12

Cognitive Development

The topics into which the field of developmental psychology can be divided are as many and as diverse as the topical divisions for psychology as a whole. The organizational scheme that we will follow in the next two chapters—a division into "cognitive" and "social"—is perhaps the broadest and most general cut that can be made. This scheme does not encompass every possible topic in the field, nor is the borderline between cognitive and social always clear. Nevertheless, the division is a typical and generally useful one.

The coverage of cognitive development in the present chapter is divided into five sections. The chapter begins with discussion of two of the most influential general approaches to the study of children's intelligence: the Piagetian approach and the intelligence test or IQ approach. The remaining sections are devoted to three of the most active topics for current research in cognitive development: memory, theory of mind, and conceptual development.

The age period for most of the research considered in this chapter is from 2 to 16—that is, postinfancy childhood. I discussed the Piagetian approach to the study of infant cognition in chapter 11, and we will briefly consider some other forms of infancy research at various points in this chapter. We will also return to several of these approaches in chapter 14 in the discussion of research on aging.

The Piagetian Approach

Piaget's Studies

We saw in chapter 11 that Piaget's work dominated the study of infant cognition for years but is no longer as influential as was once the case. The same can be said of the Piagetian approach to later childhood. This work is even harder to summarize briefly, for it constitutes a much larger literature than the infancy studies—some 25 books by Piaget and associates, as well as literally thousands of related studies by others. The goal of this section is to highlight some central themes and important research examples. Fuller discussions can be found in Flavell (1963), Ginsburg and Opper (1988), Voyat (1982), and Miller (1982).

I begin with an example. Box 12.1 presents two protocols from Piaget and Szeminska's (1952) *The Child's Conception of Number.* The concept under examination is **conservation**: the realization that the quantitative properties of an object or collection of objects are not changed by a change in perceptual appearance. The specific form of conservation at issue is conservation of

Box 12.1. Examples of Responses to Piaget's Conservation of Number Task

Hoc (4; 3): 'Look, imagine that these are bottles in a cafe. You are the waiter, and you have to take some glasses out of the cupboard. Each bottle must have a glass' He put one glass opposite each bottle and ignored the other glasses. 'Is there the same number?—*Yes.*—(The bottles were then grouped together.) Is there the same number of glasses and bottles?—*No.*—Where are there more?—*There are more glasses.*' The bottles were put back, one opposite each glass, and the glasses were then grouped together. 'Is there the same number of glasses and bottles?—*No.*—Where are there more?—*More bottles.*—Why are there more bottles?—*Just because.*'

Boq (4; 7): 'Put as many sweets here as there are there. Those (6) are for Roger. You are to take as many as he has.—(He made a compact row of about ten, which was shorter than the model.)—Are they the same?—*Not yet* (adding some).—And now?—*Yes.*—Why?—*Because they're like that* (indicating the length).—(The 6 in the model were then spread out.) Who has more?—*Roger.*—Why?—*Because they go right up to there.*—What can we do to make them the same?—*Put some more* (adding 1).—(The 6 were then closed up and his were spread out.)—*Now I've got more.*'

SOURCE: From *The Child's Conception of Number* (pp. 44, 75), by J. Piaget and A. Szeminska. New York: Humanities. Copyright ©1952 by Humanities Press.

number: the realization that number is invariant in the face of an irrelevant perceptual change. As the protocols reveal, young children do not at first understand conservation; rather, they tend to judge quantities in terms of immediate perceptual appearance. Thus, to 4-year-old Boq it is obvious that the longer of two rows must contain more sweets.

We can see immediately several similarities between Piaget's studies of infancy and his approach to later periods of development. He again focuses on basic, epistemologically central kinds of knowledge. Just as the object concept represents a major achievement in the infant's mastery of the sensorimotor world, so are principles of conservation central to the older child's capacity for more advanced forms of thought. Object concept and conservation show a more specific similarity as well. Both represent important *invariants:* aspects of the world that stay the same even though other, more obvious aspects are changing. Throughout his research career Piaget was interested in the invariants that the child comes to understand at different points in development. The phrase "comes to understand" reflects still another similarity: Both object concept and conservation, basic though they seem, are not always present; rather, they must be developed in the course of childhood. Certainly one of the reasons for interest in Piaget's work has always lain in his ability to surprise us with respect to what children, at least for a while, do *not* know.

In addition to these similarities in content, the comparison of object concept and conservation reveals some more general similarities in approach to research. Once again Piaget eschews a highly standardized approach in favor of a flexible, discovery-oriented method of probing the child's knowledge. And once again Piaget reports his results mainly in terms of individual protocols rather than group means and statistical tests.

There are also some important differences between the studies of infancy and the studies of later childhood. One obvious difference concerns sample size. As we saw, the sample for the infancy work was limited to Piaget's own three

children. The samples for the work on later childhood are considerably larger and more representative. Beyond this rather general statement it is difficult to say much, for Piaget seldom provided precise information about sample size or composition (although one book, *The Early Growth of Logic in the Child,* does report a total sample of 2,159!). Piaget's failure to describe the samples that he studied is just one of the sins of scientific reporting that he routinely committed. Nevertheless, it is safe to say that his samples for later childhood are much larger than those for infancy.

Some further differences follow from this difference in samples. Piaget's studies of infancy were longitudinal. With the exception of some work on long-term memory (Piaget & Inhelder, 1973), the research on later childhood is all cross-sectional. Similarly, the studies of infancy were within-subject, in the sense that the phenomena of interest were examined in all three babies and interrelations in development were probed for each child. The work on later childhood is, apparently, just about all between-subject. The "apparently" stems from the fact that Piaget often failed to make clear whether a particular conclusion was based on within-subject or between-subject comparisons. Generally, however, the latter appears to be the case.

A final difference concerns the locus for the observations. The studies of infancy were carried out in the home and focused largely on naturally occurring situations and naturally occurring behaviors. The studies of older children have concentrated mainly on elicited responses to tasks presented in some laboratory context. As the protocols from the *Number* book suggest, the procedure may still be more game-like than test-like, and the interchanges between adult and child may resemble spontaneous conversations more than school-like inquisitions. The fact remains, however, that the measurement is of task-elicited behavior in experimentally contrived situations, and not of spontaneously occurring cognitive activities. We will return later to this issue.

It is time to add some more illustrations of Piagetian tasks to the conservation-of-number example. I will note first that Piaget studied conservation in many quantitative domains in addition to number. Indeed, most of the books devoted to cognition in early and middle childhood include tests of conservation. There are studies of conservation of mass, weight, and volume; of length, area, and distance; of time, speed, and movement. All embody the same general approach: Two stimuli are shown to be equal on some quantitative dimension, one of the stimuli is then perceptually transformed so that the quantities no longer *look* equal, and the child is asked whether the quantities are now the same or different. All also show the same developmental progression from perceptually based nonconservation to logically based conservation.

Conservation is just one of dozens of basic logical or physical concepts that Piaget and his coworkers have studied. I settle here for briefly describing two other important examples. One is **class inclusion**: the principle that a subclass cannot be larger than the superordinate class that contains it. Class inclusion is the knowledge, for example, that there can never be more poppies than flowers, or more ducks than birds. Both of these problems were in fact included in Piaget's studies of classification (Inhelder & Piaget, 1964; Piaget & Szeminska, 1952). Box 12.2 presents a third example. The stimuli for the task are a set of wooden beads, most of which are brown but two of which are white. The response of 6-year-old Bis indicates that class inclusion, like conservation, is another basic concept that is not at first present but rather must develop.

