

Pretense and Representation: The Origins of "Theory of Mind"

Alan M. Leslie

Medical Research Council, Cognitive Development Unit, University of London

One of the major developments of the second year of human life is the emergence of the ability to pretend. A child's knowledge of a real situation is apparently contradicted and distorted by pretense. If, as generally assumed, the child is just beginning to construct a system for internally representing such knowledge, why is this system of representation not undermined by its use in both comprehending and producing pretense? In this article I present a theoretical analysis of the representational mechanism underlying this ability. This mechanism extends the power of the infant's existing capacity for (primary) representation, creating a capacity for *metarepresentation*. It is this, developing toward the end of infancy, that underlies the child's new abilities to pretend and to understand pretense in others. There is a striking isomorphism between the three fundamental forms of pretend play and three crucial logical properties of mental state expressions in language. This isomorphism points to a common underlying form of internal representation that is here called *metarepresentation*. A performance model, the *decoupler*, is outlined embodying ideas about how an infant might compute the complex function postulated to underlie pretend play. This model also reveals pretense as an early manifestation of the ability to understand mental states. Aspects of later pre-school development, both normal and abnormal, are discussed in the light of the new model. This theory begins the task of characterizing the specific innate basis of our commonsense "theory of mind."

Pretending ought to strike the cognitive psychologist as a very odd sort of ability. After all, from an evolutionary point of view, there ought to be a high premium on the veridicality of cognitive processes. The perceiving, thinking organism ought, as far as possible, to get things right. Yet pretense flies in the face of this fundamental principle. In pretense we deliberately distort reality. How odd then that this ability is not the sober culmination of intellectual development but instead makes its appear-

ance playfully and precociously at the very beginning of childhood.

Reality-oriented play, which responds to an object's actual properties or expresses knowledge of its conventional use, raises many interesting problems. But pretense poses deeper puzzles. How is it possible for a child to think about a banana as if it were a telephone, a lump of plastic as if it were alive, or an empty dish as if it contained soap? If a representational system is developing, how can its semantic relations tolerate distortion in these more or less arbitrary ways? Indeed, how is it possible that young children can disregard or distort reality in any way and to any degree at all? Why does pretending not undermine their representational system and bring it crashing down?

In this article I shall deal with the significance of the emergence of pretense in terms of the infant's capacity for internal representation.¹ To conceptualize representation, an information-processing, or cognitivist, approach is taken (Fodor, 1976; Marr, 1982). In the course of this, pretense will acquire a new theoretical definition. The resulting model has implications for both normal and abnormal development.

Current Approaches

In one of his major works on infancy, Piaget (1962) argued that pretend play is an extreme form of assimilation. A present object that is only vaguely comparable to an absent one can evoke a mental image of it and be assimilated to it, resulting in the creation of a symbol. The ability to pretend depends on this capacity to represent absent objects and situations. This capacity is said to emerge during the second year of life.

I am deeply indebted to Uta Frith, John Morton, and Josef Perner for long hours of discussion on the topic of this article and for detailed comments on earlier versions. I am also grateful to Annette Karmiloff-Smith, Rick Cromer, John Macnamara, Jean Mandler, Jerry Bruner, and four anonymous reviewers for their comments on earlier versions of this article as well as to many people, especially Heinz Wimmer, Henry Wellman, John Flavell, and Susan Carey, for helpful comments on talks based on this article. The flaws that remain are entirely my responsibility. And last, but by no means least, I should like to thank my darling Sarah-Jane for her brilliance, fun, and 4 years of hard work down at Mrs. McDog's Farm as her dazzling capacity for pretense unfolded.

This article began as a poster at the British Psychological Society Developmental Conference, Oxford, 1983, and has developed through a number of talks including the Institute of Psychiatry, London, 1984; Association Pour la Recherche sur l'Autisme et la Psychose Infantile, Paris, 1984; Zangwill Club, Cambridge, 1986; the International Conference on Developing Theories of Mind, Toronto, 1986; the Workshop on Children's Early Concept of Mind, Oxford, 1986; and the Advanced Course on Symbolism and Knowledge, Fondation Archives Jean Piaget, Geneva, 1986.

Correspondence concerning this article should be addressed to Alan M. Leslie, Medical Research Council, Cognitive Development Unit, University of London, 17 Gordon Street, London WC1H 0AH, England.

¹ Many of the issues dealt with in this article are discussed at greater length in Leslie (in press-b).

For Piaget, early pretense symbolizing develops in a hierarchical fashion from familiar self-directed actions performed out of context, through the symbolic identification of one object with another, to increasingly complex symbolic combinations (Piaget, 1962). This account has been elaborated by McCune-Nicolich (1981), who suggested that late in the second year a fundamental shift in the child's symbolic play "allows games to be generated mentally," which requires "the coordination of at least two representational structures" (p. 787).

Fischer (1980; Fischer & Pipp, 1984) has also presented a hierarchical account of these developments, but from a different theoretical viewpoint. In Fischer's behaviorist skill theory, this shift in pretend results from the coordination of two (or more) sensorimotor systems. Such a combination defines an elementary representation that can then show up in the infant's pretend play. The child can now adopt a behavioral role (e.g., of doctor) or treat an object as an agent (Fischer, 1980; Watson & Fischer, 1977, 1980).

Vygotsky (1967) placed great emphasis on the affective aspects of pretense. Imaginative play "originally arises from action" (p. 8) and from generalized "unsatisfied desires" (p. 9). Play teaches the child "to sever thought . . . from object" (p. 12) and provides a means for developing abstract thought.

Fein (1975) proposed that pretense can be thought of as involving *transformations*. By transformation she meant a process that mediates the selecting of some features of an immediate object or situation and the ignoring of others, comparing such subsets with others drawn from memory, and thereby coming to see an analogy between disparate entities. Such transformations could involve role shifts, animating inanimates and substituting one object for another.

All these views have influenced recent empirical research on the early development of pretend play. Several excellent reviews of this work have appeared recently (Fein, 1981; McCune-Nicolich, 1981; McCune-Nicolich & Fenson, 1984). Because of a general consensus on basic theoretical questions, effort has concentrated on documenting certain sorts of behavior change. Three main developmental trends have been studied: *decentration*—a move from self-directed to other-directed pretend (Belsky & Most, 1981; Corrigan, 1982; Fein & Apfel, 1979; Fenson & Ramsay, 1980; Lowe, 1975; McCune-Nicolich, 1981; Nicolich, 1977; Watson & Fischer, 1977); *decontextualization*—the use of less and less realistic objects as symbols (Bretherton, O'Connell, Shore, & Bates, 1984; Cole & LaVoie, 1985; Elder & Pederson, 1978; Fein, 1975; Field, De Stefano, & Koewler, 1982; Golomb, 1977; Jackowitz & Watson, 1980; Pederson, Rook-Green, & Elder, 1981; Ungerer, Zelazo, Kearsley, & O'Leary, 1981; Werner & Kaplan, 1967); and *integration*—the ability to combine schemes into sequences (Fenson & Ramsay, 1980, 1981; McCune-Nicolich & Fenson, 1984; Nicolich, 1977). I shall take advantage of the fact that these studies have recently been reviewed elsewhere and move on to consider a new approach.

A Cognitivist Approach

The theory I want to sketch is primarily concerned with underlying mechanisms and with the information-processing tasks these mechanisms have to perform in generating pretense.

The aims of the present approach are thus different from and complementary to previous approaches. Piaget was fundamentally interested in the nature of intelligence and how it changes with development, and his interest in pretense was subordinate to this. Fischer provided a behavioral analysis of skills and levels of skills bearing on the developing complexity of behaviors shown in pretend. McCune-Nicolich was concerned with the different forms of symbolizing in early development and with working out their operational definitions.

The present approach uses the computational metaphor and seeks to develop a mechanistic theory of the ability to pretend. I am led eventually to a major theoretical distinction, in terms of underlying mechanisms, between pretense and other forms of symbolic play. To arrive at such a distinction it is necessary to examine the special properties of internal representations required for pretense. But first one should look at intuitive grounds for distinguishing pretense.

Pretense and Acting as If

Error Acting as If

Pretending is one kind of "acting as if" something is the case when it is not. Another kind that needs to be distinguished is "acting in error." There are many ways in which one can come to do something in error and so act as if something were the case when it is not. If I jump up suddenly because I mistakenly think I see a spider on the table, I act as if a spider were there. But I certainly do not *pretend* a spider is there. Likewise, there are many ways in which young children could come to act in error. For example, they might simply make a mistake (and think the lump of wood is a lump of soap), or not be able to discriminate (e.g., shells from cups), or not possess a relevant conceptual distinction (e.g., pillows vs. cushions). In none of these cases would we say they are pretending. Pretend is a special case of acting as if where the pretender correctly perceives the actual situation.

McCune-Nicolich (1981) called this *double knowledge*. This double knowledge has to be operating at the time the pretense takes place, because most of the time the child may be able to discriminate one kind of object from another but still, on a particular occasion, fail to do so. Thus for pretense to occur it is essential that the pretender actually be "telling the difference" at the time the pretend takes place. This is something my model must capture.

Functional Play

Huttenlocher and Higgins (1978) posed a set of difficult questions for anyone interested in early pretend. Their mode of argument was to adopt a skeptical position with regard to claims that infants show symbolic activity. Their basic point was this. Suppose an infant is observed setting out a tea set in the conventional way or pushing a toy car along the ground while making "brrrrmm" noises. Can we be sure that the child is really pretending? Perhaps the child is simply demonstrating knowledge of the conventional use of objects. The toy tea set is a pretend replica to us, but to the young child they may just be ordinary objects with socially conventional uses. Even the sound effects in the toy car example are not conclusive evidence that the child

is pretending. Again, such sound effects may simply be for the child part of the conventional use of this object. If so, this functional play (Piaget, 1962; Ungerer & Sigman, 1981) certainly demonstrates sophisticated knowledge on the child's part and is a kind of acting as if, but it does not constitute pretending anything. Similar sorts of arguments can be made for the other common accompaniments of pretend play, like knowing looks, smiles, and exaggerated gestures (Bretherton et al., 1984; McCune-Nicolich, 1981; Piaget, 1962; Rosenblatt, 1977).

