, after all, only imperfectly rational" (p.255). We now have almost 20 years of data on the topic and Dennett's prediction has been borne out. The literature on children's early intentional interpretation is foggy indeed, but this is not due to a failure to replicate, it is due to an uncertainty about what one is committed to when introducing propositional attitudes into their characterizations of young children's abilities.

While Dennett's prediction may appear to be pessimistic, it is not at all; it simply follows from his theory. Dennett is often seen by scientists as an authority on propositional attitudes and this has led some to believe that he actually buys the classical cut. This is far from the case. Not one in the habit of sugar-coating, he writes "the large and well-regarded literature on propositional attitudes… is largely a disciplinary artefact of no long-term importance whatever, except perhaps as history's most slowly unwinding unintended *reductio ad absurdum*" (1994, p. 241). His distrust of this literature, and in clean qualitative leaps in general, relates directly to a central question in our literature: "When does the child first have a theory of mind?" There will be no clean cut between intentional and non-intentional phenomena, and indeed, the same phenomena may be described using various levels of description. What is important is what we gain or lose in terms of prediction within these levels. This is why he invites scientists to take the intentional stance, while at the same time warning us of its pitfalls. Thus, the mindreading bar can be set as high or as low as we like, as long as we remember that "what counts as mindreading is a less important question than the question of how such an apparent mindreading competence might be organized" (Dennett, 1996, p.124).

THE COGNITIVE NEUROSCIENTIFIC BASIS OF THE INTENTIONAL STANCE

We turn now to a brief review of the burgeoning literature on the brain basis of mental state attribution. A belief-desire psychology is not a simple process and many brain regions have been suggested to underlie the broad competence. We will concentrate only on the regions of considerable interest at present, which consist of the medial prefrontal cortex (paracingulate cortex, PCC), the superior temporal sulcus (STS), the temporal parietal junction, and the orbitofrontal-amygdala-temporal circuit. We concentrate first on the phylogenetically older substrates involved in social perception and cognition, and later on studies that look directly at theory of mind in humans.

SOCIAL PERCEPTION AND COGNITION

Superior Temporal Sulcus

Woodward's (1998) results with five-month old children suggest that they are capable of encoding the goal of an actor's reach. Though Woodward doesn't appeal to this literature in the discussion of her findings, there is quite a bit of evidence from single-cell recording in monkeys that neurons in and around the superior temporal sulcus, as well as the inferior frontal cortex, are sensitive to various goal-directed reaching motions.

Perrett and colleagues have studied 'hand action' cells in the ventral areas of the STS region and have discovered several interesting response properties. For instance, many of these cells respond better to particular kinds of actions, such as reaching, grasping, picking, tearing, etc. This responsiveness generalizes across the objects being acted on, across various visual perspectives, and across several spatiotemporal trajectories of the actions (e.g., different speeds and distances). Furthermore, the responsiveness of these neurons is greater when the actions are goal-directed (Perrett et al. 1989).

Cells in the inferior frontal cortex complement the STS cells and code both visual and motor components of these actions. These cells, called "mirror neurons" have the interesting property of firing not only when the monkey witnesses an action on an object, but also when the monkey executes an action on that object, hence the 'mirror' rubric. Unlike the STS cells, the mirror neurons will continue to fire in the dark and during forced delays in reaching. Other cells in the STS region help to avoid a potential confusion over who is acting, firing only when another acts. These 'other' cells respond continually and can not be habituated, even after long exposure to predictable (rhythmic) actions (Emery & Perrett 2000). The STS region contains other neurons which fire preferentially to head direction and eye direction, both of which have been proposed as dedicated mechanisms in the mindreading competence (e.g., shared attention and eye direction detection, see Baron-Cohen 1995; Puce et al. 1998).

The perception of biological motion has also been attributed to regions in and around the STS. In addition to the perception of various hand actions and head and eye direction described above, movements of the entire body activate cells in this region. A series of Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) studies found activation in posterior STS and anterior superior temporal gyrus to meaningful motions such as walking, dancing, and throwing (Bonda et al. 1996; Howard et al. 1996; Grossman et al. 2000). These studies used point-light displays, where lights are attached to various body parts of actors filmed in total darkness. Humans readily interpret meaningful motion from point lights alone (Johansson 1973).

These areas are dorsal and anterior to V5, the homologue of monkey V5/MT, which is specialized for the detection of motion more generally. Damage to V5 can selectively damage the ability to perceive motion and we might expect that damage to these regions could selectively impair the recognition of biological motion. Such a case has recently been reported (Cowey and Vaina 2000). Interestingly, just the opposite has also been found: a motion-impaired patient with damage to V5 was able to interpret the point-light displays as meaningful, while at the same time unable to determine whether various objects were moving (Vaina et al 1990; see also McLeod et al 1996).

A potential explanation for this strange finding may come from studies on *implied motion.* Increased regional cerebral blood flow (rCBF) is found in V5 and in the STS when participants are shown still photographs of implied motion (e.g., stills of someone in the act of throwing a discus or stills of hands acting on objects) (Kourtzi and Kanwisher 2000; Senior et al. 2000). Thus, the intact STS regions may have been sufficient to interpret the point-light displays as meaningful if received in a form analogous to still-frames, despite the lack of continuous motion.

Orbitofrontal-Amygdala-Temporal circuit.

