Thought Without Language

Edited by
L. Weiskrantz

A Fyssen Foundation Symposium

CLARENDON PRESS • OXFORD
1988

8

The necessity of illusion: perception and thought in infancy

ALAN M. LESLIE

Introduction

I am going to discuss three examples of the way in which events are understood early in development. In the first example, infants perceive a specifically causal property of a simple event. In the second, infants show a thoughtful reaction to a more complex causal situation. In the final case, full-blown counterfactual causal reasoning is involved at the start of childhood.

These examples of ‘surprising’ early abilities are interesting in their own right. We can add them to the growing catalogue of such things. But my motive here goes beyond this. I think these cases can give us insight into how the infant mind is organized. Recent advances in experimentation have led to the collapse of the sensori-motor theory of infancy, but they have not automatically produced a framework to replace it. A new theoretical understanding of the mental architecture of infancy, however, is very much on the agenda (Leslie 1986, 1987*a*; Mandler, in press; Spelke 1987, 1988).

One view of the infant mind is that it is essentially homogeneous, without differentiated powers, and without symbolic processes—a single network that acquires structure gradually through associative learning or through some other principle of equal generality.

I want to discuss a quite different framework for infant cognition. This framework argues for an infant with a wide variety of mental structures and powers (Leslie 1986, 1987*a*; Leslie and Keeble 1987; Spelke 1987, 1988). It is this variety of specific mechanisms and the overall design into which they fit that holds the key to understanding the competence of the infant and his powers to develop.

Each of the three examples of causal understanding I shall deal with illustrates a different level of mental organization. Each level has its own distinct tasks and mechanisms suited to their execution. Carrying out these tasks requires symbolic representation and creates systems of knowledge with logical and conceptual structure. In building this knowledge, the role of perception is to provide thought with a conceptual identification of current input from the environment (Fodor 1983; Sperber and Wilson 1986). I shall argue that recent results from the study of infancy reveal that this mental architecture is the basis for development, and not its outcome.

The significance of illusions

Part of my method in addressing infant cognitive organization will be to consider the nature and existence of illusions. The essence of a perceptual illusion is that a bit of the world appears to us in a way we know is not or cannot be the case but which, despite such knowledge, appears this way repeatedly and incorrigibly. Illusions are important because they reflect inherent limitations either in the models of the world that brain mechanisms build, or in the way the mechanisms build them, or in the way these mechanisms interact (Coren and Ward 1979; Gregory 1974, ch. 30; Robinson 1972).

A strong case can be made that perceptual mechanisms are organized on a modular basis (Fodor 1983; Marr 1982; Ullman 1984). The computational task of maintaining a detailed description of current input to the organism appears to be broken down into a number of independent subtasks. These are then carried out by devices dedicated to these subtasks, operating automatically, independently of other devices, and without access to knowledge or information represented centrally.

The modularity of perception provides an explanation both for the existence of illusions and for their incorrigibility in the face of what we know about the distal stimulus. Illusions are an inevitable consequence of automatic computation of limited solutions to limited problems with limited information access (Leslie 1986). But the incorrigibility of illusion implies something more than simply the impenetrability of input-processing. Illusions often create incongruities in a perceived situation. We lack the ability to modify the percept, but we do have the ability to detect the incongruities. The mechanisms of thought that detect such incongruities must have a different character from those that created them.

My aim is to exploit such phenomena to prise apart the hidden seams of perception and thought in infancy and to understand their relationship to one another in development.

A causal illusion

To suggest that there is such a thing as a *perceptual illusion* of causality is to imply that there is a rather humble perceptual mechanism operating automatically and incorrigibly upon the spatio-temporal properties of events[1] yet producing abstract descriptions of their causal structure. It also implies that the idea of cause and effect does not originate in prolonged learning. It was Michotte (1963) who discovered that adults are, under certain circumstances, subject to just such an illusion. I have been trying to determine whether or not young infants are subject to a similar illusion (Leslie 1982, 1984). I have recently obtained evidence which indicates that they are (Leslie 1986; Leslie and Keeble 1987).

Experimental studies: a first question

My investigation of a causal illusion in infants has gone in a number of steps, each using the habituation-dishabituation of looking technique. The infant watches a film of a red object colliding in a variety of ways with a green object. The film is presented repeatedly until the infant begins to lose interest as measured by the length of succeeding unbroken looks. After this, a variety of slightly changed events can be presented and any recovery of interest, measured in the same way, can be compared with a base-line established with an unchanged event. The pattern of recovery across a number of event comparisons can then provide a basis for inferring how the events are being perceived.

The basic event of these studies I call *direct launching*. This corresponds to a billiard ball collision type event where one object launches another by colliding with it (see Fig. 8.1(a)). The first question was: Can infants distinguish the *submovements* involved in direct launching or is it simply perceived as a single unanalysable 'whoosh' going from one side of the screen to the other?

I argued (Leslie 1984) that if direct launching is seen as an event with a particular internal structure (i.e. composed of submovements), then reversing the event, by playing the film backwards, should rearrange that structure. If, however, an event has *no* submovements, then reversing it would affect only properties such as spatial direction which do not depend upon structured subcomponents.

The idea then was to use reversal to probe for the infant's perception of structure in direct launching. I compared the effect of reversing direct launching with the effect of reversing a single movement made by a single object (see Fig. 8.1 (a) and (b)). Since a single movement has no subcomponents, reversal will change only its spatial direction. Using the looking technique, one can predict the following from the subcomponent hypothesis: those infants habituated to direct launching and tested on its reversal will

[1] Such a device could also consider properties of the objects involved, if it operated sufficiently late in the input systems (i.e. after object recognition). Unfortunately, the evidence bearing on whether it does or not is scanty (for discussion, see Leslie 1986).

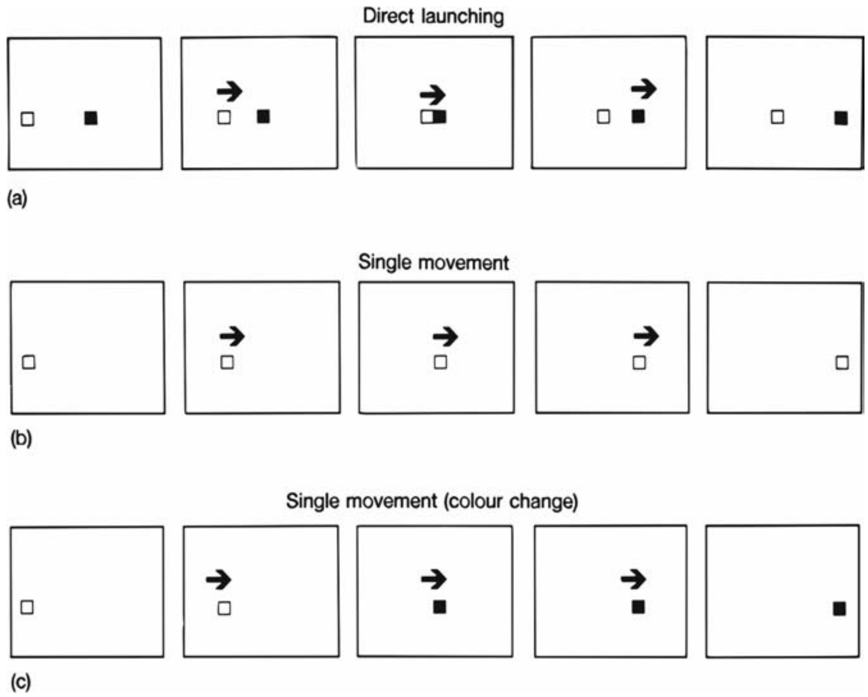

Fig. 8.1 Illustration of films used by Leslie (1984) to test for the perception by infants of internal structure in direct launching (from Leslie and Keeble 1987).

recover their looking *more* than those habituated to a single movement and tested on its reversal. The results of this experiment (Leslie 1984, experiment 1A) showed, as predicted, little recovery in the single movement group and significantly higher recovery in the direct launching group.