In addition to classes, Piaget's research includes a focus on the child's understanding of relations. Of particular interest is the relational concept of **transitivity**. Transitivity is embodied in reasoning of the following sort: If A is equal to B and B equal to C on some quantitative dimension, then A must be equal to C. Or if A is greater than B and B is greater than C, then

Box 12.2. Example of Response to Piaget's Class Inclusion Task

Bis (6; 8): 'Are there more wooden beads or more brown beads?—*More brown ones, because there are two white ones.*—Are the white ones made of wood?—*Yes.*—And the brown ones?—*Yes.*—Then are there more brown ones or more wooden ones?—*More brown ones.*—What colour would a necklace made of the wooden beads be?—*Brown and white* (thus showing that Bis clearly understood the problem).—And what colour would a necklace made with the brown beads be?—*Brown.*—Then which would be longer, the one made with the wooden beads or the one made with the brown beads?—*The one with the brown beads.*—Draw the necklace for me. (Bis drew a series of black rings for the necklace of brown beads, and a series of black rings plus two white rings for the necklace of wooden beads.)—Good. Now which will be longer, the one with the brown beads or the one with the wooden beads?—*The one with the brown beads.'* Thus, in spite of having clearly understood, and having correctly drawn the data of the problem, Bis was unable to solve it by including the class of brown beads in the class of wooden beads!

SOURCE: From *The Child's Conception of Number* (p. 164), by J. Piaget and A. Szeminska. New York: Humanities. Copyright ©1952 by Humanities Press.

A must be greater than C. Such reasoning has been studied most often with respect to length and weight, the usual stimuli being sticks of different length or clay balls of different weight. Whatever the specific quantity involved, the approach is the same: demonstration of the A-B and B-C relations, followed by a request to judge A and C. Note that the quantitative relation between A and C is not perceptually apparent; hence the child must use the information in the initial two comparisons to deduce the correct answer. According to Piaget's studies it is not until about age 8 or 9 that children are capable of such logical reasoning.

The tasks described thus far are directed to thought in middle childhood, or what is labeled the concrete-operational period in Piaget's theory. In a book entitled *The Growth of Logical Thinking From Childhood to Adolescence,* Inhelder and Piaget (1958) examined a more advanced form of thinking, which they found to emerge only around adolescence. This final period in Piaget's theory is the period of *formal operations.* The essence of formal operations is the capacity for hypothetical-deductive reasoning, the ability to go beyond immediate reality to work systematically and logically within the realm of the possible. The prototype of such reasoning is scientific problem solving; Inhelder and Piaget's tasks for studying formal operations in fact consisted mainly of problems drawn from the physical sciences. One of these problems, the pendulum task, is the source for the protocols reprinted in Box 12.3. The child's job is to determine what factor or factors influence the frequency of oscillation of a simple pendulum. Solution requires identifying each of the potentially important variables (weight, length of string, force of push, etc.), systematically testing out each variable while holding the other variables constant, and finally drawing logical conclusions from the overall pattern of results. The table presents examples of both the failure of the younger child and the success of the older child on such problems.

We will consider one final example of a Piagetian task. Not all of the studies fit the logical or physical mold of the work described so far. Piaget's first book, *The Language and Thought of the Child* (1926), was directed to an important component in the child's understanding of the social world: the ability to take

Box 12.3. Examples of Responses to the Inhelder and Piaget Formal-Operational Pendulum Task

PER (10; 7) is a remarkable case of a failure to separate variables: he varies simultaneously the weight and the impetus; then the weight, the impetus, and the length; then the impetus, the weight, and the elevation, etc., and first concludes: *"It's by changing the weight and the push, certainly not the string."*—"How do you know that the string has nothing to do with it?"—*"Because it's the same string."*— He has not varied its length in the last several trials; previously he had varied it simultaneously with the impetus, thus complicating the account of the experiment.—"But does the rate of speed change?"—*"That depends, sometimes it's the same. . . . Yes, not much. . . . It also depends on the height that you put it at [the string]. When you let go low down, there isn't much speed."* He then draws the conclusion that all four factors operate: *"It's in changing the weight, the push, etc. With the short string, it goes faster,"* but also "*by changing the weight, by giving a stronger push*," and "*for height, you can put it higher or lower*."—"How can you prove that?"—*"You have to try to give it a push, to lower or raise the string, to change the height and the weight."* [He wants to vary all factors simultaneously].

EME (15; 1), after having selected 100 grams with a long string and a medium length string, then 20 grams with a long and a short string, and finally 200 grams with a long and a short, concludes: "*It's the length of the string that makes it go faster or slower; the weight doesn't play any role*." She discounts likewise the height of the drop and the force of her push.

SOURCE: From *The Growth of Logical Thinking from Childhood to Adolescence* (pp. 71, 75), by B. Inhelder and J. Piaget. New York: Basic Books. Copyright © 1958 by Basic Books, Inc.

someone else's point of view. Can the child figure out what someone else sees at the moment, or thinks, or feels, or wishes? In particular, can the child make such judgments when the other's perspective is different from the child's own? Such *perspective taking* is critical to social understanding and social interaction. Perspective taking has a converse in **egocentrism**: the inability to break away from one's own perspective to take the perspective of others. It should come as no surprise, in light of the cognitive deficits discussed so far, to learn that Piaget finds that young children are often egocentric. Figure 12.1 shows one basis for this conclusion, the "three mountains" test of visual or spatial perspective taking. After walking around the display, the child is seated on one side of the model; the child's task is then to describe what would be seen by a doll placed at various locations around the display. To 4- or 5-year-olds the answer is clear: The doll sees whatever they see.

I return to several of the tasks just described when I discuss follow-up work. Before doing so, however, it is worth noting a few more points about the general Piagetian approach to research. The interview procedure illustrated in the various protocols in this chapter is referred to as the **clinical method** of testing. Piaget (1929) adopted the term "clinical method" because of the similarity of his approach to that of a skilled clinician attempting to diagnose and treat emotional problems. In both cases the essence of the approach is flexibility: the freedom for the investigator to deviate from preset procedure to probe the individual participant's response in a variety of nonpredetermined ways.

Piaget's fullest discussion of the clinical method is found in one of his early books, *The*

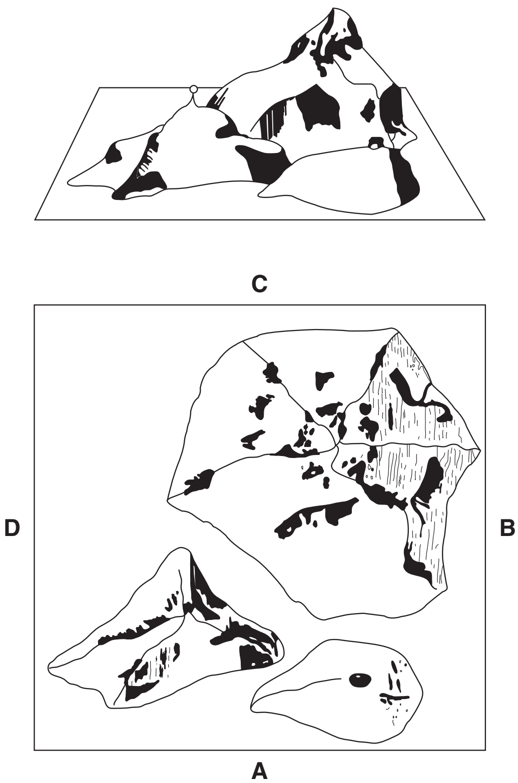


Figure 12.1 Three-mountains task for assessing visual perspective taking

SOURCE: From *The Child's Conception of Space* (p. 211), by J. Piaget and B. Inhelder, 1956, London: Routledge & Kegan Paul. Copyright 1956 by Routledge & Kegan Paul. Reprinted by permission of Taylor & Francis.