Huttenlocher and Higgins argued that the only really clear evidence for pretense is provided when the child's verbalization reveals a symbolic or pretense attitude prior to his or her act (1978, p. 124). But one could continue with the skeptic's role even here and insist that prior verbalization can be explained away, just as subsequent verbalization can, as part of a specifically learned response.

The point to take from Huttenlocher and Higgins's examples, however, is not that there is an inherent contradiction between specific learning and pretense (they are quite compatible), but simply that one and the same piece of behavior can, in principle, be produced under different internal states. This is what makes it so extremely difficult to produce a watertight behavioral definition of pretense.

To help sort out pretense from sophisticated functional play, one must consider whether at least one of three things has happened cognitively. These correspond to three fundamental forms of pretense: object substitution, attribution of pretend properties, and imaginary objects. Has one object been made to stand in for another, different object? (Has the child pretended a shell was a cat?) Has a pretend property been attributed to an object or a situation? (Has the child pretended the dolly's [clean] face is dirty?) Has the child invented an imaginary object? (Has the child pretended that a spoon is there when it is not?) If we have reason to believe that the child's play involves any one of these, we have reason to believe the child is pretending. Otherwise, we have no compelling reason to assume pretense.

This, then, is how I shall use the term *pretense* in this article. As I shall show, there are important theoretical reasons for such a narrow definition and for thus excluding functional play. It seems that play exhibiting pretense forms (in this sense) emerges roughly between 18 and 24 months of age for most children. This corresponds to McCune-Nicolich's (1981) shift to "mentally generated pretend."

In both functional play and error acting as if, the as-if component really only exists from the observer's point of view. From the actor's point of view, the actions are serious. But in pretense, the actor is acting as if from the actor's point of view as well. These considerations make important demands on any competence theory of pretense.

A Metarepresentational Theory of Pretense

What I mean by representation will, I hope, become clear as the discussion progresses. It has much in common with the concepts developed by the information-processing, or cognitivist, approach to cognition and perception (Chomsky, 1980; Dennett, 1983; Fodor, 1976; Haugeland, 1978; Mandler, 1983; Marr, 1982; Rock, 1983; Ullman, 1980). In particular, I shall

try to explain the external symbolic activity of pretending in terms of properties of the internal mental representations that underlie it.

Representation in Infancy

The basic evolutionary and ecological point of internal representation must be to represent aspects of the world in an accurate, faithful, and literal way, in so far as this is possible for a given organism. Such a *basic* capacity for representation can be called a capacity for *primary representation*. Primary representation is thus defined in terms of its direct semantic relation with the world. Its being literal and "sober" in representing the world determines its usefulness relative to the needs of the organism.

Assume that infants possess a capacity for primary representation from the outset of development. Of course this general statement does not say what aspects of the world are representable by the infant, nor with what degree of adequacy, nor how the capacity might develop. These are questions for detailed investigation and are beyond the scope of this article. But one major manifestation of primary representational capacity is the infant's perceptual abilities. There are an increasing number of studies that approach infant perception of objects, people, events, and scenes from the point of view of this representational capacity (e.g., Baillargeon, 1986; Bower, 1974, 1978; Leslie, 1982, 1984, 1986; Leslie & Keeble, 1987; Mandler, 1983, 1984; Meltzoff, 1981; Spelke, 1982).

Perception of the world and the things in it are a major source of the infant's stored knowledge. Such encyclopedic knowledge also forms structures of primary representation (cf. Keil, 1984). Again, the design principle for these representations is that they represent situations seriously and literally. The question I will now address is, "Could primary representation account for the emergence of pretense?"

Representational abuse. As already seen, one important difference between pretend and error acting as if is that in pretend there are two simultaneous representations of the situation. One representation is for how the situation is actually perceived, whereas the other represents what the pretense is. But this is not enough. The pretense relates to the actual situation in specific ways. It is *this* banana that I pretend is a telephone; it is *this* doll's face that I pretend is dirty. This must mean that pretend representations relate in specific ways to primary representations. The problem for current theory is to say what exactly this relation is.

Having simultaneous representations may suggest that pretense requires an ability to coordinate two primary representations. The emergence of pretense would then depend on the emergence of this coordinating ability. This idea is reminiscent of Piaget's notion of simultaneous distorting and generalizing assimilation in pretend (1962, p. 103), of McCune-Nicolich's coordination by internal definition of two representational structures in object-substitution pretend (1981, p. 787), of Huttenlocher and Higgins's assigned linkages between (internal) symbol elements and mental entries (1978, p. 109), of Fein's pretend transformations determining the intersection of two feature lists (1975, p. 293), and of Fischer's coordination of sensorimotor sets (1980, pp. 490-493).

If both representations are *primary*, however, then both have a literal meaning. And because the pretense relates to the same actual situation in the serious cognition, both representations have to be representations of the same situation. But typically the pretense representation contradicts the primary representation. Consequently, something has to give here.

Consider more closely how two primary representations would relate to one another in pretense. One may assume that the pretend representation uses a different code from that of the perceptual situation. Call these the symbolic and sensory codes, respectively. An item (CUP) in the symbolic code may represent the class of cups by having a *reference linkage* with percepts of the right sort (Huttenlocher & Higgins, 1978). Thus, having such a symbolic item, the infant could now recognize a member of a category of *cups* and distinguish them from noncups (e.g., *shells*). Pretending that a shell is a cup would involve establishing a reference link between the symbolic item CUP and an item \$shell\$ in the sensory code.

However if the reference link between symbolic CUP and sensory \$shell\$ has the same status as that between symbolic CUP and sensory \$cup\$, then CUP will change its meaning. CUP now refers to \$shell\$ as well as to \$cup\$. The more the child pretend plays, the more symbolic code items will change their meaning and the more chaotic and useless the symbolic code will become as a result. Instead of developing a code with greater precision and definiteness, pretending will make it more and more amorphous. And paradoxically, the more the child pretends the less able to pretend he or she will become! After all, if CUP comes to mean \$shell\$, the child can hardly pretend again that a shell is a cup—now, for the child, it really is a cup. This can be called the problem of *representational abuse*.

Piaget's and McCune-Nicolich's proposals raise similar problems of representational abuse. In discussing an infant's pretense that her mother's hair was a cat, Piaget (1962, p. 126) said that a "symbolic identification (hair = cat)" precedes the child's pretend act, whereas McCune-Nicolich (1981, p. 787), giving the example of pretending a stick is a horse, said that "an internal definition (stick = horse) is implied." Clearly, if *identification*, *definition*, and *=* are taken at face value, arbitrary changes of meaning must occur. If a representation of a stick is defined as a horse, it changes its meaning. A stick becomes a horse by definition. Representational abuse will undermine important distinctions and create spurious commonalities just at the point when it is assumed that systematic representations are developing for the first time. Obviously, taking these proposals at face value in this way was not what was intended. But this still leaves unanswered the question of what exactly the relation is between pretend and primary representations.

More abuse. The problem of representational abuse does not affect only reference links and thus object-substitution pretend. It also affects the other two basic kinds of pretense: attribution of pretend properties and imaginary objects. To effect the attribution of pretend properties, the symbolic code cannot consist only of isolated items or unstructured lists of isolated items. It must allow propositionlike expressions as well. Accordingly, pretend attribution of properties cannot place links between corresponding items in the two codes as if the expressions were just lists. Pretending to wash doll's dirty face with a clean cloth

differs from pretending to wash doll's clean face with a dirty cloth. This requires links not just between individual items but between whole expressions. Such a theory would need truth links as well as reference links.

Representational abuse can strike such propositionlike expressions in at least two ways. First, one may pretend that a red car is yellow. The pretend representation **this car is yellow** applies to a situation in which the car is red; this extends and changes the meaning of **yellow**. However, one would know that there was something odd about this representation only by looking at the external situation to which it applied. But another kind of abuse applies internally to the expression. For example, in a pretend representation such as **this empty cup contains water**, one can tell that abuse has occurred without looking further than the expression itself. So **empty** would now include situations in which cups contain water as well as situations in which cups contain nothing. Perhaps even worse, one could no longer infer from **the cup is empty** to **the cup contains nothing**. Its meaning internal to the system has been undermined as well.

Accounts based on reference linkages and definitions between primary representations will have additional difficulties with imaginary-object pretend because there is no particular thing in the perceptual situation that the symbolic item could link to or be defined in terms of.

Pretense affects the normal reference, truth, and existence relations of the representations it uses. These relations become highly deviant. Any primary representational system affected would quickly be undermined by arbitrary meaning changes. To prevent this, pretend representations must somehow be marked off, or "quarantined," from primary representations. Indeed, so deviant are the reference, truth, and existence relations of pretend representations that it begins to seem unlikely that they are primary representations at all.

In talking about these relations, bear in mind that reference, truth, and existence are really relations holding between primary representations and the world and not, therefore, links between one primary representation and another. In the case of the pretend representation, these relations appear either to be suspended altogether or to hold only at one remove through primary representation. Does yet another code need to be postulated, then, one specific to pretense—a very symbolic code perhaps? And what sort of relation is there between these quarantined and primary representations? Before pursuing these questions, one must consider a quite different reason for the need for quarantining pretend.

Understanding pretense in others. Early pretense is not always undertaken in solitude but can form part of infant social interaction. This is shown not only by everyday observation but also by experimental studies that require the infant to imitate various kinds of adult-modeled pretend play (e.g., Bretherton et al., 1984; Jackowitz & Watson, 1980; Watson & Fischer, 1977) and by studies by Dale (1983; Dunn & Dale, 1984) showing that 2-year-olds and even 18-month-olds can share pretend games with older siblings. They may even show more advanced forms of play in shared than in isolated pretense (e.g., adoption of reciprocal pretend roles). The infant must therefore in some way understand pretending in others. Spelling out the nature of

this understanding is another important objective for a cognitivist theory of pretense.