The orbitofrontal-amygdala-temporal circuit has been implicated in several models of social cognition in monkeys, apes, and humans (Brothers and Ring 1992; Baron-Cohen 1995; see Adolphs, 1999a for a recent review). These areas share reciprocal feedforward and feedback connections and their proper functioning is to a large degree interdependant. Monkeys with lesions in these areas develop compulsive behaviours with objects, especially via oral examination; they present with unusual tameness, social isolation and avoidance, and appear to lose the ability to mark the emotional significance of stimuli (e.g., picking up snakes) (Kluver and Busey 1939; Myers et al. 1973). We concentrate below on the role of the amygdala and orbitfrontal/ventromedial (OFC/VM) cortices in this circuit.

The amygdala has been implicated as playing a causal role in autistic spectrum conditions (Baron-Cohen 2000). In a task requiring subjects to attribute social information (intentions) based on information from eyes alone, individuals with Asperger syndrome (a mild form of autism) perform worse than controls (Baron-Cohen et al. 1997) and an fMRI version of this task revealed significantly less amygdala activation in the individuals with Asperger syndrome (Baron-Cohen et al. 1999).

The amygdala has a well-known role in fear conditioning in animals and can trigger flight mechanisms even before many features of the stimulus are recognized (Ledoux 1996). Several recent imaging studies with normal subjects and studies with individuals with damage to the amygdala point to an important role for the amygdala in the recognition of emotional expressions, particularly fearful faces (Adolphs 1999b; Young et al. 1995). Patients with bilateral amygdala damage have a tendency to judge faces as far more ‘approachable’ and ‘trustworthy’ than do controls (Adolphs et al. 1998). Thus, the majority of research on the amygdala converges on its important role in the processing of fearful and dangerous stimuli. The amygdala may have a more general function however, in the modulation and allocation of processing resources when a stimulus is ambiguous, regardless of valence (Whalen 1999).

Complementing the amygdala, single-cell recording studies have shown that cells in the inferotemporal cortex respond preferentially to information about faces, such as identity, social status, emotional expression (Young and Yamane 1992) and, along with cells in the amygdala, are found to be active during scenes of complex social stimuli (Brothers et al. 1990).

The role of the OFC/VM regions in social cognition has been discussed extensively by Damasio (1989, 1994; Damasio and Anderson 1993), who argues that the OFC/VM aspects of the frontal lobe act as ‘convergence zones’ and have a special role in coordinating the marking of the emotional significance of events. This marking consists of a circuit which includes amygdala, various limbic and cortical structures, and uses the body (state of the soma) as a ‘sounding board,’ in effect, against which to base decision making. Patients with damage to orbital and ventromedial regions may lose the ability to use the state of the soma as a value marker for potential outcomes, resulting in the odd and often inappropriate social behaviour observed in patients with prefrontal damage.

While damage to these areas may have no effect on standard measures of intelligence, patients with OFC/VM lesions often have difficulty with planning and on reasoning tasks involving ‘hunches’ and social scenarios. For instance, a series of studies have shown that OFC/VM patients perform poorly on gambling tasks, where choices are determined by emotional hunches in the face of incomplete information. Normal subjects learn to maximize gains by combining statistical contingencies with the value of the payoff over time, though no explicit reasoning appears to be involved (Bechara et al. 1997). Variations on the Wason deductive reasoning task which involve familiar or social information, but not those requiring abstract reasoning, are also difficult for OFC/VM patients. Correct reasoning on these tasks relies on deciphering threats, promises, and so on, against backdrop of knowledge of social contracts. Thus, these regions may subserve what Cosmides and Tooby (1992) refer to as a ‘cheater detection module’; a cognitive adaptation crucial for the maintenance of an evolutionary stable strategy of social exchange.

THEORY OF MIND STUDIES

Brain Injury

There has only been a handful studies directly concerned with mental state attributions in adults with acquired brain injury. Nonetheless, some striking findings have emerged: To date, only patients with damage to the right hemisphere (RHD), but not the left (LHD), have shown ToM deficits (Happé et al. 1999; Stone et al 1998; Griffin et al., in preparation).

Varley and Siegal (2000) tested an LHD severe aphasic on first-order and second-order ToM tasks and found no deficits despite the subject's almost total lack of language. Stone et al. (1998) tested LHD and bilateral orbitofrontal patients on first and second-order ToM and social faus pax tasks. Neither group had difficulty on the first-order and second-order tasks, though the bilateral orbitofrontal patients had difficulty on the social faux pas measure. Finally, Happé, Brownell and Winner (1999), in the first direct comparison between the groups, tested both RHD and LHD patients on a series of ToM stories and cartoons. The RHD but not the LHD group were found to have a selective impairment in ToM reasoning. This result was recently replicated and extended by Griffin and colleagues, and a closer analysis revealed that RHD patients did not differ on first-order ToM attributions, but only on second-order ToM and in their ability to detect deception. Only one of these patients had damage to the orbitofrontal regions, the right amygdala, and the right temporal lobe, though she performed as well as normal age-matched controls on the ToM measures.