Child's Conception of the World. The following passage is worth quoting in full:

> It is our opinion that in child psychology as in pathological psychology, at least a year of daily practice is necessary before passing beyond the inevitable fumbling stage of the beginner. It is so hard not to talk too much when questioning a child, especially for a pedagogue! It is so hard not to be suggestive! And above all, it is so hard to find the middle course between systematisation due to preconceived ideas and incoherence due to the absence of any directing hypothesis! The good experimenter must, in fact, unite two often incompatible qualities; he must know how to observe, that is to say, to let the child talk freely, without ever checking or side-tracking his utterance, and at the same time he must constantly be alert for something definitive, at every moment he must have some working hypothesis, some theory, true or false, which he is seeking to check. To appreciate the real difficulty of the clinical method one must have taught it. When students begin they either suggest to the child all they hope to find, or they suggest nothing at all, because they are not on the look-out for anything, in which case, to be sure, they will never find anything. (Piaget, 1929, pp. 7–8)

The clinical method is an important component in Piaget's research, but it would not mean very much unless there were interesting concepts to which it could be applied. And here we have reached perhaps the greatest strength of Piagetian research: the incredible range of interesting forms of knowledge and insightful procedures for probing this knowledge. This richness can be only hinted at in the brief sampling of tasks and findings included here. Book after Piagetian book is full of novel and informative methods for studying how children think—methods that have set the mold for a substantial proportion of later research on cognitive development. I noted in chapter 1 that technical skill in executing research must always be joined with good ideas about what is interesting to study and how to go about studying it. It is doubtful that anyone else in the history of developmental psychology has had as many good ideas as Piaget.

Issues and Follow-Up Studies

I turn immediately from praise to problems. Piaget's studies have elicited an enormous amount of follow-up research, much of it motivated by perceived deficiencies in the original

Piagetian research. Although this follow-up work has addressed a variety of specific issues, I settle for discussing just one of them here. It is the same basic question that was discussed with regard to the infancy work: that of *assessment*. Do Piaget's procedures really give an accurate picture of the child's abilities?

As with the infancy studies, the main concern has been that Piaget may have underestimated what the young child really knows. Although various possible sources of misdiagnosis have been identified, probably the most common criticism concerns the heavy verbal emphasis in many Piagetian tasks. This emphasis should be clear from the protocols quoted earlier in the chapter. Consider the conservation task. The purpose of this task is to assess the child's understanding of the logic underlying conservation. This logic, however, is not directly observable; rather, its discovery depends on the use of language, both in the questions that are asked of the child and in the responses that the child must make. It seems quite possible that a young "nonconserver" does not really believe in nonconservation but is simply confused by the use of words like "same," "more," "less," and "number." Perhaps, for example, the child thinks that "more" refers to length of row rather than number of objects.

Various approaches have been taken in response to this possibility. Some investigators have used verbal pretests in an attempt to ensure that the child understands the words that are used on the conservation test (e.g., Miller, 1977). Others have gone beyond pretesting to attempt verbal pretraining—that is, teaching the relevant words to the child prior to the test (e.g., Gruen, 1965). And some have attempted to do away with the potentially confusing language altogether by devising "nonverbal" procedures for assessing Piagetian concepts—for example, having the child select a row of candies to eat or a glass of juice to drink (e.g., Miller, 1976a). Such procedures are seldom literally nonverbal; they are simplified linguistically, however, and they do avoid such potentially troublesome words as "same" or "more."

Not all diagnostic revisions have been directed to language. Another concern has been with the general context within which conservation is typically assessed. This context is in fact a rather strange one, and it includes a number of features that may bias the child toward a nonconservation answer. Among these features are the explicit focus on quantity, the seemingly arbitrary nature of the transformation (why is the adult spreading out the candies?), and the presentation of the identical conservation question twice within a short period, a repetition that may suggest to young children that they should change their answers. Perhaps if the context could be made more natural and familiar, the child would be less likely to look like a nonconserver.

This possibility has been tested in various ways. Rose and Blank (1974) examined the effects of the usual two-questions format by simply omitting the initial pretransformation question for half their participants and asking only the final conservation question. McGarrigle and Donaldson (1974) replaced the intentional transformation by an adult experimenter with an apparently accidental transformation by a "naughty" teddy bear, the expectation being that children might find the latter more familiar and less imbued with magical quantity-changing properties. Light, Buckingham, and Robbins (1979) performed a similar manipulation but with an incidental rather than accidental change—that is, a transformation that was not directed solely to the issue of conservation but occurred naturally in the course of an ongoing game.

So far nothing has been said about results from such modified-assessment studies. Three general conclusions seem tenable (for further discussion, see Chandler & Chapman, 1991; Miller, 1976b; Siegal, 1991). The first is that Piaget's methods do in fact result in some underestimation of the young child's abilities, for children often perform better on modified tests than on standard Piagetian tests. The second is that the underestimation is probably not

great, and that phenomena such as nonconservation are by no means totally explicable on the basis of verbal confusions or contextual biases. The third is that the development of concepts such as conservation and perspective taking is more extended and multifaceted than Piaget imagined, with a number of earlier levels and precursor skills not tapped by Piaget's own procedures. This conclusion comes not only from the studies discussed but also from programs of research whose explicit focus has been on simpler and developmentally earlier skills than those examined in Piaget's research. Notable in this regard is the work of John Flavell on perspective taking (Flavell, 1992) and Rochel Gelman on concepts of number (Gelman, 1991).

The Intelligence Test Approach

The Nature of IQ Tests

The intelligence test, or IQ approach, is in many ways quite different from the Piagetian approach. Because the emphasis so far has been on Piaget, it is instructive to begin by noting what some of the differences are.

Piaget's interest was always in commonalities of development—that is, ways in which all children are alike as they grow. A concept like conservation of number, for example, is eventually mastered by all normal children. More broadly, a stage like concrete operations is eventually attained by virtually every child. What individual differences there are seem to lie in the rate of development, and such differences were never of interest to Piaget. In contrast, the whole point of IQ tests is to identify individual differences among children. Such tests, moreover, measure not only differences but *ordered* differences—we say that one child is "higher" or "lower" in intelligence than another, or that a particular child is "above" or "below" average in intelligence. There is an evaluative component that is impossible to escape from in using IQ tests. The fact that such tests force us to make value judgments about children is one reason that their use has always been so controversial.

From the start Piaget's work was theoretically guided, the goal being to answer basic epistemological questions about the child's understanding of domains such as number, space, time, and causality. Whatever practical applications the work has had (e.g., influences on school curriculum) have come later, and Piaget himself was never a major contributor to such applications. IQ tests, in contrast, have been pragmatically oriented from the start. The first successful IQ test, an instrument designed by Binet and Simon in Paris in 1905, was constructed for the very pragmatic purpose of predicting how well children would do in school. Indeed, the ability to predict school performance was an explicit criterion in the selection of items for the test. IQ tests ever since have had similar practical groundings and practical applications.

A final difference concerns the quantitative emphasis in IQ tests. IQ tests are directed to questions of how much and not to questions of how. What the tests yield is a number that tells how much intelligence a particular child has. Some tests yield several numbers, corresponding to different kinds of intelligence; the approach, however, remains basically quantitative. All that is of concern in scoring an IQ test is how many right answers the child gives, and all that is done with the right answers is to add them together to get an overall point total. The focus is thus on the products of cognitive activity, and not the underlying processes from which these products came. For Piagetians, in contrast, the interest is always more in processes than in products. The attempt in Piagetian research is to move beyond the child's right or wrong answer to identify the qualitative nature of the underlying thought system and the qualitative changes that the system undergoes as the child develops. The Piagetian focus is thus more on the how than on the how much.

What has been said so far about IQ tests has a rather negative sound to it. Such tests are more quantitative than qualitative, more concerned with products than underlying processes, more pragmatic than theoretical in their origins and construction. Nor are these the only criticisms that can be lodged against IQ tests; I have not even mentioned the common complaint that such tests are biased against certain groups. Given these various problems, the obvious question becomes: Why should anyone take IQ tests seriously? What is the evidence that such tests are really measuring intelligence? What, in short, is the evidence for the *validity* of the tests?