But suppose for a moment that the 18-to-24-month-old had no such understanding. Such an infant must engage in observations of other's use of objects because for some months he or she has been demonstrating knowledge of such use in his or her functional play (see also Abravanel & Gingold, 1985). So what would such an infant make of someone pretending, for example, that a banana is a telephone if he or she could in no way understand pretense in others? The infant would only be capable of representing the activity in a literal way; thus, he or she might be puzzled by hearing and seeing mother talking to a banana. There again, the infant might be no more puzzled by this than by seeing mother talking seriously to a telephone. Presumably infants eventually come to understand what telephones really are at least partly by representing and storing away information about people talking into telephones. Such seriously construed information will be useful in providing clues as to the real properties and functions of telephones. But treating the information from the pretend context in the same way will be highly misleading. The infant would end up with some funny ideas about either bananas or mother or both.

It would be useful for the infant to have some way of marking information from pretend contexts to distinguish it from information from serious contexts. It would be more useful if the infant had some way of representing that someone was pretending, what the pretend was, and how the pretend related to the literal acts. If an infant could do all this, he or she might be able to join in the fun and elaborate on the pretense begun by someone else. Infants late in the second year seem somehow to be capable of this.

There is a parallel here with the quarantining of pretend representations in order to avoid the problem of representational abuse. There, the need was to preserve the integrity of the representational system. Here, the need is to preserve the integrity of the infant's developing knowledge of the world. This parallel may point to the existence of a common underlying mechanism. Such a mechanism would provide a single explanation for the ability to pretend and for the ability to understand pretense in others. In the next section I outline ideas that will eventually explain why representational abuse does not occur, why multiple codes are not needed for pretending, and why understanding pretense in others is simply part and parcel of being able to pretend oneself.

Metarepresentation and Pretense

I will now bring a major feature of the present theory into focus. The emergence of pretense is not seen as a development in the understanding of objects and events as such, but rather as the beginnings of a capacity to understand cognition itself. It is an early symptom of the human mind's ability to characterize and manipulate its own attitudes to information. Pretending oneself is thus a special case of the ability to understand pretense in others (someone else's attitude to information). In short, pretense is an early manifestation of what has been called *theory of mind* (Premack & Woodruff, 1978).

An isomorphism. I have identified three basic kinds of pretending, each linked with its own form of "abuse": (a) object

substitutions (abuse by deviant reference), (b) attributions of properties (abuse by deviant truth), and (c) imaginary objects (abuse by deviant existence). I now want to point out a striking similarity between these properties of pretend play and the logical properties of sentences containing *mental state terms*. By mental state terms I mean words such as *believe*, *expect*, and *want*. Philosophers have long recognized that from a logical point of view, propositions behave strangely when placed in the context of such terms.

Three properties have commonly been identified. First, the reference of terms in such embedded propositions becomes *opaque* (Quine, 1961). For example, "the prime minister of Britain" and "Mrs. Thatcher" refer at the time of this writing to the same person. Therefore, anything asserted about the prime minister of Britain, if true, must be true of Mrs. Thatcher as well (and, likewise, false for one, false for the other). If it is true that the prime minister of Britain lives at No. 10 Downing Street, then it must be true that Mrs. Thatcher lives at No. 10 Downing Street. But put this proposition in the context of a mental state term and this no longer holds. Thus "Sarah-Jane believes that the prime minister of Britain lives at No. 10 Downing Street" in no way entails the truth (or falsehood) of "Sarah-Jane believes Mrs. Thatcher lives at No. 10 Downing Street." In a mental state context one can no longer "look through" terms to see what they refer to in deciding such issues. The mental state term suspends normal reference relations. Quine (1961) called this *referential opacity*.

Second, propositions involving mental state terms do not logically imply the truth (or falsehood) of propositions embedded in them. Thus "John believes the cat is white" says nothing about whether or not the cat really is white. Again, one cannot look through the embedded proposition to the world.

Third, assertions involving mental state terms do not logically entail the existence or nonexistence of the things mentioned in the embedded proposition. Thus "The king of France is bald" is a strange statement because it logically implies or presupposes the existence of a French king. It is just as hard to say it is false because that would still entail the king of France's existence. But "Jacqueline believes the king of France is bald" has no such problems. The existence is not entailed.

Thus for each of these semantic properties of mental state expressions there appears to be a corresponding basic form of pretense: (a) referential opacity—object substitution (deviant reference pretend); (b) nonentailment of truth (or falsehood)—attribution of pretend properties (deviant truth pretend); and (c) nonentailment of existence (or nonexistence)—imaginary object (deviant existence pretend). I suggest that these connections are not coincidental. At the very least, mental state expressions can provide a model with which to characterize the representations underlying pretend play. But I want to go beyond this to explain why an isomorphism between mental state expressions and pretense exists. I shall do this by positing an underlying form of internal representation that possesses these semantic properties. I shall then argue that mental state expressions and pretense both depend cognitively on these representations and therefore inherit their properties.

Decoupling. To the organism who entertains them, primary representations are by definition *transparent*—that is, they directly represent objects, states of affairs, and situations in the

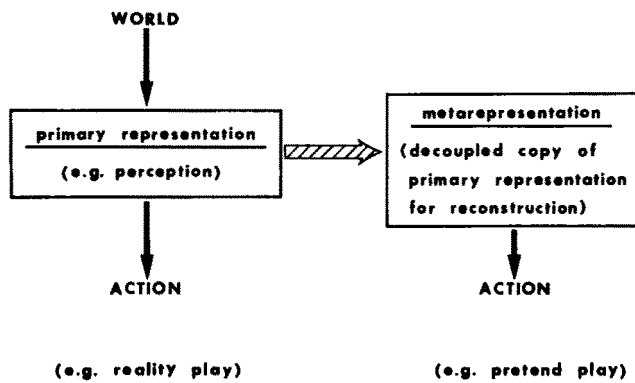


Figure 1. A general model for pretend play.

world. Pretend representations, by contrast, are *opaque*, even to the organism who entertains them. They are in effect not representations of the world but representations of representations. For this reason I shall call them second order or, borrowing a term from Pylyshyn (1978), *metarepresentations*.

The opacity of metarepresentations explains how representational abuse is avoided. The basic feature of my model is the creation of a pretense by the copying of a primary expression into a metarepresentational context. This second-order context in effect gives a report or quotation of the first-order expression. In doing this, it renders *opaque* the expression that was previously transparent. Its reference, truth, and existence relations are suspended while it appears in this context. Using an appropriately mechanistic metaphor, one can say that the metarepresentational context *decouples* the primary expression from its normal input-output relations. Meanwhile the original primary representation, a copy of which was raised to a second order, continues with its definite and literal reference, truth, and existence relations. It is free to continue exerting whatever influence it would normally have on ongoing processes (see Figure 1).

Already the model expresses a clear difference between pretending about a situation and simply being in error about it. In the error case the primary representation (of the situation) is simply not accurate from an objective point of view. No metarepresentation is involved. Bear in mind that from an observer's point of view, all the representations of another organism are opaque. Whereas for me, this infant only thinks he or she sees a cup before him or her (I know it's really a shell), for the infant, his or her (primary) representation, *here is a cup*, is completely transparent. To the infant, *cup* is not a feature of his or her representation waiting to be interpreted, it is something in the world. Pretense, however, uses metarepresentations. These are opaque, even from the subject's point of view, and have to be actively interpreted each time they are used. Meanwhile, the primary system continues unabused.

Another feature of this model is that a single code will suffice—the code of primary representation. For metarepresentations, it needs to be extended by adding only two new items. First, the opacity of decoupled expressions must be marked as such. I shall employ the device of enclosing such expressions in quotation marks. Again, this is borrowed from lan-

guage. Sentences enclosed within quotations are also rendered opaque. So for example, in reports like John said, "The king of France is bald," the quotations mark the embedded expression as toothless, suspending its normal service, in much the same way as mental state contexts. This fact was used by the logician Carnap (1947) in his *quotation theory* of mental state expressions. Church (1950) subsequently showed that this account of the logic of mental state expressions was fatally flawed. However, the reason for this was that Carnap applied the idea to sentences in natural language instead of to an underlying canonical notation and so ran into problems connected with the surface forms of the various languages. Fodor (1981, chap. 7) has recently argued that these problems do not arise if the quoted expression is interpreted not as a sentence but as an expression in a system of internal mental representation (see also Jackendoff, 1983, and for critical discussion, Barwise & Perry, 1983). It is in this guise that I adopt the quotation approach.

Suppose we start with a representation of the current perceptual situation, for example, *this is a banana*. This is decoupled to "*this is a banana*." Because its normal semantics has been suspended, the expression can be manipulated freely without fear of abusing the normal representational system existing outside this context. So, for example, it will be possible to transform the expression "*this is a banana*" into "*this banana is a telephone*" while disregarding its interpretation. An expression like *this banana is a telephone* could not arise in primary representation. Such nonsense violates the basic design principle of primary representation that it represent in a literal fashion. Decoupling, however, allows such expressions to be treated and worked on as purely formal objects.

Form of metarepresentations. Pretend representations do not pose the problem of abuse precisely because their semantics is suspended. The quarantining of information from pretense in others can be handled in the same way, that is, by decoupling. Here some way of representing who the decoupled expression belongs to is needed. And again one can turn to natural language for a model. Language has its mental state terms that denote relationships between agents and opaque propositions. In fact, the verb *pretend* is just such a term. I can add to my model formal elements that play a corresponding role in underlying mental representation. The second extension to primary code will be an item, **PRETEND**, representing an informational relation. This relation will hold between whatever primary representations of agents (e.g., mother) the infant has and decoupled expressions. Pretend metarepresentations might thus have the general form: **Agent-Informational Relation—"expression."** **Agent** ranges over, for example, persons and self, whereas "**expression**" can be filled by any decoupled representation. Two points need to be made about informational relations at this point. First, **PRETEND** is not equivalent to the English "pretend" because it does not itself decouple associated expressions—that is the job of the decoupling marks. Separating these components has significance, as will be shown later. Second, by implication there are other informational relations that the infant is or will become able to represent (e.g., **UNDERSTANDING** a message to be communicated). This, too, will be discussed later.