RHD patients have well-known difficulties in the expression and perception of emotion, difficulties interpreting nonliteral language, and difficulties using context for inferential purposes, deficits that are shared with the autistic spectrum (Brownell et al. 2000; see also Happé et al. 1999, Table 1, and Sabbagh 2000, for a longer catalogue of shared impairments). Indeed, while the developmental language delays that accompany autism were originally taken to suggest deficient left hemisphere function, the nature of the impairments, such as deficits in prosody, social use of language, and inability to read facial expressions, is more suggestive of impairments following RHD (Prior and Bradshaw 1979). Shields et al. (1996) tested a group of children with autism and a group with semantic-pragmatic disorder on a battery of tasks sensitive to left and right hemisphere injury. Both groups performed better on the left hemisphere battery relative to the right hemisphere battery, leading the authors to suggest that semantic-pragmatic disorder is part of the autistic spectrum, and implicating a link with right hemisphere dysfunction.

Neuroimaging Studies

A series of ToM imagining studies using stories, wordless cartoons, and moving geometric shapes have found peaks of activation in many different areas, although a rough pattern is starting to emerge. Nearly all of these studies have noted selective activation in areas in and around the medial prefrontal cortex at the border of the anterior cingulate (PCC). Whether this activation is lateralized to the left or the right may depend whether the stimulus items are language- or visual-based, with visual-based stimuli resulting in more rightward activation, although a proper meta-analysis has yet to be done (Brunet et al. 2000: Fletcher et al. 1995; Gallagher et al. 2000; Castelli et al. 2000).

Additionally, selective activation in the temporoparietal junction (at the border of the STS) has been found in several of these studies (Baron-Cohen et al 1999; Castelli et al. 2000; Gallagher, et al. 2000). The studies on biological motion mentioned above also fall within this region (Bonda et al. 1996; Puce et al. 1998), although very few of the materials in these studies involved motion. The concept of implied motion may again be useful in interpreting these results, though Gallagher et al. take their results to indicate more, suggesting that temporoparietal junction "is sensitive not merely to biological motion but, more generally, to stimuli which signal intentions or intentional activity."

Allison et al (2000) suggest that the putative coding of intentions discussed by Gallagher et al. is also true for an STS cell which responds to downward flexion of the head, a sign of submission in primates. They write "such gestures are probably *intended* by the viewed monkey to signal submission, and are probably interpreted as such by the viewing monkey"(emphasis ours). There is no evidence that this cell responds to the intention to be submissive, as opposed to the downward (ventral) flexion of the head, but the seduction of discovering the 'intention' cell appears to be too much to resist.

Dennett recommends that we employ the IS to provide the ideal against which to test explanations (and find the bargain). There is no reason to employ it, however, if a lower level strategy is more predictive and the IS description adds nothing. The characterization of neuronal cell assemblies involved in social perception such as those discussed above, are good cases where several of levels of description may be useful. Dennett (1989) provides one such example from Braitenberg's (1984) discussion of the bilateral symmetry detectors common in animal vision systems. These mechanisms do detect bilateral symmetry, and a description at this level is more predictive of the features of stimulus items that will trigger their response than, say, their ability to detect *that some other organism is looking at me* (p.109). But the existence of such a mechanism would be confusing unless it is looked at as quick and dirty discriminatory mechanism that detects this kind of evolutionarily important data. The intentional characterization provides the ideal and an explanation of what the device is for. With this information we can make assumptions about the costeffectiveness of the mechanism. Moreover, a symmetry detection description, while more predictive about the parameters of the mechanism, will miss the rationale for its triggering of fight, flight, and other mechanisms. Dennett might be amused by the unabashedly intentional title of a paper on the analysis of social signals at the cell level “Someone is looking at me, something touched me, something moved!” (Perrett et al. 1990).

The Medial Prefrontal Cortex/Paracingulate Cortex

The medial prefrontal cortex, bordering on the anterior cingulate, has been the central region of discussion in the ToM imaging literature. This region has been activated in several studies using different materials and has been significantly more active than control conditions. Frith and Frith (1999) optimistically suggest “that a brain system dedicated to mentalising can be localized” (p. 1693) and point to the areas bordering the anterior cingulate and medial prefrontal cortex (PCC) as the locus of the mechanism that represents the mental states of self and other. They point to other imaging studies which invite similar interpretation of the PCC as the mindreading centre, such as the reporting of one’s emotions, speech and response monitoring, self-generated thoughts, and the perception of pain and tickling (Lane et al. 1997; Rainville, et al.1997; Blakemore et al. 1998; Carter et al. 1998; McGuire et al. 1996; Frith and Frith, 1999; see also Castelli et al. 2000, Table 4).

Frith (1996) cites Dennett’s IS and immediately thereafter suggests that the distinction made between propositional attitudes and propositional content similarly applies to prefrontal and posterior cortices respectively, with the prefrontal cortices representing the attitudes and posterior cortices representing various contents. Frith contends that prefrontal cortices subserve the ‘X believes that,’ or the ‘Y intends that’ representations while the content , the Ps, Qs, and Rs, are housed in posterior, mostly temporal, cortices. The same is true for determining ones own mental states, where prefrontal cortices remain active representing ‘I believe that’ while the proper content is elicited from posterior regions. This is one of many misreadings of Dennett, who has been arguing for most of his life that propositional attitudes are not in the head.