The general issue of test validity was discussed in chapter 4. Recall that the validity of a test is generally established through demonstrating that the test correlates with other measures to which it ought to relate, "ought to" either for purposes of pragmatic prediction or for reasons of theoretical cogency. The validation of IQ tests has always rested upon such correlational power. The first such test, that of Binet and Simon, was deemed successful because it was able to differentiate among children who were likely to do well in school and those who were likely to do poorly in school. Ever since the original Binet and Simon test, correlations with school performance or academic achievement tests have been a major validity index for IQ tests designed for children. Typically, such correlations are in the neighborhood of .5 to .6, a figure that indicates a moderately strong but certainly far from perfect relation. The correlational power of IQ is not limited to academic contexts, however. IQ also correlates with occupational status in adulthood and with performance on a wide range of learning and cognitive measures (Jensen, 1981). It is this ability to predict (albeit imperfectly) to so many contexts that clearly require intelligence that constitutes the validity argument for IQ tests as measures of intelligence.

A Sampling of Tests

A number of tests purport to measure intelligence, and they vary along several dimensions. Some provide a single overall score as a measure of global or general intelligence. Probably the best-known test of general intelligence is the Stanford-Binet (Roid, 2003), the direct historical descendant of the original Binet and Simon test. Other tests are more specific in focus. An often-used test with children, for example, is the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), which furnishes a measure of receptive oral vocabulary. The various Wechsler tests (Wechsler, 1997, 2002, 2003) all provide measures of both verbal IQ and performance IQ; summed together, the Verbal and Performance scales yield an overall IQ.

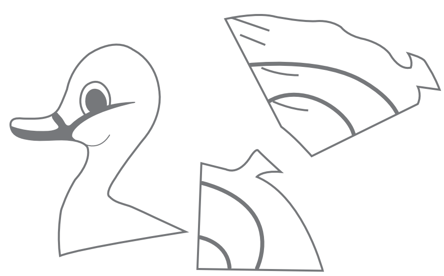
Tests also vary in the age group for whom they are intended. The three Wechsler tests are designed for three different age groups: the Wechsler Preschool and Primary Scale of Intelligence, or WPPSI, is intended for ages 4 to 6; the Wechsler Intelligence Scale for Children, or WISC, is intended for ages 6 to 16; and the Wechsler Adult Intelligence Scale, or WAIS, is intended for adults. The Stanford-Binet is broader in range; although it is most often used within the span of childhood, the test can be given at any age from 2 through adulthood. Finally, although infancy is the only age group excluded from the Wechsler and Stanford-Binet, there are tests that are specifically designed to measure development in infancy—for example, the Bayley Scales of Infant Development (Bayley, 2005).

So far little has been said about the content of IQ tests. Table 12.1 shows examples of the types of items that appear on one of the major tests of childhood IQ: the Wechsler Intelligence Scale for Children or WISC. The other leading childhood test, the Stanford-Binet, is in most respects similar to the WISC. Both tests are composed of a number of subtests directed to different kinds of abilities—11 subtests on the WISC and 10 on the

Stanford-Binet. On both tests the subtests can be grouped into larger scales: a Verbal and a Performance scale on the WISC; scales of Fluid Reasoning, Quantitative Reasoning, Visual-Spatial Processing, Knowledge, and Working Memory on the Stanford-Binet. The kinds of cognitive abilities that are stressed are similar on the two instruments. Verbal skills are important in both tests. Memory is also important, both memory for meaningful material and rote memory for unrelated items. Arithmetical ability is the focus of a number of subtests. So too is reasoning ability. And so is the child's store of real-world factual knowledge, a store that is tapped most explicitly by items such as the Information subtest of the WISC. In general, IQ tests for children are oriented to the kinds of skills that are needed for success in school—vocabulary, memory, arithmetic, problem solving. It is not surprising, therefore, that performance on such tests correlates with performance in school.

The Stanford-Binet and the WISC share other similarities as well. In both tests, a child's IQ is a function of how fast the child is developing in comparison with other children of the

Table 12.1 Simulated Items Similar to Those in the Wechsler Intelligence Scale for Children—Fourth Edition

Subtest	*Items*
Information	How many wings does a bird have? How many nickels make a dime? What is pepper?
Arithmetic	Sam had three pieces of candy, and Joe gave him four more. How many pieces of candy did Sam have altogether? If two buttons cost $.15, what will be the cost of a dozen buttons?
Vocabulary	What is a -----? or What does ----- mean? Hammer Protect Epidemic
Subtest	*Performance scale*
Object assembly	Put the pieces together to make a familiar object. 

SOURCE: *Simulated Items similar to those in the Wechsler Intelligence Scale for Children – Fourth Edition.*

same age. The original formula for calculating Stanford-Binet IQs expressed this conception directly: Intelligence Quotient equals Mental Age (as determined by how far in the test the child was able to go) divided by Chronological Age times 100. For various reasons this formula is no longer used; instead, IQ is based on the deviation between the child's score and the average score for his or her age group. The logic, however, remains the same: Children who are developing faster than average have above-average IQs; children who are developing more slowly than average have below-average IQs. Childhood IQ is thus a measure of rate of development. It is also an inherently *relative* measure. There is no absolute metric for measuring a child's intelligence, as there is, for example, for measuring physical characteristics such as height or weight. Instead, IQ is always a matter of how the child compares with other children.

A final set of similarities between the Stanford-Binet and the WISC concerns the method of administration. There are two central emphases in the administration of any IQ test. One is the need for *standardization.* As was just stressed, IQ tests are relative measures, a child's IQ being a function of how that child's performance compares with that of other children. The only way that a score is interpretable is if the test is administered and scored in the same way for all children. The second emphasis is on the need to establish and maintain *rapport.* An IQ score is supposed to be a measure of the child's optimal performance, and this optimum can be achieved only if the child remains at ease and is motivated to respond carefully. The best testers are the ones who can successfully combine these two goals, maintaining the necessary standardization while at the same time using their clinical skills to elicit the best performance that the child can give.

Issues and Research Paradigms

Many of the issues that have always surrounded IQ tests have concerned their pragmatic uses—for example, tracking children in school on the basis of IQ scores. The concentration here is on more theoretically oriented questions about the development of intelligence. Two such questions have provoked much research and much controversy: the issue of the stability of IQ and the issue of the determinants of differences in IQ.

The consideration of the stability issue can be brief, for most of the relevant points were made in chapter 3 in the discussion of longitudinal designs. Studying stability in fact requires a longitudinal approach, because our interest is in the relation between a child's performance early in life and that same child's performance later in life. The specific form of stability at issue is the stability of individual differences. Do children maintain their relative standing on IQ tests as they develop, those who are high remaining high and those who are low remaining low, or can changes occur? Typically, this question has been examined through correlations between first test and second; the higher the correlation, the greater the stability. Because IQ tests for children are designed to yield the same mean IQ at each age, it is also possible to look at the constancy of the IQ value itself. IQ, after all, *is* relative standing, and thus constancy of relative standing implies constancy of IQ. We can ask, for example, whether a child with an IQ of 90 at age 4 will still have an IQ of 90 at age 6 or 10 or 20.

IQ has been a popular topic for longitudinal study for close to 80 years now. Such studies can and often do encounter all of the problems of longitudinal research that were discussed in chapter 3. Participants who are both able and willing to be tested repeatedly may not be a representative sample of the population as a whole, a bias that limits the generalizability of the results. Dropout in the course of the study may be selective, and, if so, the sample will become even more biased. Repeated administrations of the same test can result in practice effects, thus inflating later scores relative to early ones. And questions of measurement equivalence may

arise if the study spans distinct age groups that require different IQ tests (e.g., infants and older children).