So far, I have sketched the bare outlines for a competence

theory of infant pretense. I eventually want to embody this in a performance-oriented model that can be related in greater detail to the infant's behavior. But first, the competence problem, or the "semantics of pretense," must be looked at more closely.

Semantics of Pretense

Bateson (1972) addressed questions related to some of those discussed here. He raised the paradoxical nature of some play behavior. Such play actions in effect signal their own lack of normal meaning. Bateson suggested that a message "This is play" establishes a sort of psychological "frame" embodying this paradox. He illustrated with this example:

<p style="text-align: center;">All statements within this frame are untrue.</p> <p style="text-align: center;">I love you. I hate you.</p>
--

Frames in Bateson's sense have to do with delimiting messages in a certain way to aid in their interpretation by the receiver. Messages within the frame are to be interpreted one way, messages outside the frame another way. Bateson believed that in some sense "the psychological frame has some degree of real existence" (1972, p. 186).

There are a number of connections between the ideas presented here and Bateson's frames. The notion of quarantining could be viewed in this way. Even the notion of metarepresentation has certain points of contact, though Bateson presented his arguments in terms of the theory of logical types and so missed what is central to the present theory, namely, the logic of opacity and thus the connection between pretending and theory of mind.

For expository purposes, in this section I shall borrow a version of Bateson's frames notation. By *the semantics of pretense*, I mean the relation of a pretend metarepresentation with a primary representation of the current actual situation. Given that infants and very young children are under consideration, I can restrict myself to the current situation because I doubt their pretend will normally relate to anything else. Here is a hypothetical frame *perceived situation* that might reflect current perceptual processing.

<p>Perceived Situation</p> <p><i>here is a table_i</i> <i>the table_i is dry</i></p> <p><i>here is a cup_j</i> <i>the cup_j is empty</i> <i>the cup_j is red</i> <i>the cup_j is on the table_i</i></p>
--

Another frame, *pretend situation*, might be derived from the *perceived situation* above.

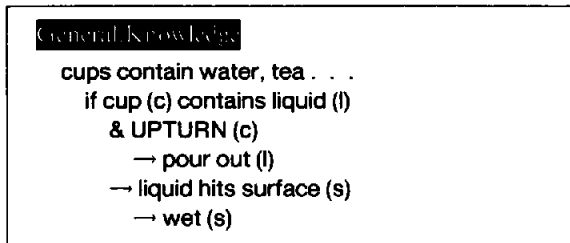
<p>Pretend Situation</p> <p>I PRETEND <i>"this empty cup_j, it_j contains tea"</i></p>

The italicized elements in pretend situation were raised from perceived situation. Hence, it is natural to interpret this part of the decoupled expression as relating to its unraised counterpart. However, in principle it is possible that a pretend could be constructed using this raised expression in which it did not relate to an unraised counterpart. Decoupled expressions no longer have an automatic reference. Indeed, I shall not use the term *reference* in connection with them because they do not relate directly to the world. I shall instead use the term *anchoring*. Decoupled expressions do not refer to objects, then, they are anchored to parts of primary representations. This is not automatic, but needs to be specially stipulated. It will be assumed, nonetheless, that where a decoupled expression matches a primary expression in perceived situation, the former will likely get anchored to the latter. I have represented such anchoring in the examples by using subscripts. So in the pretend situation example I pretend that this particular empty cup contains tea. It is deviantly true of this empty cup that it contains tea.

In effect, decoupling allows certain parts of the expression to act as variables that can be temporarily bound to parts of primary representations. The appropriate predicates in pretend situation then become deviantly true of perceived situation as a whole. Thus inferences can be made without abuse based on "*it_j contains tea_j*," with "*empty cup_j*" being read effectively as a variable. However, anchoring can still take advantage of the fact that an *empty cup* is formally specified (I don't want to pretend that a tea-filled cup contains tea). On the one hand, representational abuse of the inferential sort is avoided, whereas, on the other hand, the pretense is related correctly to the actual situation. This solution also accords with the intuition that in pretending that this empty cup contains tea I am not pretending that the cup is both empty and contains tea.

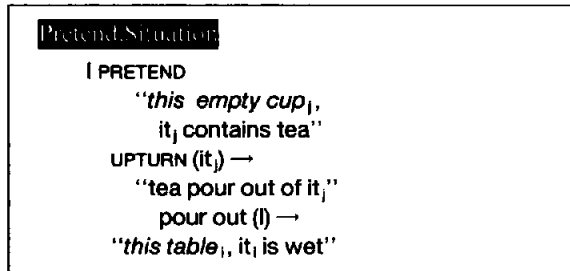
Inferences in pretense. I have raised the question of drawing inferences from pretend representations. Some simple events may already be understood by 18-month-olds with reference to elementary implication rules (Keil, 1979). Keil, using a surprise paradigm, found evidence suggesting that such children could infer that if an object is not supported, it will collapse. Assume, for the sake of argument, that 18-to-24-month olds can command some such rules in their primary dealings with the world and that some of their knowledge is represented in this way. Also assume that at some point in development the child becomes able to apply such rules in pretending and that this will be an important means for elaborating pretense. So, for example, having pretended that a cup contains tea, the child can apply the rule(s) that say(s) that a container of liquids, if upturned, will spill its contents, which will then make wet whatever surface the liquid falls onto. So the child upturns the cup and pretends that the tea pours out and that the table becomes wet.

We could put such inference rules and other general information into another box called *general knowledge*.



Such general knowledge is developed (and will continue to develop) in direct relation to understanding the world and is thus part of primary representation. Jackowitz and Watson (1980) have suggested, following Fein (1975), that the child has to learn a separate set of inferences, or transformations, for pretending. However, this does not seem to be necessary if one assumes that pretend situation can make use of general knowledge. It is important, however, that inference rules preserve the semantics of the pretend representation. The principle I propose is as follows: If the input representation to the inference rule is primary, then the output representation will also be primary; if the input is decoupled, then so is the output. For example, one does not want the infant to infer from "the empty cup contains water" that the act of upturning the cup over a table will result in the table really becoming wet! Instead, it should result in a pretend attribution of "wetness" to the table (deviant truth). Thus inference rules apply within the decoupling marks and do not remove them. As long as this principle of preserving opacity is adhered to, there is no general need for a special set of inferences for pretense.

This example can be illustrated as follows:



I have abbreviated the sequence above, but the main lines are clear. UPTURN applies to containers and implies their contents spilling out. UPTURN can have the dual role of inference rule and command for the action scheme of upturning containers. So long as it is not itself decoupled, it can result in the action actually being performed. Further inferences can be drawn from the output of UPTURN (as a rule) resulting in a series of pretend representations.

Summary. In this section I have been concerned with the semantic properties of pretend representations as part of a competence theory. I propose to construe the semantics of pretense in terms of a three-term relation PRETEND (*a*, "*e_i*", *e_j*) between an agent *a*, a decoupled expression "*e_i*", and a primary expression *e_j*.

Another way to look at pretense is in terms of an information-processing system that embodies these semantic properties and

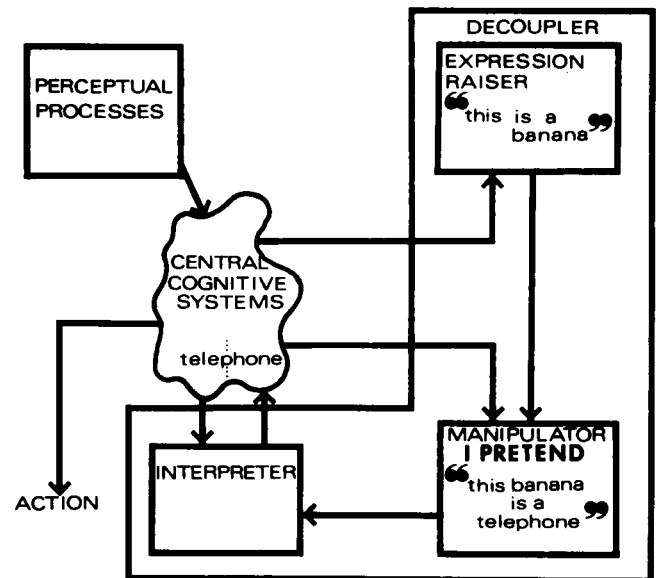


Figure 2. The decoupler model of pretense.

that can compute with such representations. Already I have come close to thinking about pretense as a processing activity that takes place in real time. For example, the series of inferences illustrated earlier can easily be thought of as successive cycles of processing. In the next sections, I begin to outline a general performance model along these lines. Having sketched the processing model, I shall then consider briefly how it relates to various phenomena in preschool development, both normal and abnormal.

Decoupling Model of Pretense

The decoupling model is illustrated in Figure 2. It has three main components. First, there are the *perceptual processes* whose job is to feed representations of the current situation to the central processes. Second, there is the set of processes labeled *central cognitive systems*. These include structures corresponding to perceived situation, memory systems (including, for example, general knowledge), systems for planning action, and so on. There is, of course, nothing novel in postulating these two components, but it helps in locating the architecture of the third major component, the decoupler. In modeling the decoupler, three major sets of processes have been posited. I call these the *expression raiser*, the *manipulator*, and the *interpreter* to suggest their functions.

The expression raiser's job is to copy primary representations from the central systems. It raises copies into the opaque context of the decoupling marks. The copy of the primary expression is thus removed from its normal input-output relations and from the normal computational consequences it would otherwise have. It will now form the nucleus of a metarepresentation: In short, it will exhibit the semantics of opacity.