An appeal to the non-ToM imaging literature, which is far larger, may help clarify some of the confusion. In a review of 107 PET studies reporting activation in this region, Paus and colleagues (1998) argue that task difficulty is the best predictor of paracingulate activation. This region is closely associated with the anterior cingulate, which has been implicated in lexical retrieval, semantic encoding, and monitoring of action (Posner and Dehaene 1994). Similarly, Duncan and Owen’s (2000) recent review argues that this region (which they call the dorsal anterior cingulate) is recruited for diverse cognitive demands, such as response conflict, novelty, number of elements and time delays, and perceptual difficulty. They suggest that this region has a specialized function: In concert with mid-dorsolateral and mid-ventrolateral cortex, it is specialized for the solution of diverse cognitive problems, that is, it is specialized for hard problems. Activity in this region increases when errors due to response competition are likely (Carter et al. 1998).

With this in mind, we can see where Frith's analogy between propositional attitudes and propositional content arises. The tasks cited above involve the inhibition and monitoring of competing elements and hence elicit these regions in the same way that tasks in many domains other than ToM elicit similar regions. The PCC remains active and serves to elicit or inhibit competing responses, or as it were, competing contents. It is a loose analogy, to be sure, and perhaps a very misleading one. If the PCC is activated for the reasons described above we should not see this activation on simple or well-rehearsed ToM scenarios. This is an empirical question.

The neuroimaging and adult lesion literature tends to lack a developmental perspective. It is one thing to learn a skill and another thing to perform that skill after it is well-learned. The fact that our faculty of folk psychology is so well-rehearsed should lead us to expect that much of it becomes automatized, at least insofar as the causal stories we tell about everyday behaviour (e.g., "He brought his umbrella because he thought it was going to rain"[3]). Hence that the multi-step tasking characteristic of the frontal lobes, and the various task demands mediated by the PCC may not be necessary in simple or well-rehearsed ToM reasoning, though these regions will be quite important in the development of ToM and its proper functioning in more difficult mental state attribution tasks. It is also possible that some of these regions are biological adaptations for complex social problems and have been co-opted for novel problems, though this literature can not speak directly to these questions. Finally, it is important to note that no coherent patient ever loses the ability to supply mental state descriptions, even if their damage is to MPFC/ PCC, although their facility with determining the reasons and causes of behaviour may be affected by damage to a number of substrates underlying our folk psychology. The nature of the deficit will depend not only on the functional role of the damaged substrate(s) but also on the task demands.

There are several other interesting threads in this literature, such as the activation of the right middle frontal gyrus and precuneus in imaging studies on both metaphor comprehension and ToM, suggesting possible shared substrates for alternative readings and weak associations, which may explain why RHD patients have trouble with non-literal speech and ToM (Bottini et al. 1995; Brownell et al. 2000). Moreover, bilateral temporal pole activation has been found in ToM studies and in studies on sentence and narrative comprehension, pointing to a potential story-telling component in these tasks (Happé et al. in preparation; for more on the brain and ToM, see Baron-Cohen et al. 2000). The information processing biases of various

---

[3] Moreover, it is quite doubtful that the brain follows the putatively important philosophical distinction between the following explanations: "He *wants* to go to the bathroom" and "He *has* to go to the bathroom." That is, we should not expect the first of these explanations to be supported by the ToM substrates and the latter to be supported by non-ToM substrates.

brain regions are becoming clearer almost daily and a picture of how the brain's role in parsing the social world and constructing mentalistic narratives as a tool for prediction and explanation is beginning to emerge.

Cognitive Neuroscience Summary

Although we have interspersed commentary and criticism throughout the above section, there are a few other points which may serve to highlight the differences between the empirical research in this area and the IS. The main tension lies with concerns about reducing mental states to brain states, specifically regarding claims about cause. Many in philosophy, Dennett included, subscribed to a form of externalism, according to which contentful states are seen as relational properties, and are identified by reference to entities outside the brain. Thus, if content ascriptions are extrinsically relational, then they can not refer directly to the local, causal, nexus in the brain. The features of the cause must be local to the causal interaction. We treat the mind as a semantic engine, yet when we look at the brain all we see is a syntactic engine, where the shape and orthography of neurons and neurochemicals are intrinsically causal, and its hard to see how to get semantics out of syntax.

A reductionist sees semantics as reducible to the brain's syntax, whereas a non-reductionist, such as Dennett, sees them as different levels of description. Dennett does not eschew representational talk, on the contrary, he invites it, provided that we don't treat representational states as brain states. These different levels usually describe different phenomena, and one should be abandoned in favour of another only if one proves more successful at describing/predicting the same phenomenon. Thus, representational talk is fine if it is seen in non-reductive, functional terms.[4]

BIOLOGICAL MECHANISMS AND CULTURE

In this essay we have concentrated largely on phylogenetically older abilities that underlie the intentional stance proper. We have paid little attention to a central feature of the IS and of folk psychology more broadly – language. While Dennett considers the intentional stance the craft of folk psychology, there is much more to the craft

[4] An example may help to clarify how functional talk is not causal talk. A physicist would say that heating a gas causes it to expand, and could provide laws that would make this predication. A biologist would say that heating a mammal causes it to sweat, and that the *function* of sweating is to keep the animal's temperature constant. The physicist would never say that the function of the gas expanding was to keep its temperature constant, even though that is precisely what happens. Thus, functions are effects, not causes, and can not be seen from the physical stance alone. A claim such as "the heart pumps *in order to* circulate the blood" is teleological, not causal, because effects do not bring about their causes.

than laid out in his rough and ready principles of intentional attribution, as we hope to have shown. Chimpanzees and other primates navigate their complex social landscapes quite impressively, despite the absence of a generative language ability, and we can expect to share many of the mechanisms that allow for this prowess. Their absence of language, however, and their inability to trade reasons and causal stories about behaviour severely limits their predictive and explanatory capabilities.