Despite this finding, the possibility remained that direct launching was perceived as a single movement but with differently coloured halves—as a single moving entity that changes colour from red to green half-way across. I made a film in which exactly this happened (see Fig. 8.1(c)). If infants do see direct launching this way, they should not readily discriminate these two sequences. In a new experiment, however, infants easily made this discrimination (Leslie 1984, experiment IB). Taken together, these two studies showed that six-month-olds did detect internal structure, and thus parsed the submovements, in direct launching.

A question about connections

I now asked what kind of internal structure, beyond submovements, infants could perceive in direct launching. Two further experiments (Leslie 1984, experiments 2 and 3) tried to find out how they perceive the relationship between the submovements. Do they perceive causal relationships or simply spatio-temporal properties?

I want to skip over many of the details here so that I can get to broader issues. Suffice it to say that a set of films were prepared which varied the spatio-temporal relations between the submovements. One version had a short delay between the impact of one object and the reaction of the other, another had a small gap between the objects so that they did not actually make contact, while yet another had both the delay and the gap combined. These sequences are illustrated in Fig. 8.2. Only the first sequence, direct launching, appears directly causal to adult observers.

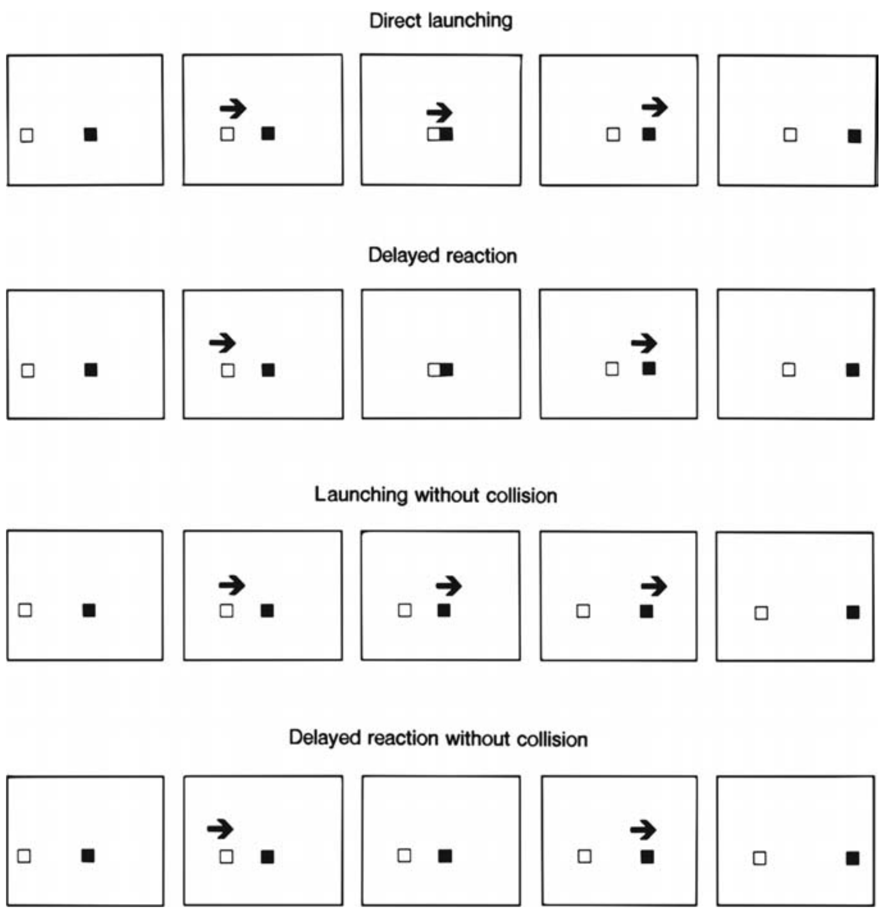

Fig. 8.2 Illustration of films used by Leslie (1984) to test infant perception of the relation between submovements in collision events (from Leslie and Keeble 1987).

All the possible comparisons between pairs of sequences were presented to the infants. Would a given contrast in spatio-temporal parameters be more effective in producing recovery of interest when it involved a causal contrast than when it did not? For example, in direct launching vs. delayed reaction without collision (see Fig. 8.2) a delay and a gap are introduced in going from a 'causal' to a 'non-causal' sequence. A delay plus a gap is also the difference between delayed reaction and launching without collision, but here both sequences are 'non-causal'. Would the infants perceive a greater difference between the first pair than between the second? It seems they did, suggesting a causal property had been perceived.

The other comparisons, however, did not support a causal conclusion. In fact, the overall results seemed simpler to account for in terms of a fairly abstract, but not causal, property which I called *spatio-temporal continuity.* The infants seemed to be encoding the sequences in terms of the degree of continuity between the submovements, but without regard for whether discontinuity came from a spatial gap or a temporal delay (Leslie 1984).

Reversing causation

It seemed to me that there was a good possibility that the previous experiments were just missing the infant's sensitivity to causality. I tried to think of a way both to minimize and to control for the spatio-temporal differences between the sequences presented so as to 'isolate' the causal structure. I returned to the technique of reversing the event.

The idea behind this new experiment was as follows. In some causal events, reversal of spatio-temporal direction entails reversal of causal direction as well. Launching is such an event. For example, billiard ball A directly launches billiard ball B by colliding with it in a rightward direction—A causes B to move. In the reverse of this event, billiard ball B comes back and directly launches ball A in a leftward direction—B causes A to move. Thus, causal direction, *as well as* spatio-temporal direction, reverses.

But in 'delayed reaction', causal direction is, by hypothesis, absent. That is, if delayed reaction is not perceived as causal, then reversal will affect *only* its spatio-temporal direction (left/right orientation and order of movement). At the causal level, however, it will lack internal structure.

Figure 8.3 illustrates the sequences. If infants perceive causal direction only in direct launching and not in delayed reaction, they will be differentially sensitive to their reversal. They ought to respond to causal *and* spatio-temporal reversal in the case of direct launching, but only to spatio-temporal reversal in the case of delayed reaction. Reversal of direct launching should therefore produce greater recovery of interest.

This is exactly what we found (Leslie and Keeble 1987). Infants around 27 weeks recovered more to reversal of an apparently causal event than to the

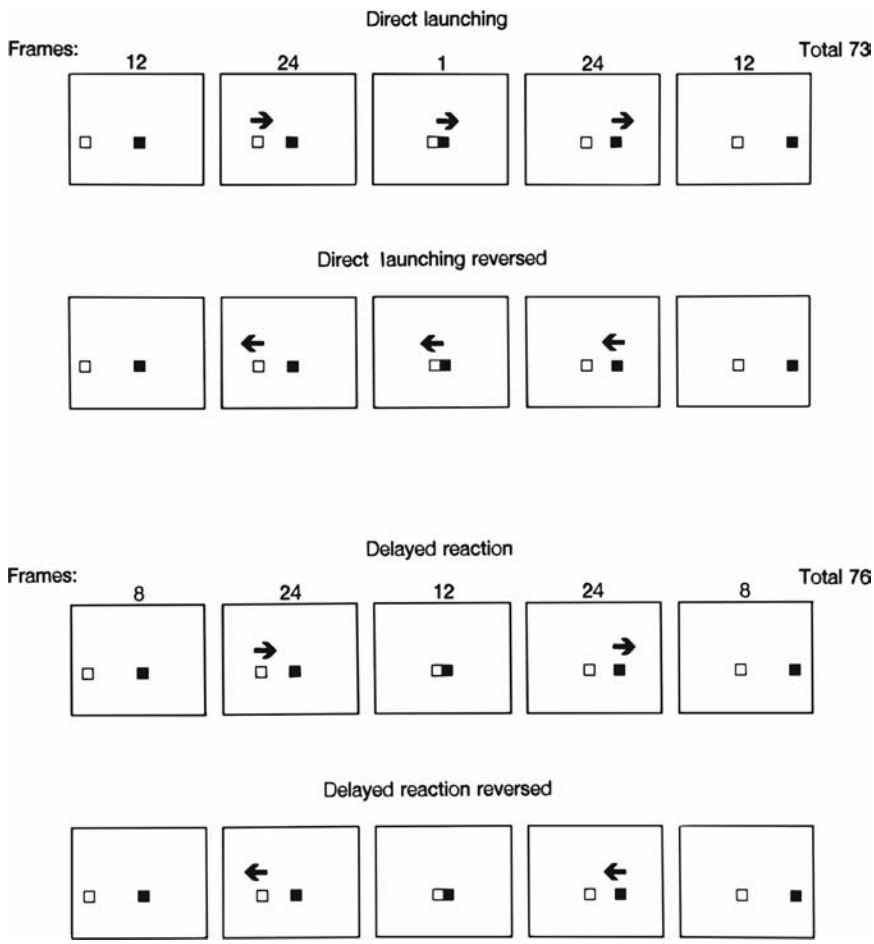

Fig. 8.3 Illustration of films used by Leslie and Keeble (1987) to test for infant perception of causal direction (from Leslie and Keeble 1987).