The manipulator's job is to transform decoupled expressions by integrating (primary) information from memory within the decoupling marks or by applying inference rules from memory.

In doing this, the manipulator forms the pretend representation, supplying the context **Agent PRETEND** _____. In addition, the manipulator can also receive previously decoupled expressions from central systems that have been stored in central memory. The manipulator outputs pretend representations to the interpreter.

The interpreter can access primary representations in central systems. It performs anchoring functions and relates decoupled expressions to the current perceptual representation. It can access inference rules and other information for passing to the manipulator in a further cycle. It can pass metarepresentations to central cognitive systems for storage.

The model makes the following additional assumptions. There is a single representational code usable throughout the perceptual processes, central cognitive systems, and decoupler. The decoupling marks that fix the scope of decoupling are an item of (meta)representational code. Like the informational relations, they can be thought of as an extension to the primary code. The decoupler is the source of these extensions.

These, then, are the basic features of the performance model. I shall illustrate how it is supposed to work with various examples of the way it generates pretense. I will consider at the same time how the model attempts to account for the developmental changes observed in early pretend play.

Early Development of Pretend Play

The ability to pretend is defined as the power to compute the relation **PRETEND** (*a*, "*e_i*", *e_j*), where *a* ranges over agents and *e* over representational expressions: More specifically, "*e_i*" is a decoupled expression and *e_j* a primary representation of the current perceived situation. This is a fundamental ability in normal children, which, once having emerged, does not develop any further.

On the other hand, there are a large number of factors that will affect the possible content of pretense. Two such factors will be especially important, namely level of conceptual development and extent of encyclopedic knowledge. The problem of how genuinely new concepts are developed is, of course, a vexed question (see Fodor, 1981, chap. 10). All one can do here is assume that whatever mechanisms are responsible, new concepts will show up in primary representation. When this happens the concept will be available for incorporation or raising into pretense. Thus social roles will enter pretense either as object substitution or attribution of pretend properties depending on how the child represents such concepts—as *types* or *properties* of persons, respectively. Or to take another example: Whereas it is a mystery how a child develops the concept of becoming invisible (in the sense of "the invisible man" rather than mundane occlusion), once developed the concept will be available for pretense.

The decoupling model can be used to characterize how various kinds of pretense might arise. Perhaps the commonest way for early pretense to start is with the raising of a primary representation of the currently perceived situation. This can be called *immediate pretense* (see Figure 3). For instance, the current situation might contain a toy horse or an empty cup. Primary representations of these may be raised. This leaves the original representation still active: For example, memory systems are

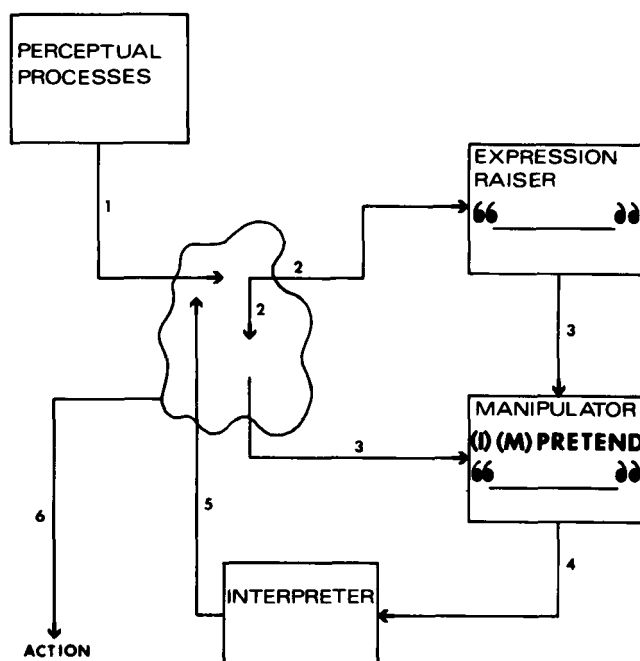


Figure 3. Immediate pretense ([1] Current perceptual representation is input to the central systems. [2] This primary representation drives further computations within the central systems—e.g., accessing information in memory—and is also copied by the expression raiser. [3] Information from memory is passed to the manipulator; decoupled expression from the expression raiser is passed to the manipulator, which integrates the two. [4] Pretend representation is passed to the interpreter. [5] which anchors the pretend representation to the current perceptual representation; [6] behavioral output.)

addressed, returning information on entities that are perceptually similar (e.g., on horses) or on the functional properties of the object (e.g., on containing). Such information may be passed to the manipulator and integrated into the pretend representation. This leads to pretense based on perceptual similarity (Elder & Pederson, 1978; Fein, 1975; Jackowitz & Watson, 1980) or on functional connection (Jackowitz & Watson, 1980; Pederson et al., 1981; Ungerer et al., 1981).

I suggest, tentatively, that anchoring proceeds by means of a best formal match between expressions in the pretend and current perceptual representations. If so, anchoring should be straightforward for immediate pretense; in the aforementioned example, **I PRETEND "this empty cup contains tea,"** the expression will have an exact equivalent in the original perception.

But where the expression raiser raises a representation from general knowledge (see Figure 4), there will be no guarantee of any close correspondence with the currently perceived situation, because the original representation is drawn from memory. For example, a representation **I had a birthday party** is raised from memory and integrated with further relevant information (e.g., about birthday cakes) to give **I PRETEND "there is a birthday cake at my party."** The interpreter may then have to "look around" for an adequately valued match to anchor "**birthday cake.**" This may result in the child's actually having to search the environment for a suitable prop. Such play would

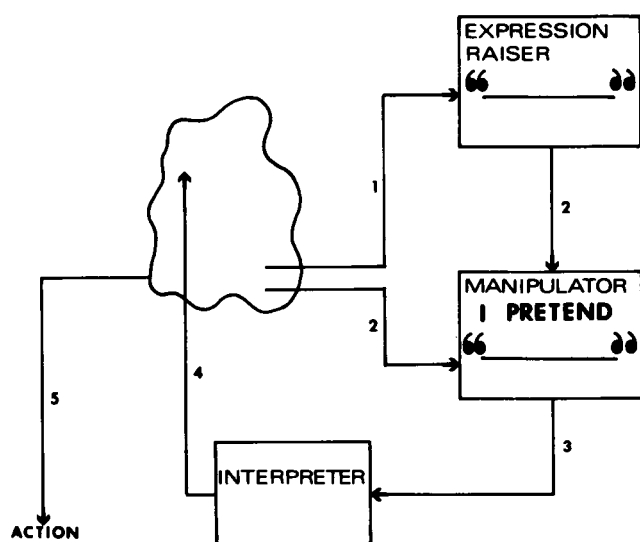


Figure 4. General knowledge pretense. ([1] Information retrieved from memory is copied by the expression raiser—e.g., *I had a birthday party*. [2] Further information is passed to the manipulator—e.g., *there are cakes at parties*. The decoupled expression is passed to the manipulator, which integrates the two: *I PRETEND "there is a birthday cake at my party."* [3] Pretend representation is passed to the interpreter, [4] which attempts to anchor the pretend representation to current perceptual representation; [5] behavioral output.)

be described as *planned pretend* (McCune-Nicolich, 1981; Nicolich, 1977).

A third type of pretense, *remembered pretense*, involves the retrieval of an already decoupled representation from memory (see Figure 5), for example, *"teddy is ill."* Such a representation can only be passed directly to the manipulator, which reinstates it as a pretend representation. Again, there may be no automatic candidate for anchoring (teddy may no longer be around). This type of pretense may be common in modeling experiments where the child retrieves and reenacts a past pretense of his or her own or one that was modeled for him or her. The problems of anchoring in such situations may be so great that offered props are rejected by the child (Bretherton et al., 1984; Golomb, 1977).

Finally, a fourth kind of pretense should be added: *understanding pretense in others*. This begins, like immediate pretense, with the raising of a perceptual representation—in this case, a representation of what someone is actually and literally doing. The child has to solve the problem of generating a pretend representation that will "explain" the other's behavior (which may be marked by exaggerated gestures and so on). For example, the child will just have to "hit" upon the similarity between bananas and telephones, aided perhaps by the functional clue of someone talking to a banana. If this can be accomplished, then information about telephones can be passed to the manipulator to arrive at the representation, *Mother PRETEND "that banana is a telephone."*

I have not considered here what broader motivations or purposes might lie behind pretending. The focus of this article is on the basic competence itself. This competence can be ap-

plied to wider and more complex topics as the development of the child's encyclopedic knowledge leads to changes in early pretend contents. But none of this need affect the underlying mechanisms. Empirical studies are needed to refine hypotheses concerning, for example, anchoring processes and repeated cycles of inference. Because the perceptions and stored knowledge of the child can be studied independently of pretense, it should be possible to examine pretense mechanisms quite carefully.

Pretense and "Theory of Mind"

The metarepresentational theory of pretense cannot be fully appreciated without considering its relation to the development of the child's commonsense or folk theory of mind. This term is borrowed from Premack and Woodruff (1978) who used it to denote the ability of a person to impute mental states to self and to others and to predict behavior on the basis of such states. To take a concrete example: John jumped into a shop doorway (behavior) because he *believed* it was raining and *wanted* to remain dry. In the most basic cases beliefs and desires are used together to explain or predict a piece of behavior. Notice that belief and desire imputations exhibit the logic of opacity. So, for instance, it does not matter if it was not really raining—John believed it was and that is why he jumped. To employ theory of mind requires that one can comprehend opaque states in oneself and in others.