Several conclusions from the stability studies can be noted. A first set of conclusions concerns predictions from infancy. Except for extremely low scores, performance on traditional infant tests such as the Bayley is generally not predictive of later IQ (Lipsitt, 1992). This finding has long been evident, and it has led to the conclusion that there is a *discontinuity* in the nature of intellectual skills from infancy to later childhood—that is, what we mean by "intelligence" is simply not the same in the sensorimotor, preverbal infant as it is in the older child or adult.

There is undoubtedly some truth to this discontinuity argument. Nevertheless, recent research suggests that there is also one important qualification to it. A number of investigators have found that *response to novelty* in infancy *does* predict later IQ—not perfectly, by any means, but with typical correlations of .35 to .40 (Kaveck, 2004; McCall & Mash, 1994). Thus, babies who are especially interested in and responsive to novelty are the ones who, on the average, grow up to have the highest IQs. This finding, in turn, has led to the creation of a new instrument for assessing infant intelligence that is built around response to novelty: The Fagan Test of Infant Intelligence (Fagan & Detterman, 1992; Fagan & Shepherd, 1986).

What about predictions from childhood IQ tests? Once we move beyond infancy, scores do begin to correlate significantly from one age period to another; the stability, however, is far from perfect. In general, correlations—and thus similarity in IQ—are higher the closer together the ages being compared; they are also higher the older the child is at the time of initial testing. The latter statement is equivalent to saying that there is increased stability of IQ with increasing age.

The question of where differences in IQ come from has been hotly debated ever since the first IQ tests were developed. Part of the reason for the debate lies in the difficulty of getting clear evidence on the question. There are two possible sources for differences in IQ: the different genes with which people are born, or the different environments in which they grow up. It is easy enough to imagine a well-designed scientific study that would disentangle these two factors; all that need be done is to hold one factor constant while systematically varying the other. It is equally easy to see that such studies are impossible to do. The result is that we must fall back upon less satisfactory sorts of evidence. Two kinds of evidence have been prominent in the heredity-environment debate: studies of twins and studies of adopted children.

We considered twin studies briefly in chapter 11 in the discussion of temperament. Twin studies capitalize on the fact that there are two kinds of twins. Monozygotic or identical twins come from the same egg and are thus genetically identical; dizygotic or fraternal twins come from different eggs and thus have only a 50% average genetic overlap—the same as ordinary siblings. We have, then, a naturally occurring experiment with variation in the genetic variable. If genes are important for IQ, identical twins should be more similar in IQ than are fraternal twins. And this, in fact, is the finding. Reported correlations in IQ for identical twin pairs are in the .80s; correlations for fraternal twin pairs are typically in the .50s or .60s (McGue, Bouchard, Iacono, & Lykken, 1993; Plomin, 1990).

There is an obvious criticism of this genetic interpretation of the twin data. Perhaps environments are on the average more similar for identical twins than for fraternal twins. Identical twins, after all, look and in some ways act more alike than fraternal twins, and they may elicit a more similar treatment from their environments. There is an obvious answer to this obvious criticism: Study identical twins who have been separated early in life and brought up in different environments. Such twins are not easy to come by—there are only a handful of such studies, and none has a very large sample size. Furthermore, no one

separates twins for the purposes of scientific study; separations occur for a variety of reasons under a variety of circumstances, and this lack of control hampers clear interpretation. Nevertheless, the data that emerge from such studies appear strongly supportive of a genetic model (Bouchard, 1997; Segal, 1999). Reported correlations in IQ for identical twins reared apart average around .75—only slightly lower than those for nonseparated identical twins, and higher than those for fraternal twins brought up in the same home.

Studies of adopted children constitute a larger literature than studies of twins. The starting point for such studies is the finding that parents and children typically correlate about .5 in IQ. Children tend, therefore, to resemble their parents; the problem is to figure out why. Each parent contributes 50% of the child's genes; thus there is a genetic basis for the correlation. But each parent also contributes a major part of the child's environment; thus there is also an environmental basis.

Studies of adopted children offer the possibility of pulling apart these two contributors. What we can look at are two sets of correlations. One is the correlation between an adopted child's IQ and the IQs of the adoptive parents. In this case the environmental basis remains; the genetic contribution, however, is ruled out. The other is the correlation between an adopted child's IQ and the IQ of the biological parents. In this case the environmental contribution (apart from the prenatal and perhaps early postbirth environment) is ruled out; the genetic basis, however, remains.

Two main findings emerge from the studies of adopted children (Turkheimer, 1991; van IJzendoorn, Juffer, & Poelhuis, 2005). One is that the child's IQ correlates more highly with those of the biological parents than with those of the adoptive parents. This is evidence in support of the importance of genetic factors. The second finding is that the mean IQ for samples of adopted children is typically above average—and typically above the mean for their biological parents. Because correlation is a measure of relative standing, a mean difference of this sort can come about even though the two sets of scores are fairly highly correlated. Part of the parent-child difference can be attributed to regression to the mean, a phenomenon that also applies to cross-generation comparisons (i.e., parents with below-average IQs tend to have children whose IQs are higher than their own). Part of it, however, is almost certainly a reflection of the above-average nature of adoptive homes. Such homes tend to be privileged in various ways, and they apparently boost the IQs of children who grow up in them. Thus, the studies of adopted children provide evidence for both genetic and environmental effects.

It should be noted that the adopted child studies, like the twin studies, do have some limitations. I mention two here. One is the possibility of *selective placement*—that is, the tendency of adoption agencies to do some matching of the adoptive home with characteristics of the biological parents. To the extent that such selective placement occurs, interpretation of the parent-child correlations becomes very difficult. A second possible problem is *restriction of range* among adoptive homes. As the preceding paragraph noted, such homes are not a random subset of the population of homes in general; rather, they tend to be above average in various ways. This also means, however, that they tend to be fairly homogeneous, and such homogeneity is a problem in a correlational study. The lower the variation in a variable (in this case, characteristics of adoptive homes, including adoptive parents' IQs), the less likely it is that that variable will correlate significantly with other variables. This factor sets an important qualification on the low correlation between adopted child and adoptive parents.

Memory

The remaining sections of the chapter shift from general approaches to specific topics in the study

of cognitive development. The first topic is one of venerable standing. Memory has been a focus of study in psychology since the earliest days of psychology as a science, and children's memory has long been of interest to both parents and psychologists. I begin with memory in infancy, then move on to work with older children.

Memory in Infancy

Memory was not one of the topics in chapter 11's discussion of research on infancy. Nevertheless, many of the procedures considered there, whatever their primary focus, necessarily tell us something about infant memory as well. Piaget's work, for example, shows infants remembering things and acting accordingly from very early in life. Similarly, many phenomena in the development of attachment, such as the infant's preference for mother over stranger, clearly imply the operation of memory.

The main procedures for the explicit study of infant memory have come from two other approaches discussed in chapter 11: the habituation-dishabituation paradigm and the Fantz preference method. Memory is an intrinsic part of the habituation-dishabituation procedure. The only way that an infant can habituate to a stimulus is if the infant can store information about the stimulus over time and recognize it as familiar when it is encountered again. If there were no memory for past events, there could be no habituation. Similarly, the only way that an infant can show dishabituation to a new stimulus is if the infant remembers the original stimulus and realizes that the new stimulus is in some way different. Habituation and dishabituation can both be demonstrated in the newborn; thus, we know that some memory capacity is present from birth.