reversal of an apparently non-causal event. Figure 8.4(a) shows the mean looking times obtained on first look to stimulus, last look following habituation and first look following reversal of the stimulus. Both groups were similar on last look to their respective films. But when these films were reversed, the direct launching group increased its looking significantly more.

The results of a replication were even clearer (Fig. 8.4(b)). We included a control group with no reversal to check that its looking would stay at the same level on the test trial and it did. As predicted, reversal of direct launching produced the most recovery.

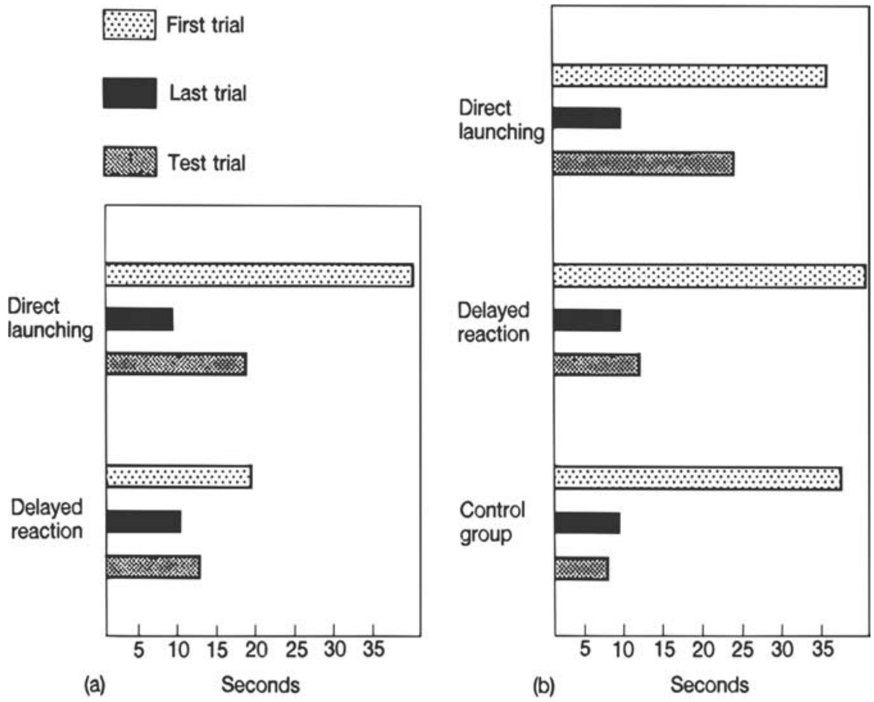

Fig. 8.4 Results showing looking times on first and last trials of habituation to films illustrated in Fig. 8.3 and looking times on test trial to their reversal, (a) shows first experiment and (b) replication (from Leslie and Keeble 1987).

Causal perception: a hypothesis

The reversal experiments suggest that young infants can perceive a specifically causal relation. Because spatio-temporal changes were controlled, and because infants recover less both to a reversed single movement (Leslie 1984, experiment 1A) and to a reversed delayed reaction, we require a *structural* explanation. I have proposed (Leslie 1986; Leslie and Keeble 1987) that at 27 weeks there is a visual mechanism already operating which is responsible for organizing a causal percept. Taking input from lower level motion-processing, this device will parse submovements, produce higher level descriptions of the spatio-temporal properties of the event, and produce a description of its causal structure.

A working hypothesis about the output of this mechanism is illustrated in Fig. 8.5. Multiple representations are computed for the same event. Succeeding representations become more abstract and a higher level description is computed from a lower one. At the first level, the spatial and temporal relations between the submovements are computed and represented orthogonally. This allows a redescription of launching and its variants in terms of continuity at the second level, produced by summing the values of the

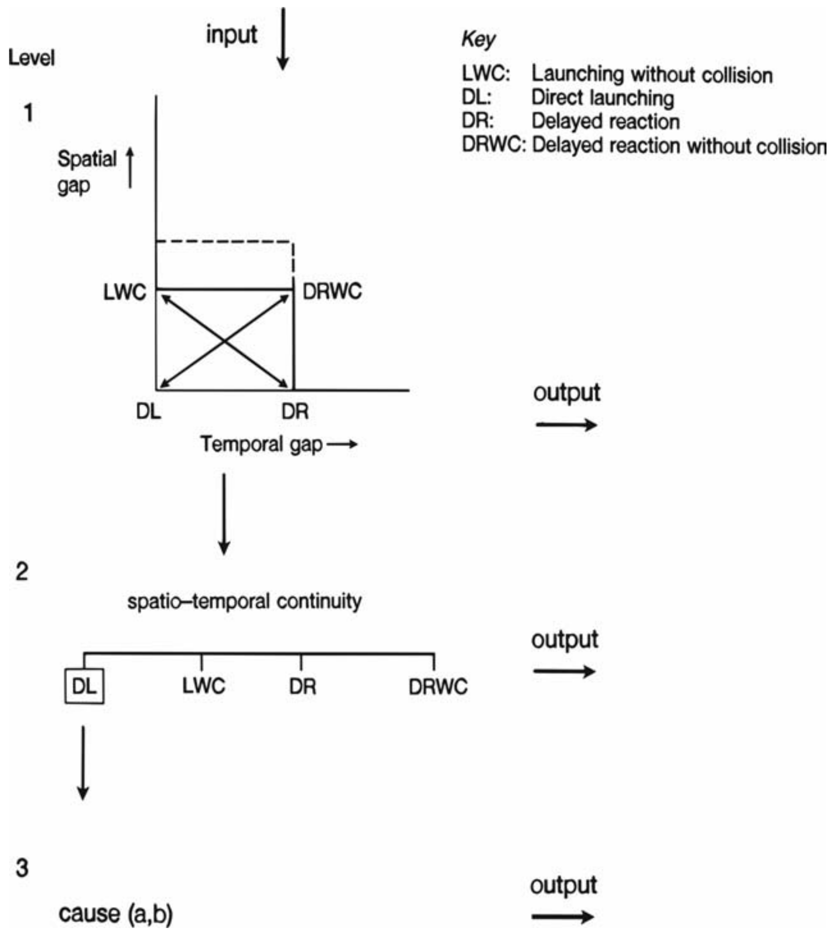

Fig. 8.5 A working hypothesis concerning the output from a module for analysing launching events (from Leslie and Keeble 1987).

parameters at level one. The second-level description then allows the selection of highly continuous events for redescription at the last level. Causal roles may be described at this third level. Further investigation of this level is at present under way.

Modular perception and development

Why should there be such a visual mechanism and why should it be operational at 27 weeks of age? The answer could be that this mechanism forms part of a major learning system. The module for perceiving launching automatically provides a conceptual identification of its input for central thought. There are a number of specific contributions such a device might make (Leslie and Keeble 1987). For example, it could help analyse visible mechanisms, distinguishing causally connected events from those which merely co-vary or are coincidental. Its descriptions could suggest plausible hypotheses for central thought to follow up. In this way it could promote a rapid build-up of mechanical understanding and thus help explain pre-school children's competence (Bullock 1985; Bullock *et al.* 1982; Kun 1978; Schultz 1982).