Our culture of story-telling allows us to anchor what Dennett calls "free-floating rationales." Free-floating rationales are reasons for behaviours that are not explicitly represented in the organism, but implicit when looking at the design of the larger system. For instance, while we can see the rationale for a piper plover's feigning of a broken wing, the piper plover can not. There are reasons not only for the broad behaviour, but also for the functional design of the mechanisms that carry out the behaviour. Dennett (1996) wishes to explain not only how we developed the ability to see these reasons, but how they came to be "captured and articulated in some of the minds that evolved," so that they become the agent's *own* reasons.

Dennett doesn't expect that a novel, species-specific, mental organ is responsible for our ability to see these rationales or to detect higher-order intentional patterns. Instead, he looks for a more parsimonious route and stresses the role of culture in anchoring these rationales, both within and outside the head. His view is a combination of hardcore evolutionary psychology along the lines of Cosmides and Tooby (1995, 1997) and socio-culturalism along the lines of Vygotsky (1979), positions that are often seen as diametrically opposed. For Dennett, intentionality is not specific to humans nor is it the mark of the mental, but it already exists in organisms built by natural selection to detect and exploit their various niches. In turn, natural selection builds systems that detect these (already intentional) systems, such as bilateral symmetry detectors ("someone is looking at me") or the neuronal ensembles which pick up on eye direction, head direction, goal-directed reaching, and emotional expressions in relation to objects or events in the world (social referencing).

We may share the mechanisms mentioned above with our primate cousins, but for them and for infants, the rationales for these mechanisms are free-floating and invisible. The Vygotskyan flavour of Dennett's view has to do with an outside-in move in the composition, revision, and endorsement of the reasons for our behaviour. This can be put in terms of the notion of 'memes'(this term was introduced in the Introduction). The memes of folk psychology exist in the cultures into which we are born, and our biological constitution renders us the perfect hosts for their instantiation and dissemination. The same is true for the memes of science, some of which are replacing the folk memes, though we are less suitable hosts for these in that we were not designed to break up the world in many of the ways illuminated by science. The process of composition, revision, and endorsement is dynamic, and exists between individuals, the memes of their culture and time, and the state of the world. In this way, the rationales for even our own behaviour – for example, that it is guided by our own goals and intentions -- moves from the outside-in.

Dennett sees the pressures of communication in our species as forcing us into declaring categories ("Are you going to fish or cut bait?") and in turn, creating the illusion of more definition of content than actually exists (see also McFarland 1989). These declarations are born from a tangled, competing web of neural circuitry, but the victorious declarations serve to convince us that our behavioural tendencies are controlled by explicitly represented goals – or intentions. Dennett calls this a form of *approximating confabulation* -- carving nature in places where there are no salient joints – and compares it checking off an answer in a poorly-designed multiple-choice exam. If "none of the above" is not an option, we're forced to settle for the nearest miss. In this way, representations of intentions enter in a backhanded way.

Almost immediately after children begin talking they create narratives of their actions (e.g., "now I go up," "now I sit here"), no doubt creating the illusion of clear-cut intentions, but these narratives also serve to change their cognitive and intentional profiles. They quickly develop a list of options and reasons to justify their own behaviours and to predict what others will do. By the time they leave preschool, they are quite adept folk psychologists indeed. Exhaustively charting the many changes between the detection of biological motion in infancy and the ability to tell socially respectable causal stories of behaviour in primary school has occupied lifetimes of research and will occupy many more. While many important details are still unknown, the bigger picture is starting to come into focus.

Folk psychology is intuitive but the intentional stance, despite the many similarities to FP, is not. It is difficult for many to accept the notion that beliefs and desires are not reducible to brain states, that is, that they are not in the head. It is less difficult to accept the fact that it is beneficial to abstract away from the messy and often irrelevant details of a complex system when attempting to predict and explain its behaviour. This is what science attempts to achieve and this is why Dennett invites science to take the intentional stance. He polices his own theory, however, and is the master of what he calls "killjoy" hypotheses that deflate the intuitive intentional characterization with lower-level explanations. Dennett is not sending mixed messages, he is simply promoting good science. Dennett is a champion of science, although he is well aware of many of the potential minefields that philosophy has charted, such as reducing meaning to the syntax of the brain, a particularly troubling problem for those of us who freely talk about representation. Yet instead of declaring the field impassable, like many in philosophy, Dennett is handing out maps of how to navigate this potentially treacherous course.

References

Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, 393(6684), 470-474.

Adolphs, R. (1999a). Social cognition and the human brain. *Trends in Cognitive Sciences, 3,* 469-479.

Adolphs, R., Tranel, D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., Anderson, A., Lee, G. P., & Damasio, A. R. (1999b). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia, 37*(10), 1111-1117.

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4*(7), 267-278.

Astington, J. W., Harris, P. L., & Olson, D. R. E. (1988). *Developing theories of mind*. New York: Cambridge University Press.

Baillargeon, R. (1993). The object concept revisited: New directions in the investigation of infants' physical knowledge. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy, Carnegie Mellon Symposia on Cognition* (Vol. 23). Hillsdale, NJ: Erlbaum.