Memory is not an intrinsic part of the preference method, but it can easily be built into the procedure. All that need be done is to make familiarity the critical dimension along which the two stimuli differ. Suppose, for example, that we run a series of trials on which one of the two stimuli is always a triangle; the other stimulus, however, changes from trial to trial. The question is whether, over trials, the infant begins to look longer at the relatively novel or the relatively familiar stimulus. Infants older than about 2 months in fact show a preference for novelty; infants younger than 2 months may prefer familiarity (e.g., Wetherford & Cohen, 1973). The point for now is that either kind of preference—for the novel or for the familiar—implies the operation of memory.

In addition simply to demonstrating the presence of memory, the paradigms just discussed can be used to probe the nature of the early memory system. For example, we can examine the duration of the infant's memory by varying the time period between the initial exposure to the stimulus and the test for whether the stimulus is remembered. Or we can explore which aspects of a stimulus are noted and retained by varying which aspects we change during the dishabituation phase. If, for example, a stimulus consists of components A, B, and C, we can test for dishabituation when A alone is changed, when B alone is changed, or when C alone changes. Studies of this sort have revealed both some impressive early competencies in young infants and some important advances in the memory system across the first year or so of life (Courage & Howe, 2004; Rose, Feldman, & Jankowski, 2004).

Other procedures have also contributed to the study of early memory. Consider the arrangement pictured in Figure 12.2. As the figure shows, the ribbon connecting ankle and mobile confers a potential power upon the infant: If the baby kicks, the mobile will jump. Infants as young as 2 months can learn this relation (a form of operant conditioning). Once the response has been established, we can alter the situation in various ways to probe the infant's memory. We can test the duration of memory by reintroducing the mobile after a delay and seeing whether the kick response still occurs. We can test the specificity of memory by presenting new mobiles that vary

Figure 12.2 An experimental arrangement for studying infants' ability to learn and remember. When the ribbon is attached to the baby's ankle (as in the right-hand photo), kicking the leg makes the mobile above the crib move. Learning is shown by increased kicking whenever the ribbon is attached and the mobile is present.

SOURCE: These photos were made available by Dr. C. K. Rovee-Collier.

in their similarity to the training mobile. These studies, too, reveal both some impressive early abilities and improvements in those abilities across infancy (Rovee-Collier, 1999).

The kind of memory that is most clearly demonstrated in all of the examples discussed thus far is what is called **recognition memory**. Psychology's definition of "recognition" is the same as the everyday, dictionary definition: realizing that something new is the same as something encountered before. An infant who habituates is thus showing recognition, as is an infant who responds to the familiarity of the mother's face. Recognition can be contrasted with another basic form of memory: **recall memory**. Recall refers to the active retrieval of some memory material that is *not* immediately present. A child who relates what happened at his or her birthday party of a week before is demonstrating recall; so too is a child who draws a picture of the party.

Recall is difficult to study in infants, because babies cannot produce the responses (such as verbal reports or drawings) that are used to study recall in older participants. Investigators have therefore had to rely on less direct measures. Probably the most informative of these measures has been **deferred imitation**—that is, imitation of a model that is not immediately present but rather occurred some time in the

past. The ability to reproduce behaviors that were seen a day or a week earlier would certainly seem to imply some capacity to recall information over time.

Piaget was the first to study deferred imitation, and he concluded that such imitation emerged at about 18 to 24 months, a finding in keeping with his belief that deferred imitation requires a capacity for representation and recall that is not available until the end of infancy. Later researchers have accepted the theoretical argument but disagreed with respect to the timing. A study by Meltzoff (1988) provided the first challenge to Piaget's account; it showed that 9-month-olds could imitate a novel action (e.g., pushing a button to produce a sound) that they had witnessed 24 hours earlier. Subsequent studies pushed the earliest signs of deferred imitation steadily earlier in development; at present 6 weeks is the youngest age at which success has been demonstrated (Meltzoff & Moore, 1994). Thus, some capacity for recall, if not present at birth, seems to emerge very early.

Infants can imitate and remember not just isolated behaviors but also simple sequences of actions. Eleven-month-olds, for example, can reproduce two-act sequences modeled by an adult (e.g., to make a rattle, first put button in box, then shake box); by 13 months babies can handle sequences with three components (e.g., to make a more complicated rattle, first put ball in larger cup, then invert smaller cup into larger, then shake cups—Bauer & Mandler, 1992). Such memories, moreover, are not necessarily only short-term; 11-month-olds show some retention of modeled sequences across delays as great as 3 months, and by 16 months of age the retention interval has stretched to 6 months (Bauer, 2004). There is even longitudinal evidence that some babies tested first at 11 months of age can remember some aspects of their experience a full year later (McDonough & Mandler, 1994)!

As soon as we move beyond infancy, studies of recall become much more common than studies of recognition. It has been clear for a long time that older children, on the average, recall things better than younger children. This finding emerges in a variety of contexts and for a variety of forms of recall, including more advanced versions of the sort of sequential memory that we just saw is first evident in infancy. The really interesting question is why such improvements in memory occur. I turn next to two kinds of research that have attempted to identify the bases for developmental improvements in memory.

Mnemonic Strategies

The basic idea behind the study of **mnemonic strategies** is that developmental improvements in recall do not result solely—or perhaps at all—from a simple quantitative expansion in the size of the memory "store." The improvements, rather, reflect the greater tendency of older children to *do* something—to utilize some mnemonic strategy—to help themselves remember. Such strategies may come at the time of initial exposure to the material, or during the delay period between exposure and memory test, or at the point of attempting to retrieve the material. These strategies may take a variety of forms: some verbal, some nonverbal, some simple, some complex. Their common property is that they do, at least usually, facilitate memory.

Let us briefly consider what mnemonic strategies look like, before moving on to the question of how to study them. Table 12.2 shows examples of three general classes of mnemonic strategies. The strategies in the table are by no means the only ones that children might use; indeed, we saw an example of another memory strategy—appropriate allocation of study time—in the Dufresne and Kobasigawa (1989) study discussed in chapter 2. Rehearsal, organization, and elaboration are, however, among the strategies that have received the most research attention.

Table 12.2 Examples of Mnemonic Strategies

General strategy	*Experimental task*	*Children's procedure*
Verbal rehearsal	Ten pictures of familiar but unrelated objects are presented one at a time and then removed. The child's task is to recall as many of the pictures as possible.	Single-item rehearsal: Label each picture repeatedly as it is presented—e.g., "apple, apple, apple, flag, flag, flag. . . ." Cumulative rehearsal: Label each picture as it is presented and then rehearse all of the labels to that point—e.g., "apple, apple-flag, apple-flag-moon. . . ."
Elaboration	Twenty pairs of familiar but unrelated words are presented (e.g., "cow-tie," "car-tree"). Subsequently one member of each pair is presented; the child's task is to recall the other member of the pair.	Pictorial elaboration: Create a mental image that relates the members of a pair in some way—e.g., picture to oneself a cow wearing a tie, a car driving up a tree. . . . Verbal elaboration: Create a sentence or phrase that relates the members of a pair in some way—e.g., "the cow wore a tie," "the car drove up the tree. . . ."
Organization	Twenty pictures of familiar objects are presented, drawn from four conceptual categories: animals, vehicles, clothing, furniture. The pictures are not grouped by category during presentation. The child's task is to recall as many of the pictures as possible.	Clustering: Organize the pictures in terms of the four categories, and recall members from the same category together—e.g., "dog, horse, camel, bear, squirrel, car, truck, plane. . . ."

How can strategies be studied? Some strategies are by their nature overt and thus relatively easy to study. Note taking, for example, is a common and easily observable mnemonic strategy. So is asking one's parents for help in remembering something. Perhaps because such external strategies seem so obvious, however, they have not received much attention from developmental psychologists. The interest, rather, has been in more internal, in-the-head strategies such as those in the table—in the kinds of strategies that are necessary precisely when external aids like notes or parents are *not* available. And here, clearly, we run into a measurement problem. How can an in-the-head strategy be measured?