Preschool Children's Theory of Mental States

There is growing evidence that the ability to use a commonsense theory of mind emerges in the preschool years in normal children (Baron-Cohen, Leslie, & Frith, 1985; Hogrefe, Wimmer, & Perner, 1986; Macnamara, Baker, & Olsen, 1976; Shantz, 1983; Shultz & Cloghesy, 1981; Wellman, 1985; Wimmer & Perner, 1983). It is manifested in a number of ways, in-

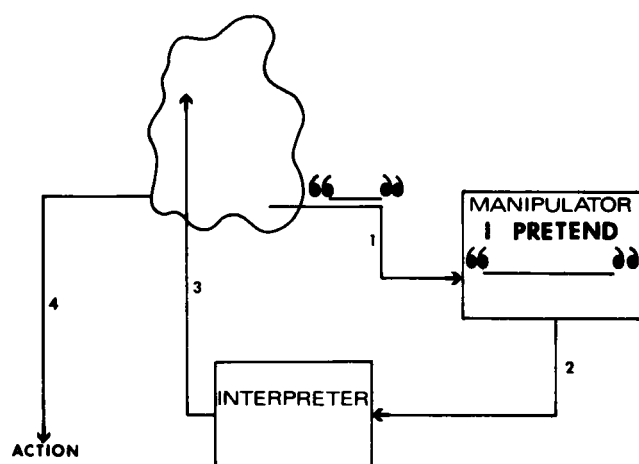


Figure 5. Remembered pretense. ([1] Decoupled expression is retrieved from memory, passed to the manipulator—e.g., *"teddy is ill"*—and reinstated as pretend representation, *I PRETEND "teddy is ill."* [2] Pretend representation is passed to the interpreter, [3] which attempts to anchor the pretend representation to current perceptual representation; [4] behavioral output.)

cluding normal moral development (Wimmer, Gruber, & Perner, 1984), understanding the consequences of ignorance (Hogrefe et al., 1986) and of false belief (Baron-Cohen et al., 1985; Wimmer & Perner, 1983), the appearance–reality distinction (Flavell, 1985; Flavell, Flavell, & Green, 1983; Harris, Donnelly, Guz, & Pitt-Watson, 1986), certain aspects of communication situations (Robinson & Whittaker, 1986), and in acquiring the language of mental state expressions (Bretherton & Beeghly, 1982; Shatz, Wellman, & Silber, 1983).

In the present view, what these developments have in common is that they require the deployment of metarepresentations. And indeed it is this that also links the ability to pretend and understand pretense in others to the employment of theory of mind. Pretend play is thus one of the earliest manifestations of the ability to characterize and manipulate one's own and others' cognitive relations to information. This ability, which is central to commonsense theory of mind, will eventually include characterizing relations such as believing, expecting, and hoping, and manipulating these relations in others, for example, getting someone to expect that something will happen by promising. In this section, I shall consider how the decoupling model might relate to some of the early developments in theory of mind.

Early communication. Around the beginning of the second year, intentional communication in gesture and vocalization emerges. These abilities may carry some implications for theory of mind. For example, in the philosophical literature, communication-intention theorists (Grice, 1957, 1968; Searle, 1969; Strawson, 1964) have argued that an essential feature of "intelligent" communication is that the speaker acts with a complex intention. The speaker's intention is said to be complex because part of its content represents the reflexive mental state of the audience. Thus, typically, the speaker intends that his audience recognize his or her intention to communicate (the message). These theories have attracted attention from those interested in the communicative abilities of infants (e.g., Bates, Benigni, Bretherton, Camaioni, & Volterra, 1979; Bruner, 1976, 1981).

Just as it is important to distinguish functional and pretend play, so in this context it is vital to ask if the child's (internally represented) goal is to influence someone's behavior or to influence someone's mental state. I assume that the goal of much infant communication is to achieve a concrete change in a situation or in behavior without reference to mental states. This is very different from having a goal to influence someone's mental state by sending a message (by means of a gesture or a sound). This latter goal would require an infant to represent his audience as having an informational relation to the message. Such a relation might simply be understanding the message (cf. Fodor, 1976, pp. 103–104). Bretherton, McNew, and Beeghly-Smith (1981) have considered observational evidence that suggests some such capacity may develop in the course of the second year. If so, it would require the deployment of metarepresentation. However, it is not yet clear, on this sort of evidence, how to distinguish complex communicative acts (generated under a metarepresentation) from communications generated by a primary mechanism that can take account of social/behavioral vicissitudes but that cannot explicitly represent mental states. The relation between metarepresentational capacity and com-

munication in infancy remains an interesting but still open question.

Talking about mental states. In the third year the child begins to acquire the mental state terms of his or her language (Bretherton & Beeghly, 1982). Shatz et al. (1983) showed that the earliest uses of mental state verbs are conversational (e.g., pause fillers) and do not refer to mental states. Reference to mental states begins in the second half of the third year, and by the third birthday, children are using about six mental state terms to refer to mental states. One child studied intensively by Shatz and her colleagues used seven terms between 2 years, 8 months and 3 years: *know*, *think*, *remember*, *pretend*, *dream*, *wonder*, and *believe*. The child appeared to use these terms appropriately with a complement clause expressing the content of the mental state.

It is hard to see how perceptual evidence could ever force an adult, let alone a young child, to invent the idea of unobservable mental states. Nor is it clear how language learning could lead to such a concept because the meaning of relevant linguistic expressions could not be grasped without first understanding the concept. But a learning mechanism drawing on the meta-representational powers emerging late in infancy could play an important role. For example, a distinction between primitive informational relations (e.g., *PRETEND*, *BELIEVE*) that take decoupling and those (e.g., *SEE*, *KNOW*) that do not could make a contribution to learning the semantics of the corresponding natural language terms. Verbs such as *see* take transparent complements, so if I saw the king of France yesterday, there must be a king of France; if there is no king of France, then it was not him that I saw. The child's task, then, would be to discover how a given linguistic expression translates into metarepresentational code. Although this sort of problem is far from trivial (see Landau & Gleitman, 1985), it is less monumental than having to invent the whole idea of mental states from scratch as well.

Pretense and false belief. By 4 years of age, children are capable of complex reasoning across metarepresentational structures. This is shown by the 4-year-old predicting the behavioral consequences of someone having a false belief (Baron-Cohen et al., 1985; Wimmer & Perner, 1983). To illustrate the use of this ability, consider the following scenario. Someone hides a piece of chocolate in a box and then goes away. Unknown to the hider, someone else transfers the chocolate to a basket and departs. The original hider returns. Where will the hider look for the chocolate? This simple test can be easily administered to young children and even to mentally retarded groups (Baron-Cohen et al., 1985). Appropriate control questions can check that the children remember the displacements of the object. The results will then show whether the child predicts merely from his own knowledge (hider looks where object really is) or whether he appreciates the hider's false belief.

Why does it take the 2-year-old pretender an additional 2 years to understand false belief? Wimmer and Perner (1983) argued that it was not until 4 years of age that the child could conceive simultaneously of two contradictory models of reality. But the early emergence of pretense shows that one must look elsewhere for an explanation.

Even a cursory comparison of pretense and false-belief understanding shows that they differ markedly in the complexity of the reasoning required. In pretense the metarepresentational

relations are essentially just stipulated. In false-belief understanding, the answer must be worked out. Even in understanding pretense in others, where the child has to infer from what the other is literally doing to what he or she is pretending, the answer is (deliberately) made obvious by the pretender. For example, mother performs a series of exaggeratedly clear telephoning actions with the banana and may even say "This is a telephone!" to emphasize the point. Indeed, if she does not go to such lengths, she runs the risk of not communicating the content of the pretend.

In contrast, understanding the previous false-belief situation requires identifying the specific events in the episode that are crucial for determining the relevant belief that someone will form, which in turn is crucial for predicting where that someone will look. The child must make inferences that go from a primary representation of the episode, through primary representations of what another person could and could not see of the episode, to a decoupled representation of what the other BELIEVES about the current situation. After that, an inference needs to be made from this metarepresentation back to a primary representation of what the other will do in that situation as a result of BELIEVING the decoupled expression. The inferential problem here clearly is more complex.

Perner, Leekam, and Wimmer (1987) showed 3-year-olds a confectionery packet well known to all British children and asked them what they thought it contained. Naturally enough, they answered, "Smarties." They were then shown that the packet actually contained a pencil. After this most of them were able to remember and report what their own false belief had been and that it was false. Despite this, when asked what their friend would think when he saw the box, nearly half were unable to predict the false belief they themselves had just suffered and said instead, "A pencil(!)."

The 3-year-old's problem therefore seems to be understanding how (false) beliefs arise, not representing and reporting beliefs per se. The 3-year-old has difficulty inferring from a person's contact with a situation to the belief the person will have as a result. They are thus happy to attribute miraculous knowledge. One should also expect to find difficulties in the other direction as well, that is, in inferring from a false belief to its consequences in behavior.

But does the greater complexity of false belief over pretense inference entirely account for the long 2-year lag? Or does the child's failure to draw the appropriate inferences in false belief reveal a deeper problem? Leslie (in press-a, in press-c) argued that the root of the child's difficulties may be in understanding the way in which mental states are part of the causal fabric of the world. Wellman (in press) has shown that the 3-year-old already has definite ideas about the way mental states exist. For example, they understand that bananas may be eaten, but that thoughts about bananas may not. Thus the 3-year-old already thinks of mental states as immaterial and abstract entities. I suggest that the next step—perhaps made more difficult for the child by his focus on nonconcreteness—is to think of mental states as abstract entities that nevertheless have concrete causes and concrete effects. Such insight would then underwrite a new interest in the predictive understanding of the relation between situations and the mental states of people exposed to them.

These ideas are discussed at greater length in Leslie (in press-b, in press-c).

I propose, then, that the basic representational structures for a theory of mind are put in place by the emergence of the decoupler mechanism. Upon this foundation, the development of specialized inferential knowledge builds a powerful causal theory.

One can also ask about abnormal development. In the following section, I will look at some recent work that suggests that the syndrome of childhood autism involves a pathology in the development of metarepresentational capacity.

Childhood Autism: Is There a Failure of Decoupling?