Baillargeon, R. (1995). Physical reasoning in infancy. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* . Cambridge, MA: MIT Press.

Baldwin, D. (1995). Understanding the link between joint attention and language: Development of joint visual attention in infants. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origin and role in development* . Hillsdale, NJ: Erlbaum.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (Eds.). (1993). *Understanding other minds: perspectives from autism*. Oxford: Oxford University Press.

Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *British Journal of Psychiatry, 165*(5), 640-649.

Baron-Cohen, S. (1994). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/ Current Psychology of Cognition, 13*, 513-552.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry, 38*(7), 813-822.

Baron-Cohen, S., Ring, H., Wheelwright, S., Bullmore, E., Brammer, M., Simmons, A., & Williams, S. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience, 11*, 1891-1898.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. (Eds.). (2000). *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford: Oxford University Press.

Baron-Cohen, S. (2000). The cognitive neuroscience of autism: implications for the evolution of the male brain. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences* (2nd ed.). Cambridge, MA: MIT Press.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. New York: Oxford University Press.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293-1294.

Beeman, M., & Chiarello, C. (1998). *Right hemisphere language comprehension: Perspectives from cognitive neuroscience*. Mahwah, NJ: Lawrence Erlbaum.

Bennett, J. (1978). Some remarks about concepts. *The Behavioral and Brain Sciences, 1*, 557-560.

Bertenthal, B. (Ed.). (1993). *Infants' perception of biomechanical motions: Intrinsic image and knowledge-based contraints.* Hillsdale, NJ: Erlbaum.

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience, 1*(7), 635-640.

Bloom, P. (2000). *How children learn the meaning of words*. Cambridge, MA: MIT Press.

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience, 15*(3737-3744).

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S., & Frith, C. D. (1994). The role of the right hemisphere in the interpretation of figurative aspects of language: a positron emission tomography activation study. *Brain, 117,* 1241-1253.

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, MA: MIT Press.

Brothers, L., Ring, B., & Kling, A. (1990). Responses of neurons in the macaque amygdala to complex social stimuli. *Behavioural Brain Research, 41*, 199-213.

Brothers, L., & Ring, B. (1992). A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience, 4*(2), 107-118.

Brownell, H., Gardner, H., Prather, P., & Martino, G. (1995). Language, communication, and the right hemisphere. In H. S. Kirshner (Ed.), *Handbook of neurological speech and language disorders* (Vol. 33, pp. 325-349). New York: Marcel Dekker, Inc.

Brownell, H., Griffin, R., Winner, E., Friedman, O., & Happé, F. (2000). Cerebral lateralization and theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: perspectives from developmental cognitive neuroscience* . Oxford: Oxford University Press, pp. 306-333.

Brunet, E., Sarfati1, Y., Hardy-Baylé, M.-C., & Decety, J. (2000). A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage, 11*(2), 157-166.

Carey, S. (1995). On the origin of causal understanding. In D. Sperber & D. Premack (Eds.), *Causal cognition: A multidisciplinary debate. Symposia of the Fyssen Foundation* (pp. 268-308). New York: Clarendon Press/Oxford University Press, pp. 268-308.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*(2), 315-330.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection and the online monitoring of performance. *Science*, *280*, 747-749.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage*, *12*, 314-325.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. H. Barkow & L. Cosmides (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 163-228). New York: Oxford University Press.

Cosmides, L., & Tooby, J. (1995). From function to structure: The role of evolutionary biology and computational theories in cognitive neuroscience. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1199-1210). Cambridge: MIT Press.

Cosmides, L., & Tooby, J. (1997). Dissecting the computational architecture of social inference mechanisms. *Ciba Foundation Symposium, 208*, 132-156; discussion 156-161.

Cowey, A., & Vaina, L. (2000). Blindness to form from motion despite intact static form perception and motion detection. *Neuropsychologia*, *38*(5), 566-578.

Csibra, G., & Gergely, G. (1998). The teleoloigcal origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science, 1*, 255-259.

Csibra, G., Gergely, G., Biro, S., Koos, O., & Brockbanck, M. (1999). Goal attribution without agency cues: the perception of 'pure reason' in infancy. *Cognition, 72*, 253-284.

Damasio, A. (1989). The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation, 1*, 123-132.

Damasio, A. R., & Anderson, S. W. (1993). The frontal lobes. In K. M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology (3rd ed.)* (pp. 409-460). New York: Oxford University Press.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* New York: Grosset/Putnam.

Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Clarendon Press.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in cognitive sciences, 3*(5), 172-178.

Dennett, D. C. (1978). Beliefs about beliefs. *The Behavioral and Brain Sciences, 1*, 568-570.

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA.: MIT Press.

Dennett, D. C. (1989). Cognitive ethology: Hunting for bargains or a wild goose chase? In A. Montefiore & D. Noble (Eds.), *Goals, no-goals, and own goals: A debate on goal-directed and intentional behaviour* . London: Unwin Hyman.

Dennett, D. C. (1991). Two contrasts: Folk craft versus folk science, and belief versus opinion. In J. Greenwood (Ed.), *The future of folk psychology: Intentionality and cognitive science* (pp. 135-148). Cambridge: Cambridge University Press.