A further implication is that this same mechanism operates in adults and gives rise to the causal illusion discovered by Michotte (1963). The existence of this illusion will be a side-effect of the modularity of the underlying mechanism: it will operate automatically and incorrigibly given the right input. Infants too will be subject to the same illusion and for the same reasons adults are. This suggests an important connection between adult illusions and infantile perceptual competence: namely, modular perceptual systems of adults are ideal for fostering early knowledge acquisition.

Recall that an essential property of a modular input process is that it is impervious to general knowledge and reasoning. It can, and does, operate without the benefit of either of these. Such a mechanism is ideal for operating early in infancy when there is little or no encyclopaedic knowledge and only limited reasoning ability. It can provide an automatic starting engine for encyclopaedic knowledge. Because it operates independently of such knowledge and reasoning, it can function at a time when these are just beginning to develop, and it can do so without suffering any disadvantage whatsoever. It can provide a conceptual identification of input from the environment, in terms of cause and effect, in exactly the right format for inferential processes, and do this even in the absence of past experience. This is perfect for a mechanism whose job is to help produce development. But do infant input systems actually feed into central inferential mechanisms, or must they await the development of thought processes which can exploit perceptual descriptions?

A causal principle

Baillargeon has made an important discovery about the young infant's understanding of mechanics (Baillargeon 1986, 1987*a*, *b;* Baillargeon *et al.* 1985). In the basic experiment, five-month-old infants watch a screen which starts flat on a table and rotates backwards in a drawbridge type movement until it is flat on the table again (Baillargeon *et al.* 1985). This is repeated until the infants habituate. With the screen back in its starting position, they are shown a box being placed behind the screen. The infants then watch the same movement of the screen as before. After the screen reaches 30 to the upright, the box is occluded from the infant's view for the remainder of its rotation.

The results showed that those infants who were tested on the impossible event in which the screen made the same movement but appeared to rotate *through* the hidden box recovered interest and appeared to be surprised.

Meanwhile, the infants who saw the new but possible event in which the screen stopped when it reached the hidden object showed less recovery of interest.

It is hard to fit this result within the standard framework of habituation-dishabituation theory. In particular, it is hard to see how the infant's dishabituation could have been the result of an automatic process of perceptual discrimination and local stimulus recognition (see Mandler, in press). It seems rather to reflect the central evaluation of the significance of a change in the real world. For it is only possible to understand why the infants dishabituate by considering the stimulus as an *event* in a sequence of events in a world where things have to make sense in certain ways.

These results have since been extended to cover a third object moving behind the screen. In this case, the hidden box is either blocking the moving object's trajectory or merely alongside it. Infants are surprised only when the moving object appears to have passed through the blocking box's position (Baillargeon 1986). In another variation on the original set-up, a compressible object is hidden behind the screen. This time the infants are not surprised when the screen rotates all the way back. Furthermore, infants' surprise is also contingent upon the orientation of the hidden object: it must be oriented such that it will be in the right place to block the screen's backward rotation (Baillargeon 1987a).

The infants in these studies understand where the hidden object is, what its orientation is, whether it is compressible or not, and retain fairly accurate information about its spatial extent. The infants use this rich representation to make judgements about the likely outcomes for mechanical interactions, even though some of these are also hidden from view.

I said these results imply an important evaluative act of understanding from infant central thought. Since the very existence of thought in young infants has traditionally been doubted, this hunch must be given very careful consideration. If correct, there will be major consequences for a theory of infant cognitive architecture. The remaining parts of this section address this question. First, I consider whether these results could stem entirely from the operation of infant input systems and therefore not imply thought. Some will find the assumptions I make about the powers of infant input systems rather liberal. Even so, evidence from illusions leads me to conclude that input systems are not responsible for the crucial feature of Baillargeon's results. I then consider what properties of infant thought account for the results.

Illusions and impossible events

Let us assume that input systems function to build and maintain a model of the perceptual world that is rich enough to allow a conceptual identification of input. At the least, this implies a description of a perceptual situation that extends far enough in time and space to allow a local identification of objects and causal interactions. Thus the infant's input systems would pass information to central systems about, for example, the size, shape, and locations of objects, their displacements through space, and some inherent properties like solid or compressible. They would also describe the mechanical interactions of objects.

Next assume that representations built by input systems can include descriptions of objects which are no longer visible or otherwise sensible, as well as for objects traditionally celebrated as being present to the senses. There is evidence that the visual system can construct illusory invisible objects: in experiments on apparent motion, for example, a shape can appear to move and 'hide' behind another (Ramachandran and Anstis 1986). It is very likely then that input systems can describe situations with hidden objects. The results of Bower (1967) indicate that this is likely in the case of infants too.

With these assumptions in mind, let us consider again what seems to be happening in Baillargeon's experiments. The infant is surprised when an apparently impossible event occurs. A block is seen in a certain position behind a screen. The screen rotates upwards but the box is stationary as it is occluded. Then either the screen or some other object appears to move through the space still occupied by the box. This sequence of events creates an incongruity between one representation that says that a certain region of space is occupied by a rigid solid object and another representation that says that another solid object has just passed through that space. The detection of such incongruities will be the task of a system which seeks to maintain consistent and non-superficial models of a region of space through time.

Is this task carried out by the input systems themselves? One might try the following argument. Since these systems work bottom-up and without access to central information, they will not be able to access the earlier representation of the blocking object when the later passing-through event takes place. They would therefore not be able to detect the contradiction. This will require instead the use of central memory resources. However, avoiding the problem of not having access to earlier representations may have led to a solution in which input systems are specifically designed by evolution to hold onto representations of the objects in the current space. In which case, input systems would be able to detect such incongruities. This sort of *a priori* argument, then, is too weak to be of much use to us.

Much better would be evidence that input systems are actually quite happy with the idea of one object passing through another. Here evidence from illusions is, for obvious reasons, crucial. And, in fact, an illusion does exist where one object appears to pass through another (Ramachandran 1985). However, there are two immediate problems with citing this particular illusion as evidence in our case. The illusion involves the apparent motion of a light through a hand. First, there may be something special about apparent as opposed to real movement (e.g. Kolers 1964), and, second, a light is not a solid object.

Sperber, speaking in the discussion period, has put the following point to me: 'Part of the function of input modules is to filter out most information and to filter *in* potentially relevant information. Incongruities in the environment are typically relevant to the organism and should therefore be filtered in to become objects of central attentive processes'. This suggests a simple explanation: the infant's input systems detect the incongruity in Baillargeon's experiment and alert central systems to pay more attention. Nothing of interest would follow as regards infant thought.

It is not the case, however, that input systems always filter in incongruities, sometimes they resolve conflicts and produce new illusions as a result. For example, stereograms can be used to create an incongruence between interposition and binocular information for the relative positions in depth of two planes (Zanforlin 1982). In this case, the visual system resolves the contradiction by bending one of the planes round the other.

The power of input systems to resolve incongruities can also be seen in intermodal illusions. In the 'McGurk effect' a listener is exposed to an auditory 'ga' while watching the speaker make the lip movements for 'ba'. Under these circumstances the looker/listener hears neither 'ba' nor 'ga' but an intermediate 'da' (McGurk and Macdonald 1976). The incongruity between visual and auditory input is resolved by the input systems in a striking illusion.

So it would be of great interest if the perceptual resolution of an incongruity resulted in an illusion of one object passing through another. It would suggest that this was more acceptable to vision than the original incongruity. The Ames trapezoidal window with rod illusion might fit this bill. In this a trapezoid seems to rotate back and forth while a rod projecting through the centre seems to rotate continuously through 360°. According to Rock (1983), however, it is not clear exactly what is seen at the moment when the rod should pass through the side of the window. Rock also points out that the conflicting interposition information is available to the visual system only very briefly at this instant. This illusion too, then, is not quite what we are looking for.