A characteristic feature of childhood autism is a severe impairment in pretend play (Baron-Cohen, 1987; Rutter, 1978; Sigman & Ungerer, 1981; Sigman, Ungerer, Mundy, & Sherman, 1987; Ungerer & Sigman, 1981; Wing, Gould, Yeates, & Brierley, 1977; Wulff, 1985). Wing et al. (1977), in their epidemiological study of 108 mentally retarded children, found that the absence of pretense was a consistent feature of children showing the "full syndrome" of autistic social impairment. Wulff (1985) concluded in her review that "the autistic child's play is striking in its lack of fantasy and all other aspects of symbolic play" (p. 146).

That this sort of impairment is not simply the inevitable result of general mental retardation is shown by two things. First, although primary representational abilities like object concept and causality appear to develop in line with mental age (MA) in the autistic child (Baron-Cohen, Leslie, & Frith, 1986; Curcio, 1978; Ungerer & Sigman, 1981), pretense is severely impaired relative to MA. Second, in other forms of mental retardation the ability to pretend is not impaired relative to these primary abilities or to MA—for example, in Down's syndrome children (Hill & McCune-Nicolich, 1981). For these reasons, it appears that the lack of pretend play in autism reflects a specific deficit and is not simply the result of whatever general mental retardation they may also suffer.

In terms of the present model, there is a ready explanation for the apparent dislocation in autism between primary and metarepresentational abilities. Although primary representational systems reflect only the general level of mental retardation, there is a specific metarepresentational deficit.

If this is so, autistic children should also show serious impairment in their later theory of mind. Such impairment should itself, moreover, reflect a specific deficit. Thus it should be found even in high-ability autistic children with borderline-to-average IQ. In addition, one should also expect that severely retarded Down's syndrome children who nevertheless pretend in line with MA would not show such a deficit in their theory of mind.

This was the reasoning behind two recent studies by Baron-Cohen et al. (1985, 1986). In the first of these, three groups of children were given the Wimmer and Perner test. One group consisted of clinically normal 4 1/2-year-olds, another of Down's syndrome children (mean IQ = 64, mean age = 11 years), and the third of children diagnosed as autistic (mean IQ = 82, mean age = 12 years). This high-ability autistic group was used to allow a conservative test of the hypothesis of specific impairment. The results showed consistent success by both the normal and the Down's groups with 85% and 86%, respectively, passing.

In contrast, 80% of the autistic children failed, consistently pointing to the location where the object actually was and not to where the hider should think it was. All the children in all three groups passed the control questions, showing that they understood and remembered the basic task.

In a follow-up study on the same children, Baron-Cohen et al., (1986) used a picture-sequencing task to assess the children's understanding of various kinds of events. The results confirmed that the autistic children had poor understanding of events involving mental states, performing significantly worse than the Down's and the clinically normal children. Autistic performance on picture sequences depicting mechanical events, on the other hand, was very good—significantly better than the Down's and the young normals. Verbal protocols taken from the children following sequencing confirmed this pattern and suggested that the autistics had a paucity of mental state language.

There is at least preliminary evidence, then, to suggest that a large proportion of autistic children have a specific deficit in theory of mind. This confirms the prediction of our theory of metarepresentational development based on prior findings that autistic children show a specific deficit in pretend play. The profound social impairment characteristic of childhood autism (American Psychiatric Association, 1980; Kanner, 1943; Rutter, 1978) may to some extent be explicable in terms of this metarepresentational deficit, leaving the autistic child unable to comprehend or predict a lot of the behavior of others.

Given that many autistic children appear to manifest a theory of mind neither in a primitive form (pretense) nor in a more advanced form (false-belief understanding), one can hypothesize that they are specifically impaired in their power of decoupling in this domain. This means that they should not show a specific deficit in understanding mental states that are not opaque, for example, *seeing*. Recent work by Hobson (1984) suggests that this is so. Hobson found that autistic children performed in line with MA and with MA-matched Down's children in line-of-sight and "three-mountains" type tasks. Thus, on perceptual perspective taking, autistic children do not appear to be specifically impaired.

This fairly complex pattern of deficits and abilities can be succinctly explained by the hypothesis that such children are decoupling impaired—for example, suffer a dysfunction in expression raising. The decoupler model may thus contribute to an understanding of questions in pathological as well as in normal development.

Final Remarks

The metarepresentational theory reveals pretend play in a new light as a primitive manifestation of the ability to conceptualize mental states. It allows new links to be made between normal and abnormal development. The emergence of metarepresentation through the growth of a decoupling mechanism implies a major developmental discontinuity. It seems likely that this constitutes a major part of the specific innate basis for our commonsense theory of mind.

The cognitivist framework focuses on the representational mechanisms underlying behavior. I have tried to apply it to some developmental problems of infancy and early childhood.

The result has been a more detailed modeling of the cognitive mechanisms underlying pretense than was available before. It should now be possible for empirical studies to exploit this analysis and to examine in some detail the child's thought processes during pretense. This should result in a greater understanding than we have at present of the child's central inferential processes and general knowledge.

The view of early representation that emerges lends support, I believe, to the language-of-thought hypothesis (Fodor, 1976, in press). The semantic properties of decoupling characterize the distinctiveness of pretend representations and explain why abuse does not occur. They also permit a single representational code to be used in both primary thought and pretense. This result is not only welcome on grounds of parsimony, but more important still, it captures significant generalizations that would otherwise be lost. Thus, for example, the **telephone** that features in pretense is the **telephone** that features in general knowledge, and the inference about (say) liquid containers made in pretense is the same inference made in problem-solving thought—or at least very nearly so. To the extent that such representations and such inferences do share common properties across contexts, there exist systematic generalizations that must be captured by a theory of human cognition. In short, the view advanced here offers for the first time a principled explanation for both the peculiarities of pretense and for the existence of these generalizations.

References

- Abravanel, E., & Gingold, H. (1985). Learning via observation during the second year of life. *Developmental Psychology*, 21, 614–623.
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month old infants. *Cognition*, 23, 21–41.
- Baron-Cohen, S. (1987). Autism and symbolic play. *British Journal of Developmental Psychology*, 5, 139–148.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, 21, 37–46.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioral and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology*, 4, 113–125.
- Barwise, J., & Perry, J. (1983). *Situations and attitudes*. Cambridge: MIT Press/Bradford.
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.
- Bateson, G. (1972). A theory of play and fantasy. In G. Bateson (Ed.), *Steps to an ecology of mind* (pp. 177–193). Aylesbury, England: Chandler.
- Belsky, J., & Most, R. K. (1981). From exploration to play: A cross-sectional study of infant free play behavior. *Developmental Psychology*, 17, 630–639.
- Bower, T. G. R. (1974). *Development in infancy*. San Francisco: Freeman.
- Bower, T. G. R. (1978). Concepts of development. In *Proceedings of the 21st International Congress of Psychology* (pp. 79–97). Paris: Presses Universitaires de France.
- Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology*, 18, 906–921.