One possibility is to *infer* the use of a strategy from the participant's overt memory performance. Suppose, for example, that we have a memory task that is heavily verbal, that we study verbally mature older children and verbally immature younger children, and that we

find that the older children perform better. A reasonable (although of course not certain) inference is that the older children perform better because they are using their verbal skills to help themselves remember. Perhaps, for example, they verbally rehearse the items during the delay period, whereas the younger children do not. Such an inference becomes more certain if we can identify developmental changes not only in the level but in the *pattern* of performance. For example, a classic finding from the adult memory literature concerns the so-called *primacy effect,* which is the tendency for the first items in a stimulus list to be remembered better than later items are remembered. This effect is typically attributed to verbal rehearsal of the early items. When comparable studies are done developmentally, the finding is that older children show the primacy effect; younger children, however, usually do not (e.g., Cole, Frankel, & Sharp, 1971). Such a pattern is consistent with the hypothesis of a developmental increase in verbal rehearsal.

A second general approach is to *induce* the use of the strategy. In this case we do not guess at the presence of the strategy; rather, we instruct children in its use and then observe the effects. We might, for example, tell half of our participants to rehearse the items during the delay period, but give no such instructions to the other, control half. A typical finding is that such instruction is beneficial; children who are helped to use a strategy generally perform better than those who are not. Note that we have here the kind of convergence of evidence that has been discussed at various points. Studies of inferred strategy use suggest that some helpful mnemonic strategy is being used; studies of induced strategy use confirm that the strategy is in fact helpful.

Although the kinds of studies just discussed are informative, they do have their limitations. Studies that infer strategy use have the obvious limitation that the use is inferred; we do not know for certain what the child is doing. Studies that induce strategy use avoid this problem but run into another: Because we have forced the children to use a strategy, we do not know what they would have done on their own.

A breakthrough in the study of children's memory occurred in the 1960s with the discovery of experimental situations in which children would spontaneously produce strategies that were at least somewhat overt and therefore measurable. This discovery made it possible to combine the best elements of the other two approaches: to study strategies that were both spontaneous and observable. Because John Flavell was a pioneering researcher in this area, I will use one of his early studies as an example. It should be stressed, however, that this example is just one of many; both the Flavell research group (e.g., Flavell, Friedrichs, & Hoyt, 1970) and other investigators (e.g., Bjorklund, Coyle, & Gaultney, 1992; Miller, 1990) have been ingenious in devising experimental settings in which strategies can be directly observed.

Flavell, Beach, and Chinsky (1966) presented pictures of seven common objects to kindergarten, second-, and fifth-grade children. On each trial the experimenter pointed in a particular order to a subset of the pictures; the child's task was to recall the designated pictures in the correct order. A delay of 15 seconds intervened between pointing and recall test. The child wore a toy space helmet throughout the study, and during the delay period the visor of the helmet was pulled down, thus ensuring that the child could not see the pictures. The visor had a second purpose as well, which was to allow one of the experimenters to stare at the child's mouth during the delay. This experimenter's job was to record any verbalizations by the child, both overt verbalizations and semiovert ones (the experimenter had been trained to lip-read prior to the study). Of particular interest, of course, were instances of apparent rehearsal.

Children did rehearse, but the probability of rehearsal was strongly tied to age: 17 of 20 fifth-graders showed detectable rehearsal, whereas only 2 of 20 kindergarteners did so.

A subsequent study with a similar procedure (Keeney, Cannizzo, & Flavell, 1967) demonstrated that individual differences in rehearsal within an age group (first grade) correlated with recall performance—that is, children who spontaneously rehearsed showed better recall than children who did not. Both the developmental difference in the tendency to use strategies and the benefit when strategies *are* used are common findings in the memory literature.

In studies of the sort just described, children younger than 5 or 6 often fail to generate mnemonic strategies. Lest it be thought that this is an absolute deficit, consider a study by Wellman, Ritter, and Flavell (1975). Their participants were 3-year-old children, and their experimental procedure is summarized in the following passage:

> I want to tell you a story about this dog. See, here he is on the playground [table top]. He loves to play, he runs, he jumps, . . . but he was playing so hard he got very hungry. So he went to look for some food. When he was looking he went by this dog house, and this dog house, and this dog house, and this dog house [dog is walked by all four cups]. And then he went in this dog house to find some food [dog is hidden]. You know what, I have another toy I could get to help us tell the story. I'll go get it because we need it for the story. (p. 781)

At this point instructions diverged for the two experimental conditions. Children in the Wait condition were told simply to wait with the dog. Children in the Remember condition were told to remember where the dog was. The question, of course, was whether the children who were told to remember would behave differently from those who were told simply to wait. The answer is that they did behave differently, and in a very sensible way: Children instructed to remember were much more likely to spend the delay period with their eyes glued to the critical cup and possibly their fingers touching it as well. These are, it is true, very simple mnemonic strategies, but they *are* strategies, and they are available by age 3. Other studies using a similar hide-and-seek procedure have found the rudiments of strategic behavior in children as young as 18 months (DeLoache, Cassidy, & Brown, 1985). We can see here another example of what has become a recurrent theme in developmental psychology: It is dangerous ever to assert that some ability (in this case, mnemonic strategies) is totally lacking in the young child.

Constructive Memory

Important though they are, strategies do not account for all memory phenomena of interest or for every important developmental change in memory. Consider some of the limitations of the studies of strategic memory. Such studies have typically focused on memory for arbitrary and meaningless material (e.g., lists of unrelated words); clearly, however, much of real-life memory is for meaningful material. Studies of strategic memory have focused on intentional memory; much (perhaps most) of real-life memory, however, is unintentional or incidental, in the sense that we remember things that we never attempted to commit to memory. Finally, studies of strategic memory concern definite and discrete techniques for storing or retrieving the memory material. But not all of the cognitive activities involved in memory can be accounted for in terms of such definite and intentional strategies.

The study of **constructive memory** concerns the effects of the general knowledge system on memory. The basic idea behind this approach is that memory is simply a form of applied cognition, that form that has to do with storing information over time and retrieving information from the past. Like any form of cognition, memory involves action and understanding, not merely passive registration of input. And like any form of cognition, memory shows definite developmental changes as the child's understanding of the world changes.

Let us consider an example. Paris (1975) read the following story to kindergarten through fifth-grade children:

> Linda was playing with her new doll in front of her big red house. Suddenly she heard a strange sound coming from under the porch. It was the flapping of wings. Linda wanted to help so much, but she did not know what to do. She ran inside the house and grabbed a shoe box from the closet. Then Linda looked inside her desk until she found eight sheets of yellow paper. She cut up the paper into little pieces and put them in the bottom of the box. Linda gently picked up the helpless creature and took it with her. Her teacher knew what to do. (p. 233)

Following the story, children were asked eight questions:

1. Was Linda's doll new?
2. Did Linda grab a match box?
3. Was the strange sound coming from under the porch?
4. Was Linda playing behind her house?
5. Did Linda like to take care of animals?
6. Did Linda take what she found to the police station?
7. Did Linda find a frog?
8. Did Linda use a pair of scissors? (p. 233)

Note the basic difference between the first four questions and the last four. The first four concern specific information that is given directly in the story. The last four, however, can be answered only on the basis of inferences that go beyond the information that is explicitly provided. A basic finding—not only from the Paris study but from other such studies as well—is that even young children do make the kinds of inferences that are required by the second set of questions. The tendency to make inferences, as well as the complexity of the inferences that are possible, increases with development. But from early in life, memory—whether for a story, a conversation, an event seen, or whatever—is for meaning and not simply for verbatim detail.

Let us consider another example of this point. We saw earlier that even infants demonstrate some memory for the sequence in which events typically occur. So, of course, do older children—and for events that are a good deal lengthier and more complex than the simple sequences that fall within the span of infant memory. Memories that have to do with the order and the structure of familiar events are referred to as **scripts**. Children form a number of scripts as they develop; among the scripts that have been studied are those for taking a bath, making cookies, going to McDonald's, and having a birthday party.