The following kind of evidence is needed; a robust and clearly describable illusion in which one solid rigid object is seen to pass through another solid rigid object; the illusion arises from the visual system's attempt to resolve an incongruity; and it occurs despite the continuous availability of perceptual information that conflicts with the resolving (illusory) percept. This is quite a complex specification and I despaired of ever finding such a phenomenon. Then Wilson and Robinson (1986) published their observations on the Pulfrich double pendulum illusion.

Seeing is not believing

The Pulfrich double pendulum (PDP) illusion is actually a set of simultaneous illusions. Wilson and Robinson (1986) constructed two pendulums using rigid metal rods with plastic detergent bottles filled with sand on the end. The

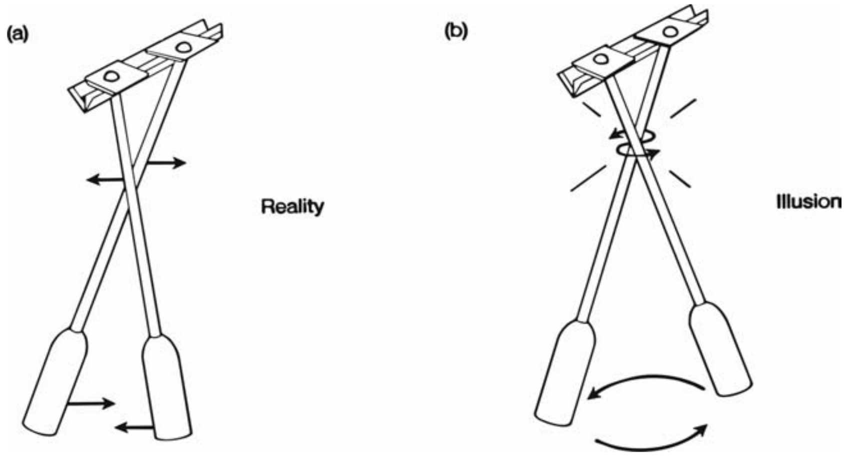

Fig. 8.6 The Pulfrich double pendulum illusion: (a) what really happens, and (b) how the illusion appears to an observer viewing with reduced luminance to one eye (direction of apparent rotation depends upon which eye).

pendulums are then mounted so that they swing in parallel, one slightly behind the other. The arrangement is viewed in fairly dim light with the pendulums set to swing in opposite phase in a frontal plane (see Fig. 8.6). The observer holds a neutral density filter over one eye but looks with both. The reduced luminance to one eye creates a time delay in signals from it and thus a stereoscopic discrepancy in the position of the pendulum which varies with the velocity and direction of swing. Stereoscopic fusion interprets this discrepancy as a variation in depth and the pendulum is seen to swing in an ellipse. With two pendulums in opposite phase, two elliptical paths are seen and the pendulum bobs appear to be chasing each other around without, somehow, the rods twisting round each other.

Wilson and Robinson (1986) also describe a concomitant size illusion due to inappropriate size constancy scaling with the pendulum bobs appearing to grow as they recede and shrink as they approach. What Wilson and Robinson do not describe, however, is what observers see happening to the rods. They say that observers do *not* see them twisting round each other, but they do not say what observers *do* see. How does the visual system resolve the incongruity in the overlapping orbits of the two pendulum swings created by the stereoscopic illusion? It seemed there might be a chance that it would have the rods pass through each other. Robinson (pers. comm.) confirmed that Wilson and Robinson (1986) had not studied this aspect of the illusion.

I have therefore investigated this myself with a similarly constructed PDP.

The results were clear (Leslie, in preparation *a)*. First, I can confirm that the PDP illusion as reported by Wilson and Robinson, including the elliptical paths, the 'chasing round', and the size illusion, is striking and easily obtained. Equally striking is the clear perception of the rigid solid rods passing through each other. Most observers were able to find an angle of view where even the pendulum bottles appear to pass through one another despite their large size and marked surface texture.

The PDP illusion satisfies the conditions I laid down. First, the illusion is robust and easily described, most viewers spontaneously offering the observation that the rods were passing through each other. Second, this seems to arise from the visual system resolving a perceptual conflict which is itself due to an illusion. Most impressively, however, there is conflicting interposition, convergence, and retinal size information continuously available that the pendulums are not varying in depth and not passing through each other. Presumably, the visual system could have resolved this in some other way; for example, by bending the rods and momentarily twisting and untwisting them, or by simply not specifying clearly what happens at the cross-over point, or indeed by suppressing the stereoscopic illusion altogether. Instead, an illusion of passing through occurs. This suggests that the visual system is really rather happy with the idea of solid objects passing through one another.

Baby knows better

Let us return to the infants in Baillargeon's experiments. These infants seem to have knowledge that solid objects cannot cohabit the same space even temporarily. The adult visual system, on the other hand, despite a great deal of time to detect this regularity about the behaviour of objects in the familiar world—despite never having seen a counterexample in 40 years—does not seem to have learned it and is perfectly prepared to advance this bizarre percept as soon as it is shown the PDP illusion. Such obstinate ignorance would be difficult to understand if input systems were simply mechanisms of associative learning. Instead it points to a different kind of organization—one which is designed to provide central learning mechanisms with the right conceptual identification of input. Such identifications may carry an *implication* of mechanical incongruity which input systems cannot detect, nor resolve.

A *central* learning mechanism is, I believe, the key to understanding Baillargeon's results. First the infant input systems provide central thought with the representation that a solid rigid object is in a certain location throughout. Then a little later they advance the representation that another solid rigid object has just passed through this location. So far there is no contradiction. Contradiction only arises in conjunction with a third proposition, namely, that solid objects cannot occupy the same space. But the results from the PDP illusion show that, unlike the other two, the source of this third proposition cannot be perception.

Yet infant thought does appear to apply a principle of no cohabitation to solid objects. What I have to do now is to try to understand how this principle is embodied and applied in thought. I shall follow Sperber and Wilson (1986) and assume that central thought employs a general system of spontaneous *deductive* inferences.

An engine of development

I think the reason the principle of no cohabitation exists in the form it does is that there is a central system which, in conjunction with input mechanisms, maintains a consistent and non-superficial model of the infant's current environmental situation. We all need this bit of architecture no matter what age. But the same system has another related function which is particularly important in development. This function is to build encyclopaedic knowledge and common-sense theories about the mechanics of the physical world.

I am going to postulate two parts to this system. First, a set of spontaneous deductive inferences, and, second, a set of principles which enter into these inferences along with other representations. These other representations may include further principles, perceptual representations, and encyclopaedic knowledge. This system in Baillargeon's infants detects the *logical* contradiction in holding three things to be simultaneously true: (a current perceptual representation that) one solid object has traversed a certain trajectory, (a representation received previously and now in memory that) another solid object has all the while sat astride that trajectory, and (a principle representation that) solid objects cannot share the same space.

What distinguishes principles from other representations, aside from their origins, is their inviolability. That is, in the face of apparent counterevidence principles are not disconfirmed. Instead such evidence is immediately doubted or the system looks for other ways to escape from interpretations that lead detectably (by spontaneous deduction) to contradiction of a principle. In short, apparent violation of a principle creates paradox and not disproof.

The privileged status of no cohabitation gives this principle its power in the learning system—the engine of development—-that builds and constrains the child's encyclopaedic knowledge and common-sense theories about the physical world. This system is apparently functioning by four months of age (Baillargeon 1987*b*) and probably serves us in essentially the same role throughout life.

Spelke's objecthood principles: distinguishing perception and thought

Spelke (1987, 1988) has made important proposals regarding the infant's core concept of an object. According to her theory, this core concept consists of four principles: *boundedness, cohesion, spatio-temporal continuity,* and *substance.* This last principle is what I have discussed as 'no cohabitation', though Spelke may not agree with my proposals for how it is embodied.