Dennett, D. C. (1996). *Kinds of minds: Toward an understanding of consciousness*. New York: Basicbooks.

Dennett, D. C. (1998). *Brainchildren: Essays on designing minds.* Cambridge, MA: MIT Press.

Duncan, J., & Owen, A. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23*, 475-483.

Elman, J. L., Bates, E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking Innateness: Connectionism in a Developmental Framework*. Cambridge, MA: MIT Press.

Emery, N. J., & Perrett, D. I. (2000). How can studies of the monkey brain help us understand 'theory of mind' and autism in humans? In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding Other Minds: Perspectives from developmental cognitive neuroscience* (pp. 274-305). Oxford: Oxford University Press.

Feinman, S., & Lewis, M. (1983). Social referencing at 10 months: a second order effect on infants' responses to strangers. *Child Development, 54*, 878-887.

Field, T., Vega-Lahar, N., Scafidi, F., & Goldstein, S. (1986). Effects of maternal unavailability on motion-infant interactions. *Infant Behavior and Development, 9*, 473-478.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, D. J., Frackowiak, R. S. J., & Frith, C. D. (1995). Other minds in the brain: A functional imaging study of "theory of mind" in story comprehension. *Cognition, 57*(2), 109-128.

Fodor, J. A. (1976). *The Language of Thought*: Brighton, Sussex: Harvester.

Fodor, J. (1987). *Psychosemantics: the problem of meaning in the philosophy of mind.* Cambridge: MIT Press.

Frith, C. (1996). Brain mechanisms for "having a theory of mind". *Journal of Psychopharmacology, 10*(1), 9-15.

Frith, C., & Frith, U. (1999). Interacting Minds: A Biological Basis. *Science, 286*, 1692-1695.

Frith, C., & Frith, U. (2000). The physiological basis of theory of mind: functional neuroimaging studies. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen

(Eds.), *Understanding other minds: perspectives for developmental cognitive neuroscience* . Oxford: Oxford University Press.

Frye, D., & Moore, C. (1991). *Children's theories of mind*. Hillsdale, NJ: Lawrence Erlbaum.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483-527.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., & Frith, C. D. (2000). Reading the mind in cartoons and stories: an fMRI study of 'theory of the mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11-21.

Gergely, G., Nadasdy, Z., Csibra, C., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165-193.

Gergely, G., & Csibra, G. (1997). Teleological reasoning in infancy: the infant's naive theory of rational action. A reply to Premack and Premack. *Cognition, 63*(2), 227-233.

Gopnik, A., & Graf, P. (1988). Knowing how you know: young children's ability to identify and remember the sources of their beliefs. Child Development, 59, 1366-1371.

Griffin, R., Brownell, H., Friedman, O., Winner, E., Ween, J., & Happé, F. (in preparation). Theory of mind and the right cerebral hemisphere: An examination of modularity, executive function, and central coherence. .

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience, 12*, 711-720.

Hains, S. M. J., & Muir, D. W. (1996). Effects of stimulus contingency in infant-adult interactions. *Infant Behavior and Development, 19*(1), 49-61.

Happé, F., Brownell, H., & Winner, E. (1999). Acquired "theory of mind" impairments following stroke. *Cognition*, *70*(3), 211-240.

Happé, F., Gallagher, H. L., Brunswick, N., Fletcher, P., Frith, U., & Frith, C. (in preparation). Brain activity associated with extracting the meaning from verbal and non-verbal narratives. .

Harris, P. (1993). Pretending and planning. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding other minds: perspectives from autism* . Oxford: Oxford University Press.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behaviour. *American Journal of Psychology*, *57*, 243-259.

Hood, B. M., Willen, J. D., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131-134.

Johansson, G. (1973). Visual perception of biological motion and a model of its analysis. *Perception and Psychophysics*, *14*, 202-211.

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science, 1*(2), 233-238.

Johnson, S. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences, 4*, 22-28.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Keil, F. C. (1992). The origins of an autonomous biology. In M. R. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition. The* Minnesota symposia on child psychology, Vol. 25 (pp. 103-137). Hillsdale, NJ, USA: Lawrence Erlbaum.

Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 234-254). New York: Cambridge University Press.

Kluver, H., & Bucy, P. (1939). Preliminary analysis of function of the temporal lobe in monkeys. *Archives of Neurology, 42*, 979-1000.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience, 12*, 48-55.

Lane, R. D., Fink, G. R., Chau, P. M. L., & Dolan, R. J. (1997). Neural activation during selective attention to subjective emotional responses. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 8*(18), 3969-3972.

LeDoux, J. E. (1996). *The emotional brain: The mysterious underpinnings of emotional life.* New York: Simon and Schuster.

Legerstee, M. (1994). Patterns of 4-month-old responses to hidden silent and sounding people and objects. *Early Development and Parenting, 3*(2), 71-80.

Leslie, A. M. (1987). Pretence and representation: the origins of "theory of mind". *Psychological Review, 94*, 412-426.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*(3), 265-288.

Leslie, A. (1994). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfied & S. Gelman (Eds.), *Mapping the mind: domain specificity in cognition and culture* (pp. 119-148). New York: Cambridge University Press.

McFarland, D. (1989). Goals, no-goals, own goals. In A. Montefiore & D. Noble (Eds.), *Goals, no-goals, own goals: A debate on goal-directed and intentional behaviour* . London: Unwin Hyman Ltd.