- Bretherton, I., McNew, S., & Beehly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a "theory of mind"? In M. E. Lamb & L. R. Sherod (Eds.), *Infant social cognition* (pp. 333-373). Hillsdale, NJ: Erlbaum.
- Bretherton, I., O'Connell, B., Shore, C., & Bates, E. (1984). The effect of contextual variation on symbolic play: Development from 20 to 28 months. In I. Bretherton (Ed.), *Symbolic play and the development of social understanding* (pp. 271-298). New York: Academic Press.
- Bruner, J. S. (1976). From communication to language—A psychological perspective. *Cognition*, 3, 255-287.
- Bruner, J. S. (1981). Intention in the structure of action and interaction. In L. P. Lipsitt & C. K. Rovee-Collier (Eds.), *Advances in infancy research* (Vol. 1, pp. 41-56). Norwood, NJ: Ablex.
- Carnap, R. (1947). *Meaning and necessity*. Chicago: University of Chicago Press.
- Chomsky, N. A. (1980). *Rules and representations*. Oxford, England: Blackwell.
- Church, A. (1950). On Carnap's analysis of statements of assertion and belief. *Analysis*, 10, 97-99.
- Cole, D., & LaVoie, J. C. (1985). Fantasy play and related cognitive development in 2 to 6 year-olds. *Developmental Psychology*, 21, 233-240.
- Corrigan, R. (1982). The control of animate and inanimate components in pretend play and language. *Child Development*, 53, 1343-1353.
- Curcio, F. (1978). Sensorimotor functioning and communication in mute autistic children. *Journal of Autism and Childhood Schizophrenia*, 8, 281-292.
- Dale, N. (1983, September). *Early pretending in parent-child and sibling-child play*. Paper presented at the meeting of the British Psychological Society, Oxford, England.
- Dennett, D. (1983). Styles of mental representation. In *Proceedings of the Aristotelian Society* (pp. 213-226). London: Aristotelian Society.
- Dunn, J., & Dale, N. (1984). I a Daddy: Two year old's collaboration in joint pretend with sibling and with mother. In I. Bretherton (Ed.), *Symbolic play and the development of social understanding* (pp. 131-158). New York: Academic Press.
- Elder, J. L., & Pederson, D. R. (1978). Preschool children's use of objects in symbolic play. *Child Development*, 49, 500-504.
- Fein, G. G. (1975). A transformational analysis of pretending. *Developmental Psychology*, 11, 291-296.
- Fein, G. G. (1981). Pretend play in childhood: An integrative review. *Child Development*, 52, 1095-1118.
- Fein, G. G., & Apfel, N. (1979). Some preliminary observations on knowing and pretending. In M. Smith & M. B. Franklin (Eds.), *Symbolic functioning in childhood*. Hillsdale, NJ: Erlbaum.
- Fenson, L., & Ramsay, D. S. (1980). Decentration and integration of the child's play in the second year. *Child Development*, 51, 171-178.
- Fenson, L., & Ramsay, D. S. (1981). Effects of modeling action sequences on the play of twelve-, fifteen-, and nineteen-month old children. *Child Development*, 52, 1028-1036.
- Field, T., De Stefano, L., & Koewler, J. H. (1982). Fantasy play of toddlers and preschoolers. *Developmental Psychology*, 18, 503-508.
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477-531.
- Fischer, K. W., & Pipp, S. L. (1984). Processes of cognitive development: Optimal level and skill acquisition. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 45-80). New York: Freeman.
- Flavell, J. H. (1985, August). *The development of children's knowledge about the appearance-reality distinction*. Paper presented at the annual convention of the American Psychological Association, Los Angeles.
- Flavell, J. H., Flavell, E. R., & Green, F. L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology*, 15, 95-120.
- Fodor, J. A. (1976). *The language of thought*. Hassocks, Sussex England: Harvester Press.
- Fodor, J. A. (1981). *Representations: Philosophical essays on the foundation of cognitive science*. Brighton, England: Harvester Press.
- Fodor, J. A. (in press). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge: MIT Press.
- Golomb, C. (1977). Symbolic play: The role of substitutions in pretence and puzzle games. *British Journal of Education Psychology*, 47, 175-186.
- Grice, H. P. (1957). Meaning. *Philosophical Review*, 64, 377-388.
- Grice, H. P. (1968). Utterer's meaning, sentence meaning, and word meaning. *Foundations of Language*, 4, 1-18.
- Harris, P. L., Donnelly, K., Guz, G. R., & Pitt-Watson, R. (1986). Children's understanding of the distinction between real and apparent emotion. *Child Development*, 57, 895-909.
- Haugeland, J. (1978). The nature and plausibility of cognitivism. *The Behavioral & Brain Sciences*, 2, 215-260.
- Hill, P. M., & McCune-Nicolich, L. (1981). Pretend play and patterns of cognition in Down's syndrome children. *Child Development*, 52, 217-250.
- Hobson, R. P. (1984). Early childhood autism and the question of egocentrism. *Journal of Autism and Developmental Disorders*, 14, 85-104.
- Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, 57, 567-582.
- Huttenlocher, J., & Higgins, E. T. (1978). Issues in the study of symbolic development. In W. Collins (Ed.), *Minnesota Symposia on Child Psychology* (Vol. 11, pp. 98-140). Hillsdale, NJ: Erlbaum.
- Jackendoff, R. (1983). *Semantics and cognition*. Cambridge: MIT Press.
- Jackowitz, E. R., & Watson, M. W. (1980). The development of object transformations in early pretend play. *Developmental Psychology*, 16, 543-549.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217-250.
- Keil, F. (1979). The development of the young child's ability to anticipate the outcomes of simple causal events. *Child Development*, 50, 455-462.
- Keil, F. (1984). Mechanisms in cognitive development and the structure of knowledge. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 81-100). New York: Freeman.
- Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge: Harvard University Press.
- Leslie, A. M. (1982). Discursive representation in infancy. In B. de Gelder (Ed.), *Knowledge and representation* (pp. 80-93). London: Routledge & Kegan Paul.
- Leslie, A. M. (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception*, 13, 287-305.
- Leslie, A. M. (1986). Getting development off the ground: Modularity and the infant's perception of causality. In P. van Geert (Ed.), *Theory building in developmental psychology* (pp. 405-437). Amsterdam: Elsevier North-Holland.
- Leslie, A. M. (in press-a). The child's understanding of the mental world. In R. L. Gregory (Ed.), *The Oxford companion to the mind*. Oxford, England: Oxford University Press.
- Leslie, A. M. (in press-b). *Engines of development: Mechanisms of mental representation*. Oxford, England: Blackwell.
- Leslie, A. M. (in press-c). Some implications of pretense for mechanisms underlying the child's theory of mind. In J. Astington, D. Olson, & P. Harris (Eds.), *Developing theories of mind*. Cambridge, England: Cambridge University Press.
- Leslie, A. M., & Keeble, S. (1987). Do six-month-olds perceive causality? *Cognition*, 25, 265-288.

- Lowe, M. (1975). Trends in the development of representational play in infants from one to three years—An observational study. *Journal of Child Psychology and Psychiatry*, 16, 33–47.
- Macnamara, J., Baker, E., & Olsen, C. (1976). Four year old's understanding of pretend, forget, and know: Evidence for propositional operations. *Child Development*, 47, 62–70.
- Mandler, J. (1983). Representation. In J. H. Flavell & E. M. Markman (Eds.), *Manual of child psychology: Cognitive development* (pp. 420–494). New York: Wiley.
- Mandler, J. (1984). Representation and recall in infancy. In M. Moscovitch (Ed.), *Infant memory* (pp. 75–101). New York: Plenum Press.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- McCune-Nicolich, L. (1981). Toward symbolic functioning: Structure of early use of early pretend games and potential parallels with language. *Child Development*, 52, 785–797.
- McCune-Nicolich, L., & Fenson, L. (1984). Methodological issues in studying early pretend play. In T. D. Yawley & A. D. Pellegrini (Eds.), *Child's play: Developmental and applied* (pp. 81–104). Hillsdale, NJ: Erlbaum.
- Meltzoff, A. N. (1981). Imitation, intermodal coordination, and representation in early infancy. In G. Butterworth (Ed.), *Infancy and epistemology* (pp. 85–114). London: Harvester Press.
- Nicolich, L. M. (1977). Beyond sensorimotor intelligence: Assessment of symbolic-maturity through analysis of pretend play. *Merrill-Palmer Quarterly*, 23, 89–99.
- Pederson, D. R., Rook-Green, A., & Elder, J. L. (1981). The role of action in the development of pretend play in young children. *Developmental Psychology*, 17, 756–759.
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125–137.
- Piaget, J. (1962). *Play, dreams and imitation in childhood*. London: Routledge & Kegan Paul.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Sciences*, 4, 515–526.
- Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? *The Behavioral and Brain Sciences*, 1, 592–593.
- Quine, W. V. (1961). *From a logical point of view*. Cambridge: Harvard University Press.
- Robinson, E. J., & Whittaker, S. J. (1986). Children's conceptions of meaning–message relationships. *Cognition*, 22, 41–60.
- Rock, I. (1983). *The logic of perception*. Cambridge: MIT Press.
- Rosenblatt, D. (1977). Developmental trends in infant play. In B. Tizard & D. Harvey (Eds.), *Biology of play* (pp. 33–44). London: Heinemann Medical Books.
- Rutter, M. (1978). Language disorder and infantile autism. In M. Rutter & E. Schopler (Eds.), *Autism: A reappraisal of concepts and treatment* (pp. 85–104). New York: Plenum Press.
- Searle, J. N. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, England: Cambridge University Press.
- Shantz, C. V. (1983). Social cognition. In P. H. Mussen (Ed.), *Handbook of child psychology* (Vol. 3, pp. 495–555). New York: Wiley.
- Shatz, M., Wellman, H., & Silber, S. (1983). The acquisition of mental verbs: A systematic investigation of the first reference to mental states. *Cognition*, 14, 301–321.
- Shultz, T. R., & Cloghsey, K. (1981). Development of recursive awareness of intention. *Developmental Psychology*, 17, 465–471.
- Sigman, M., & Ungerer, J. (1981). Sensorimotor skill and language comprehension in autistic children. *Journal of Abnormal Child Psychology*, 9, 149–165.
- Sigman, M., Ungerer, J., Mundy, P., & Sherman, T. (1987). Cognition in autistic children. In D. Cohen, A. Donnellan, & R. Paul (Eds.), *Handbook of autism and pervasive developmental disorders* (pp. 103–120). New York: Wiley.
- Spelke, E. S. (1982). Perceptual knowledge of objects in infancy. In J. Mehler, E. Walker, & M. Garrett (Eds.), *Perspectives on mental representation* (pp. 409–430). Hillsdale, NJ: Erlbaum.
- Strawson, P. F. (1964). Intention and convention in speech acts. *Philosophical Review*, 73, 439–460.
- Ullman, S. (1980). Against direct perception. *The Behavioral & Brain Sciences*, 3, 373–415.
- Ungerer, J. A., & Sigman, M. (1981). Symbolic play and language comprehension in autistic children. *Journal of the American Academy of Child Psychiatry*, 20, 318–337.
- Ungerer, J. A., Zelazo, P. R., Kearsley, R. B., & O'Leary, K. (1981). Developmental changes in the representation of objects in symbolic play from 18 to 31 months of age. *Child Development*, 52, 186–195.
- Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Soviet Psychology*, 5, 6–18.
- Watson, M. W., & Fischer, K. W. (1977). A developmental sequence of agent use in late infancy. *Child Development*, 48, 828–836.
- Watson, M. W., & Fischer, K. W. (1980). Development of social roles in elicited and spontaneous behavior during the preschool years. *Developmental Psychology*, 16, 483–494.
- Wellman, H. M. (1985). The child's theory of mind: The development of conceptions of cognition. In S. R. Yussen (Ed.), *The growth of reflection* (pp. 169–206). San Diego, CA: Academic Press.
- Wellman, H. M. (in press). First steps in the child's theorizing about the mind. In J. Astington, D. Olson, & P. Harris (Eds.), *Developing theories of mind*. Cambridge, England: Cambridge University Press.
- Werner, H., & Kaplan, B. (1967). *Symbol formation*. New York: Wiley.
- Wimmer, H. M., Gruber, S., & Perner, J. (1984). Young children's conception of lying: Lexical realism–moral subjectivism. *Journal of Experimental Child Psychology*, 37, 1–30.
- 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.
- Wing, L., Gould, J., Yeates, S. R., & Brierley, L. M. (1977). Symbolic play in severely mentally retarded and in autistic children. *Journal of Child Psychology and Psychiatry*, 18, 167–178.
- Wulff, S. B. (1985). The symbolic and object play of children with autism: A review. *Journal of Autism and Developmental Disorders*, 15, 139–148.

Received June 9, 1986

Revision received February 18, 1987

Accepted March 2, 1987 ■