Spelke is skeptical about past attempts to distinguish perception from thought, and in particular about the role that the notion of modular organization might play in such a distinction. She argues that if there is any principled distinction between perception and thought it will be that they deliver different kinds of knowledge. Perception delivers knowledge about the continuous surface layout of the world in continuous change—producing representations like Marr's (1982) 2½-D sketch. Thought breaks this continuous layout up into units—into objects and events—and finds relations between these units. The units and relations thought finds in the world are intimately related to the theories thought builds and entertains about the world. The infant's object constructing principles are an example of this function of central thought.

I cannot hope to do justice to Spelke's ideas and results here, but I do want to respond briefly to her arguments on the differences between perception and thought. For the sake of argument, I shall assume Marr's (1982) view of the organization of visual perception. According to this, early vision culminates in a viewer-centred representation of surface layout which Marr called a '2½-D sketch' (Marr 1982; Marr and Nishihara 1978). This representation is arrived at entirely bottom-up as a function simply of the retinal array. This kind of early recoding of retinal arrays can be thought of, therefore, as a kind of extended *sensory analysis.* It is to this, however, that Spelke wants to restrict the term *perception.*

The next level of representation in Marr's account is the object-centred representation called the '3-D model'. This goes beyond sensory analysis in the sense that representations at this level are only partial functions of retinal arrays. Additional information, for example, from a catalogue of three-dimensional object shape descriptions, is used to disambiguate viewer-centred representations and to categorize objects (Marr and Nishihara 1978). There is neuropsychological evidence that a 'pure visual' object recognition module operates independently of and prior to a module for recognizing object function or 'meaning' (Warrington and James 1986). Also at this level, according to Ullman's (1984) theory of visual routines, there are processes of visual analysis which are responsive in highly restricted ways to goals set by central attention. For example, 'optional' visual analyses, like fast curve-tracing, can be performed in support of the recognition of particular objects or other special tasks (Jolicoeur *el al.* 1986), while the influence of set on the perception of illusory contours (Coren *et al.* 1986) suggests that this class of illusions may involve some kind of central triggering information. Let us call this level of input processing *perceptual analysis.*

It is the level of perceptual analysis, interfacing sensory analysis and thought, that results in a conceptual identification of input. Physical objects and physical events are parsed and related at this level. Mandler (in press) uses the term perceptual analysis in a somewhat similar way; I do not want to suggest, however, that this represents the highest level of infant cognition, as the previous discussion of Baillargeon's findings should have made clear. On the other hand, it does seem likely that some of Spelke's objecthood principles are implemented at the level of perceptual analysis. For example, boundedness and cohesion would seem to be required for Marr-Nishihara-type object recognition. By contrast, the principle of substance or no cohabitation must belong to central thought if the arguments presented earlier are accepted. This means that infants construct concrete objects over a number of cognitive levels.

Causal inference and metarepresentation

Toward the end of infancy a capacity for a new kind of internal representation emerges. This first shows itself in the ability to pretend. Instead of being directed at representing the world in a faithful and literal way, as perception and the kind of thought we have been considering are, pretence involves a deliberate distortion of the way the real situation is understood. I have been trying to understand the cognitive mechanisms that make this possible but will not say much about this aspect here (see instead, Leslie 1987*a*, in press *a).* I do want to describe briefly a study of inferential processes with respect to imaginary states of affairs in two-year-olds (Leslie, in preparation *b).* This study demonstrates counterfactual causal reasoning and has important implications for early mental architecture.

Sharing pretence with young children can be turned into a flexible experimental method. I require the child to follow what I am pretending, encouraging him to join in as much as possible. For example, I show the child two empty toy cups, a toy bottle, and some toy animals and I describe the setting, giving a birthday party for one of the animals, as a cover story for later 'events'.

I ask the child to 'pour out' some 'water' into the two cups. I then pick up one of the cups and turn it upside-down for a moment or two and then replace it. I ask the child which cup is empty/full. The child can either point, say which one, or 'refill' the 'empty' cup (both are really empty). To get this right, the child has to keep track of the pretend status of the two cups. He must watch what I do and interpret my actions with respect to the pretend world we jointly create. Somehow he must calculate the 'consequences' of those actions in the pretend world, as well as perceive the actual results in the real world. Children of around two-and-a-half years seem to enjoy this task and are very good at making appropriate causal inferences.

During the 'birthday party' a regrettable incident takes place in which one of the animals picks up a cup which the child has recently 'filled' with 'water' and proceeds to upturn the cup above the head of another animal, holding the cup upside-down in this position. I ask the child what has happened. The children usually answer that the water has gone all over the victim or that the victim is wet. Again causal inferences appropriate to the pretend but not to the real state of affairs are made. Some children draw a different conclusion and 'refill' the cup.

One regrettable incident leads to another, and soon the child has inferred that one of the animals has become 'muddy' by rolling on a certain region of the table I have designated as a muddy puddle. I suggest that the animal is in need of a bath (the children never seem to think of this themselves) and make a 'bath' using four toy bricks arranged to produce a cavity. I place the animal in this cavity and roll it around a few times, then remove it. The child might then pick up a 'towel' to 'dry' the animal. I say, 'Watch this' and pick up one of the toy cups. I then put the toy cup into the cavity formed by the bricks and make a single scooping motion. I then ask the child, 'What's in here?', pointing to the empty cup. The child replies, 'Water'.

The language of thought in pretence

The two-year-old is following the pretend scenario which we jointly construct by representing imaginary events, imaginary objects, and properties and by calculating the imaginary consequences of imaginary states of affairs. But the inferences used by the two-year-old to do this are real world inferences—they are not just fantastic and random leaps from one pretend state of affairs to another. This is interesting for several reasons. First, it shows that even early pretending involves highly constrained thought processes. It also demonstrates counterfactual reasoning which employs the same knowledge used in understanding and predicting the real world. This is exactly what was expected on the basis of the cognitive model of pretense proposed in Leslie (1987*a*, in press *a).*

These results further show that the two-year-old's real world knowledge is not represented in such a way that it is bound to specific contexts. The perceptual support provided by my pretend scenarios is minimal. The props used and my actions and words at best *suggest* a story-line, but this story-line cannot be simply perceived nor be computed from perception by central thought in the normal way (see Leslie 1987*a*). The inferences used, while being drawn from the set of real world inferences the child can make, nevertheless have to apply to different representations than those that arise in understanding the real current situation. They have to apply to representations of the pretence and not to representations of what is really happening. They must also produce other pretense representations as their output, and not serious ones (Leslie 1987*a*).

Inferential processes are only one part of this complex cognitive activity in two-year-olds. One must also consider the nature of the representations that are being processed. I will outline some main points here (but, for more extended discussion, see Leslie 1987*a*, in press *a,* in press *b).*

Pretence representations have computational properties that distinguish them not only from representations of actual situations but from any serious, literal representation, even ones considered false. They belong to the class of *metarepresentations*—that is, the class of representations that relate agents to representations of representations. Sentences which report direct speech are a natural model for metarepresentations. So, for example, in *John said, 'Computer hardware is infallible',* John is related to a sentence or representation. But because this sentence or representation is quoted and not asserted, one cannot make normal inferences from it—in particular one cannot infer that computer hardware does not break down.

In fact, there is a detailed correspondence between the inferential properties of another related class of sentences in language and the inferential properties of pretence (Leslie 1987*a*). These are sentences like *John believes computer hardware is infallible*—sentences which report mental states. This correspondence suggests that pretending and mental state reporting depend cognitively upon the same underlying form of representation. This form of representation must have certain crucial inferential properties.

Consider the following as thoughts:

(1) the cup is full of water;
(2) the empty cup is full of water;
(3) I pretend the empty cup is full of water;
(4) I pretend the cup is both empty and full of water.

There are internal contradictions in (2) and (4) but not in (1) and (3). I do not think that we or young children can have (2) and (4) as thoughts in the ordinary way because the logical contradiction is so blatant and is soon picked up by spontaneous deduction. The puzzle is why (3) does not suffer this defect while (2) and (4) do.