McGuire, P. K., Silbersweig, D. A., & Frith, C. D. (1996). Functional neuroanatomy of verbal self-monitoring. *Brain, 119*(3):907-17.

McLeod, P., Dittrich, W., Driver, J., Perrett, D., & et al. (1996). Preserved and impaired detection of structure from motion by a "motion-blind" patient. *Visual Cognition, 3*(4), 363-391.

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

Michotte, A. (1963). *The perception of causality.* Andover: Methuen.

Myers, R. E., Swett, C., & Miller, M. (1973). Loss of social group affinity following prefrontal lesions in free-ranging macaques. *Brain Research, 64*(257-269).

Paus, T., Koski, L., Caramanos, Z., & Westbury, C. (1998). Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: A review of 107 PET activation studies. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience, 9*(9), R37-R47.

Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology, 146,* 87-114.

Perrett, D., Harries, M., Mistlin, A., Hietanen, J., Benson, P., Bevan, R., Thomas, S., Oram, M., Ortega, J., & Brierley, K. (1990a). Social signals analyzed at the single cell level: someone is looking at me, something touched me, something moved! *International Journal of Comparative Psychology, 4,* 25-55.

Posner, M. I., & Dehaene, S. (1994). Attentional networks. *Trends in Neurosciences, 17*(2), 75-79.

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61,* 973-983.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences, 1,* 515-526.

Premack, D. (1988). How to tell Mae West from a crocodile. *Behavioral and Brain Sciences*. 11(3), pp. 522-525.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*(1), 1-16.

Premack, D., & Premack, A. J. (1995). Intention as psychological cause. In D. Sperber, D. Premack, & A. Premack (Eds.), *Causal cognition: A multi-disciplinary debate*. Oxford: Clarendon Press, pp. 185-199.

Premack, D., & Premack, A. (1997). Motor competence as integral to attribution of goal. *Cognition, 63*, 235-242.

Prior, M., & Bradshaw, J. (1979). Hemispheric functioning in autistic children. *Cortex, 15*, 73-81.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience, 18*, 2188-2199.

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science, 277*(5328), 968-971.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*, 12-21.

Rochat, P., Morgan, R., & Carpenter, M. (1997). Young infants' sensitivity to movement information specifying social causality. *Cognitive Development, 12*, 537-561.

Sabbagh, M. (2000). Communicative intentions and language: evidence from right-hemisphere damage and autism. *Brain and Language, 70*(1), 29-69.

Schlottmann, A., & Surian, L. (1999). Do 9-month-olds perceive causation-at-a-distance? Perception, 28(9), 1105-1113.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or 'representational momentum.'. *Current Biology, 10*(1), 16-22.

Shields, J., Varley, R., Broks, P., & Simpson, A. (1996). Hemispheric function in developmental language disorders and high-level autism. *Developmental Medicine and Child Neurology, 38*(6), 473-486.

Sorce, J., Emde, R., Campos, J., & Klinnert, M. (1985). Maternal emotional signalling: its effect on the visual cliff behavior of 1-year-olds. *Developmental Psychology, 21*, 195-200.

Spelke, E., & Van de Walle, G. (1993). Perceiving and reasoning about objects: insights from infants. In N. Eilan, W. Brewer, & R. McCarthy (Eds.), *Spatial representation*. New York: Blackwell.

Spelke, E., Phillips, A., & Woodward, A. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon Press.

Stone, V., Baron-Cohen, S., & Knight, K. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10*, 640-656.

Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology, 30*(5), 639-650.

Tomasello, M. (1999). Having intentions, understanding intentions, and understanding communicative intentions. In P. D. Zelazo & J. W. Astington (Eds.), *Developing theories of intention: Social understanding and self control* (pp. 63-75). Mahwah, NJ: Lawrence Erlbaum.

Vaina, L. M., Lemay, M., Bienfang, D. C., & Choi, A. Y. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. *Visual Neuroscience, 5*(4), 353-369.

Varley, R., & Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. *Current Biology, 10*(12), 723-726.

Vygotsky, L. (1979). *Mind in society*. Cambridge, MA: Harvard University Press.

Wellman, H. (1990). *The child's theory of mind.* Cambridge, MA: MIT Press.

Wellman, H. M., & Bartsch, K. (1994). Before belief: Children's early psychological theory. In C. Lewis (Ed.), *Early understanding of mind: Origins and development*. (pp. 331-354). Hove, Sussex: Lawrence Erlbaum.

Wellman, H. M., Cross, D., & Watson, J. (in press). Meta-analysis of theory of mind development: The truth about false-belief. *Child Development, 72*, 3: 655-684.

Whalen, P. (1999). Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science, 7*, 177-187.

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

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language, 62*(1), 89-106.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1-34.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256*, 1327-1331.

Young, A., Aggleton, J., Hellawell, D., Johnson, M., Broks, P., & Hanley, J. (1995). Face processing impairments after amygdalectomy. *Brain, 118*, 15-24.

Zaitchik, D. (1991). Is only seeing really believing? Sources of the true belief in the false belief task. *Cognitive Development, 6*(1), 91-103.

Zelazo, P. D., Astington, J. W., & Olson, D. R. (Eds.). (1999). *Developing theories of intention: Social understanding and self-control.* Mahwah, NJ: Lawrence Erlbaum.