The answer I give is roughly this (see Leslie 1987*a*, in press *a).* The internal representation of the thought (3) has more structure than is apparent in the way it is written down. Part of the expression is actually quoted or, as I say, *decoupled:*

(5) I pretend the empty cup 'it is full of water'.

Inferential processes have to respect this structure. Suppose there was a causal inference to do with what happens when things that contain water are turned upside-down. If this were to apply to (1) it might output something like (6):

(6) the water pours out and makes something wet—the container becomes empty.

Used in pretence, this inference would apply to the decoupled part of (5). Since the input to the inference is decoupled, its output too will be decoupled. This

ensures that the 'conclusion' is part of the pretence and not a prediction about the real world. Thus one of the 'conclusions' when the inference is applied to
(5) would be:

(7) I pretend the empty cup 'it is empty'.

The thought (7) is not a mere tautology. In fact, it is a particularly interesting case of pretence because it shows that pretend representations are not merely marked as false (see Leslie, in press a). In the extended pretence going from 'filling' the cup with 'water' to 'emptying' it to 'refilling' it again, of which (7) is a part, the cup really is empty throughout. If parts of pretend representations were simply marked as false, they could not be used to produce this kind of pretense. Leslie (in press *a)* gives further reasons why a 'mark as false' account of pretense will not work.

Recall the *spontaneous deductive* inferences, discussed in the middle part of this chapter (p. 200), which detected the contradiction between the location of one object, the trajectory of another, and the principle of no cohabitation. These spontaneous inferences will not detect a contradiction in (5) since the elements which would have been incongruous are at different levels; i.e. decoupled and non-decoupled. However, if I write out (4) in full, to give

(8) I pretend the empty cup 'it is both empty and full of water',

one can see that here the contradictory elements are at the same level—as they also are in (2). Spontaneous deduction should immediately detect this within levels contradiction. This is why one never finds children who think like (2) or pretend like (8).

Inferences in pretence: evidence for symbolic processing

One of the most fundamental questions about infant mental architecture concerns the computational organization of the processing hardware. Recently it has been claimed that cognitive psychology has been mistaken when it assumed that (all or any) adult computational processes involved the manipulation of symbolic codes (e.g. Rumelhart *et al.* 1986). The suggestion is based on the study of a quite different computational architecture from the familiar serial processing, by rules, of symbol strings read from and written to a memory store (Newell 1980). In 'connectionist' systems, there are no symbolic representations, no representation of the processing rules, and no distinct memory stores containing symbols. Yet these connectionist systems have interesting powers of associative learning. The question arises whether in the early stages of development a connectionist architecture might provide the entire basis for cognition.

I think that the existence of a capacity for pretence rules out this possibility. Connectionist architecture, while it may be able to simulate pretence, is, as far as I can see, inherently incapable of providing a principled explanation for the most important properties of pretence cognition and related phenomena. Whatever the hardware, it must keep serious and pretence-related cognitions apart and distinct. An organism that confused its serious knowledge of the world with its pretence would be in trouble. Because pretence is part of the capacity to represent different mental models of the world, it is a special case of a much more general system of cognition underlying our ability to model other minds (Leslie 1987*a*, *b,* in press *a* and *b).* There are thus equivalent requirements to keep apart and distinct (representations of) my pretend from your pretend, my beliefs from your beliefs, your hopes from my beliefs, my beliefs from my beliefs about my beliefs, my hopes about your beliefs from my beliefs about your hopes, and so on and on. To handle these different representational 'spaces' and the differences in their content will require in connectionist machinery functionally distinct networks.

Using functionally distinct networks would probably allow a simulation to be built. In simulation one could attempt to construct networks whose 'contents' had shared properties. On the other hand, it would be just as easy to construct networks whose 'contents' were arbitrarily different. There is nothing in connectionist architecture to prevent functionally distinct networks from differing arbitrarily. But this fact will deprive us of a principled explanation should it be the case that different 'mental spaces' are *always* related in their content.

Unfortunately for connectionist models, the contents of metarepresentational states are always deeply and systematically related to one another. In fact, this is the first thing any theory in this domain must account for. These states are individuated in three important ways: first, in terms of whose state it is; second, in terms of the relation involved (e.g. *pretend, believe, hope, expect* and so on); and third, in terms of the content of the state—whether I believe that it is raining, or that Edinburgh is a beautiful city, or that Leslie discovered the connection between heat and light. Two states then may differ but share exactly the same content: there is a non-arbitrary relationship between *pretending it is raining* and *believing it is raining.* What they have in common is the proposition *it is raining.* Or one content may be the negation of another: *believing it is not raining.* And so on with endlessly many relations.

These facts—both the differences between different metarepresentational states and the systematic relations between their possible contents—can be parsimoniously accounted for using a system of symbolic computation (Pylyshyn 1984; Leslie 1987*a*). For example, the differences between serious and pretense-related cognition can be captured by the differences between the forms of the underlying representations. Their systematic relations meanwhile are given by relations between subexpressions in the symbolic code. So the full *of water* that features in pretense is the same *full of water* that features in serious cognition. Since all the different 'mental spaces' use the same symbolic code, systematic relations of this sort are ubiquitous and inevitable.

Finally, the fact that both types of representation are subject to the same computational processes (e.g. the same rules of inference) also receives a principled explanation in a symbolic processing account. In connectionist machines there are no computational processes identifiable independently of the network. So different networks are perfectly free to vary in, for example, the rules of inference they implement. This deprives such architectures of a principled explanation of the fact that in human children pretence employs the same inferential processes as serious cognition. In symbolic processing architectures, however, it is fundamental that there are computational processes which apply to symbolic expressions and which are sensitive to the structure of those expressions; that is what a symbolic computation system is. We can therefore readily find an explanation for why the same inferences apply and why these inferences respect the structural differences, as well as the structural similarities, between primary representations and metarepresentations.

Pretence, then, provides powerful evidence in favour of an infant mental architecture that includes symbolic processing. Because metarepresentation presupposes primary representation, it is likely that symbolic processing devices have been operating throughout most of infancy. The arguments and evidence discussed earlier in this chapter, regarding the relationship between perception and thought in infancy and the logical properties of infant representations, confirm and support the existence of a symbolic processing architecture during human infancy.

Conclusion

The main organizational features of the adult mind appear to be present in infancy. I have argued for a modular organization in infant perception and pointed to its advantages for development. Central thought processes appear to operate early and, like perception, are richly structured, presumably by biological endowment. They employ powerful inferential processes which are sensitive to the logical properties of infant symbolic representation.

Towards the end of infancy, thought acquires the power to represent itself recursively and thereby to reason imaginatively. This will provide the basis for the conceptual distinction between appearance and reality (Leslie in press *a*). This distinction will allow central processes to theorize about those things in experience that are incorrigibly not what they seem. The necessity of illusion comes home to roost.

Figure 8.7 summarizes the argument of this chapter. The main conclusion appears to be that human mental architecture provides the basis for development and not its outcome. Should this seem strange, we should reflect that acquiring theoretical knowledge of the world—in the sense both of common sense and of more specialized scientific and religious theories—is

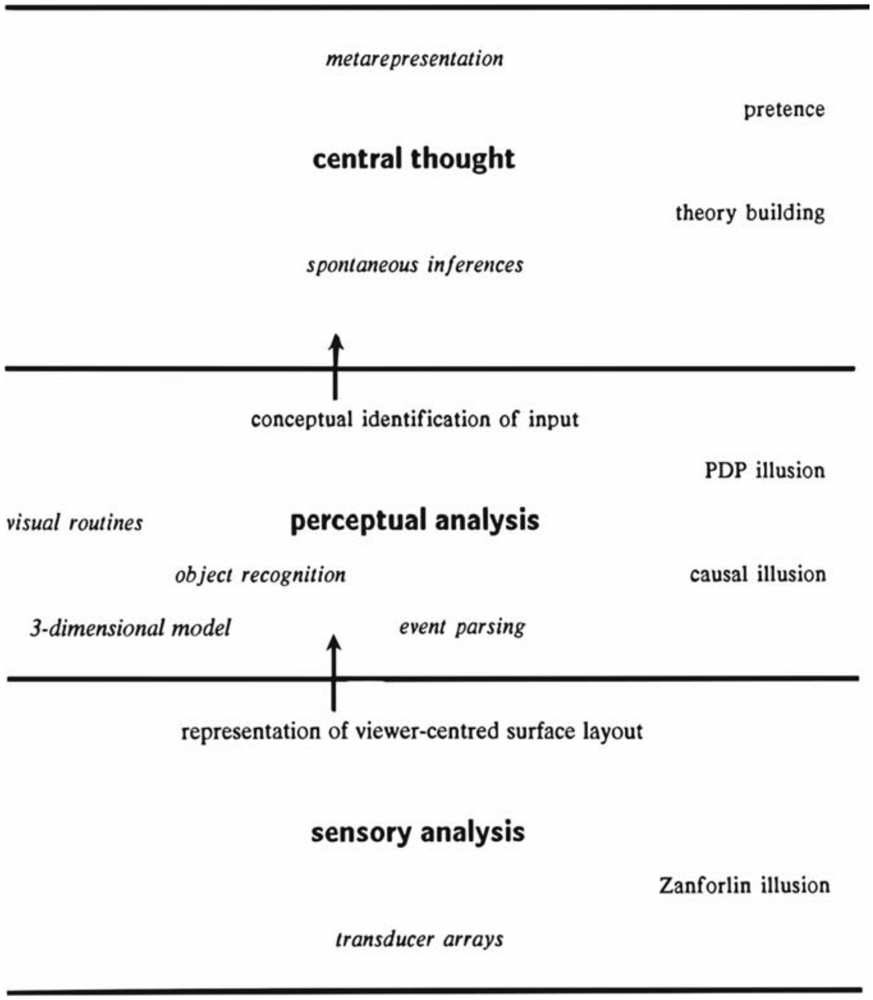

Fig. 8.7 Summary of main arguments concerning mental architecture in infancy.

uniquely the point of human development. The basic organization of the infant-adult mind is highly designed for this task.

Acknowledgements

I am grateful to John Morton for comments on an earlier draft, and to Jean Mandler and Elizabeth Spelke for highly relevant discussion.

References

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.

— (1987a). Young infant's reasoning about the physical and spatial properties of a hidden object. *Cognitive Development,* **2**, 179-200.

— (1987b). Object permanence in 3.5- and 4.5-month-old infants. *Developmental Psychology,* **23**, 655-64.

—, Spelke, E. S., and Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition,* **20,** 191-208.

Bower, T. G. R. (1967). The development of object permanence: some studies of existence constancy. *Perception and Psychophysics,* **2***,* 74-6.

Bullock, M. (1985). Causal reasoning and developmental change over the preschool years. *Human Development,* **28**, 169-91.

—, Gelman, R.,and Baillargeon, R. (1982). The development of causal reasoning. In *The developmental psychology of time,* (ed. W. Friedman), pp. 209-54. New York.

Coren, S. and Ward, L. M. (1979). Levels of processing in visual illusions: the combination and interaction of distortion-producing mechanisms. *Journal of Experimental Psychology: Human Perception and Performance***, 5***,* 324-35.

—, Porac, C., and Theodor, L. H. (1986). The effects of perceptual set on the shape and apparent depth of subjective contours. *Perception and Psychophysics,* **39,** 327-33.

Fodor, J. A. (1983). *The modularity of mind.* MIT Press, Cambridge, Ma.

Gregory, R. L. (1974). *Concepts and mechanisms of perception.* Duckworth, London.

Jolicoeur, P., Ullman, S., and Mackay, M. (1986). Curve tracing: a possible basic operation in the perception of spatial relations. *Memory and Cognition,* **14**, 129-40.

Kolers, P. A. (1964). The illusion of movement. *Scientific American,* **211**, 98-106.

Kun, A. (1978). Evidence for preschoolers' understanding of causal direction in extended causal sequences. *Child Development,* **49**, 218-22.

Leslie, A. M. (1982). Discursive representation in infancy. In *Knowledge and representation,* (ed. B. de Gelder), pp. 80-93. Routledge and Kegan Paul, Andover, Hants.

— (1984). Spatiotemporal continuity and the perception of causality in infants. *Perception,* **13,** 287-305.

— (1986). Getting development off the ground: modularity and the infant's perception of causality. In *Theory building in development,* (ed. P. van Gert), pp. 405-37. North-Holland, Amsterdam.

— (1982a). Pretense and representation: the origins of 'theory of mind'. *Psychological Review,* **94,** 412-26.

— (1987b). The child's understanding of the mental world. In *The Oxford companion to the mind,* (ed. R. L. Gregory), pp. 139-42. Oxford University Press.

— (in press a). Some implications of pretense for mechanisms underlying the child's theory of mind. In *Developing theories of mind,* (ed. J. Astington, D. Olson, and P. Harris). Cambridge University Press.

— (in press *b).* A 'language of thought' approach to early pretense. *Cahiers de la Fondation Archives Jean Piaget.*

— (in prep. *a). Further observations on the Pulfrich double pendulum illusion.* MRC Cognitive Development Unit, University of London.

— (in prep. *b). Causal inferences in pretense: evidence for symbolic processing in two year olds.* MRC Cognitive Development Unit, University of London.

—, and Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition,* **25**, 265-88.

Mandler, J. (in press). How to build a baby: on the development of an accessible representational system. *Cognitive Development.*

Marr, D. (1982). *Vision.* (W. H. Freeman, San Francisco, CA).
—, and Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, B,* 200, 187 217.
McGurk, H. and Macdonald, J. (1976). Hearing lips and seeing voices. *Nature,* **264**, 746-8.
Michotte, A. (1963). *The perception of causality.* Methuen, Andover, Hants.
Newell, A. (1980). Physical symbol systems. *Cognitive Science*, **4***,* 135-83.
Pylyshyn, Z. W. (1984). *Computation and cognition: toward a foundation for cognitive science.* MIT Press, Cambridge, MA.
Ramachandran, V. S. (1985). The neurobiology of perception. *Perception,* **14**, 97-103.
—, and Anstis, S. M. (1986). The perception of apparent motion. *Scientific American,* 255, 80-7.
Robinson, J. O. (1972). *The psychology of visual illusion.* Hutchinson, London.
Rock, I. (1983). *The logic of perception.* MIT Press, Cambridge, MA.
Rumelhart, D. E., Hinton, G. E., and McClelland, J. L. (1986). A general framework for parallel distributed processing. In *Parallel distributed processing: explorations in the microstructure of cognition. Vol. 1,* (ed. D. E. Rumelhart and J. L. McClelland), pp. 45-76. MIT Press, Cambridge, MA.
Shultz, T. R. (1982). Rules of causal attribution. *Monographs of the Society for Research in Child Development*, **47**, No. 1.
Spelke, E. S. (1987). Where perceiving ends and thinking begins: the apprehension of objects in infancy. In *Perceptual development in infancy. Minnesota symposia on child psychology,* (ed. A. Yonas), pp. 197-234. Lawrence Erlbaum Associates, Hillsdale, NJ.
—, (1988). The origins of physical knowledge. In *Thought without language,* (ed. L. Weiskrantz), pp. 168-83. Oxford University Press.
Sperber, D. and Wilson, D. (1986). *Relevance: communication and cognition.* Blackwell Scientific Publications, Oxford.
Ullman, S. (1984). Visual routines. *Cognition,* **18,** 97-159.
Warrington, E. K. and James, M. (1986). Visual object recognition in patients with right-hemisphere lesions: axes or features? *Perception,* **15**, 355-66.
Wilson, J. A. and Robinson, J. O. (1986). The impossibly twisted Pulfrich pendulum. *Perception,* **15,** 503-4.
Zanforlin, M. (1982). Figure organization and binocular interaction. In *Organization and representation in perception,* (ed. J. Beck), pp. 251-67. Lawrence Erlbaum Associates, Hillsdale, NJ.