Scripts are themselves a form of memory; once developed, however, they also influence subsequent memory. Children (and indeed any of us) often make sense of new experiences by relating them to familiar scripts. In general, children remember experiences that preserve the structure of a familiar script better than those that violate a script, and they show better memory for events that are central to a script than for those that are peripheral (Mandler, 1983; McCartney & Nelson, 1981). Children may even rearrange details in their recall to make their experiences fit a script. In one study, for example, preschoolers who heard the phrase "children brought presents" at the end of a story about a birthday party tended to move the phrase to the beginning of the story in their retelling (Hudson & Nelson, 1983). We will encounter such constructive "corrections" again in chapter 13, in the context of a consideration of how children's beliefs about sex differences influence their memory for how boys and girls or men and women behave.

As a final example of the constructive nature of memory, consider the stimuli in Figure 12.3.

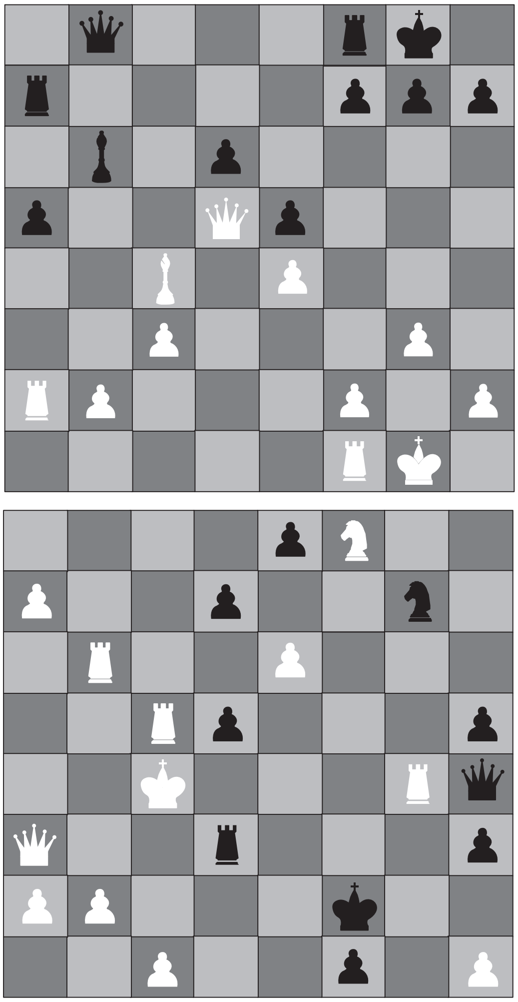


Figure 12.3 Meaningful and random configurations of chess pieces. Because they can make use of their knowledge of chess, chess players show better memory for the top array than for the bottom one.

SOURCE: From "Chess Expertise and Memory for Chess Positions in Children and Adults," by W. Schneider, H. Gruber, A. Gold, and K. Opwis, 1993, *Journal of Experimental Child Psychology, 56*, p. 335. Copyright © 1993 by Academic Press. Reprinted by permission from Elsevier.

Imagine that your task is to study each picture for 10 seconds and then attempt to reproduce the configuration from memory. If you do not play chess, the two arrays would probably prove equally difficult; both, after all, contain the same number and the same variety of stimuli. If you do play chess, however, then the top array would almost certainly be easier to remember. It would be easier because the pieces are in positions that could actually occur in a game, whereas those in the bottom array are randomly arranged. Furthermore, if you not only play chess but play it well, then your memory for the top array—as well as the difference in memory for the two arrays—would almost certainly be even greater.

The chess example illustrates the effect of expertise on memory. The term **expertise** refers to organized factual knowledge about some content domain—to what we know about some subject. In general, as in the chess example, when expertise is high memory is high—experts take in and access information more quickly and more effectively than do nonexperts. Because older children and adults possess more expertise for most topics than do young children, this phenomenon provides another explanation for improvements in memory with age. On the other hand, the concept of expertise also leads to an interesting prediction: If we can find situations in which children have more expertise than adults, then the usual developmental differences in memory might be reversed. And this in fact has proved to be the case for a variety of topics, including the memory-for-chess example with which I introduced the concept of expertise (Chi, 1978). Such a finding is instructive, because it indicates that in at least some instances it is expertise, and not other factors associated with age, that is important for memory.

This section has just touched on the many ways in which children's knowledge can affect what they remember. Fuller coverage of such effects—as well as discussion of other bases for developmental changes in memory—can be found in Flavell, Miller, and Miller (2002) and Siegler and Alibali (2005).

Theory of Mind

Methods of Study

Consider the simple scenario depicted in Figure 12.4. To any adult, the answer to the question of where Sally will search for her marble is obvious—in the basket, where she last saw it. She has no way, after all, of knowing that the marble has been moved in her absence. It turns out, however, that this answer is not obvious to many preschool children. Most 3-year-olds indicate that Sally will look in the box. They answer in terms of their own knowledge of the reality of the situation, and not in terms of what Sally could be expected to think.

The Sally/Anne task is an example of research from an exciting new area of study labeled **theory of mind**. Theory of mind has to do with understanding of the mental world—with how children think about phenomena such as desires, intentions, emotions, and (as in the example) beliefs. The interest in such questions is, to be sure, not totally new; the Piagetian approach, in particular, has long encompassed social and mental phenomena within its targets of study. The work on perspective taking discussed earlier in the chapter is in fact a clear forerunner of the contemporary interest in theory of mind. The modern emphasis, however, has gone well beyond these Piagetian beginnings, spawning a host of theories, issues, and research findings that were not available before. It has also led to the creation of a number of new methodologies for studying children's understanding of mental phenomena, and it is the methodological aspect of theory of mind on which I concentrate here.

One such methodology is the one shown in Figure 12.4. The concept at issue in the Sally/Anne task is **false belief**—the realization that it is possible for people to believe something that is not true. Correct response to the task clearly requires this realization; children must set aside their own knowledge of the true location of the marble to realize that Sally would believe something else—would hold a false belief. As already noted, most 3-year-olds fail this task, whereas by age 4 or 5 most children succeed. The false belief task has been of interest because it tells us something very basic about the child's understanding of mental representation. To appreciate the possibility of a false belief requires the understanding that beliefs are simply mental representations, not direct reflections of the world, and as representations they may or may not be true. Thus, what I believe may or may not be the same as what you believe, and what either of us believes may or may not agree with reality. It is this understanding that the young child seems to lack.

There have been two main paradigms for studying false belief. The approach shown in Figure 12.4 is labeled the *unexpected locations* (or *unexpected transfer*) *task* (a procedure invented by Wimmer and Perner, 1983). The other possibility is the *unexpected contents task* (first used by Hogrefe, Wimmer, and Perner, 1986). Suppose that we show the child a candy box and ask what the child thinks is inside. The child replies "candy," at which point we open the box and show that it actually contains pencils. We then close the box back up, bring out a puppet of Ernie from Sesame Street, and pose the following question: "Ernie hasn't seen the inside of the box. What will Ernie think is inside the box before I open it?" You or I (and most 4- and 5-year-olds) would say "candy"—we would realize that Ernie has only the outside of the box as evidence and thus would form a false belief about its contents. Most 3-year-olds say "pencils." As in the locations task, young children are unable to set aside their own knowledge of reality; rather, they assume that everyone else must hold the same true belief that they hold.

The contents task lends itself to another measure as well. Rather than making Ernie the target, we can pose the question in terms of the child's own initial belief. The issue now is whether children realize that they themselves can hold false beliefs. The first part of the procedure remains the same: presentation of the

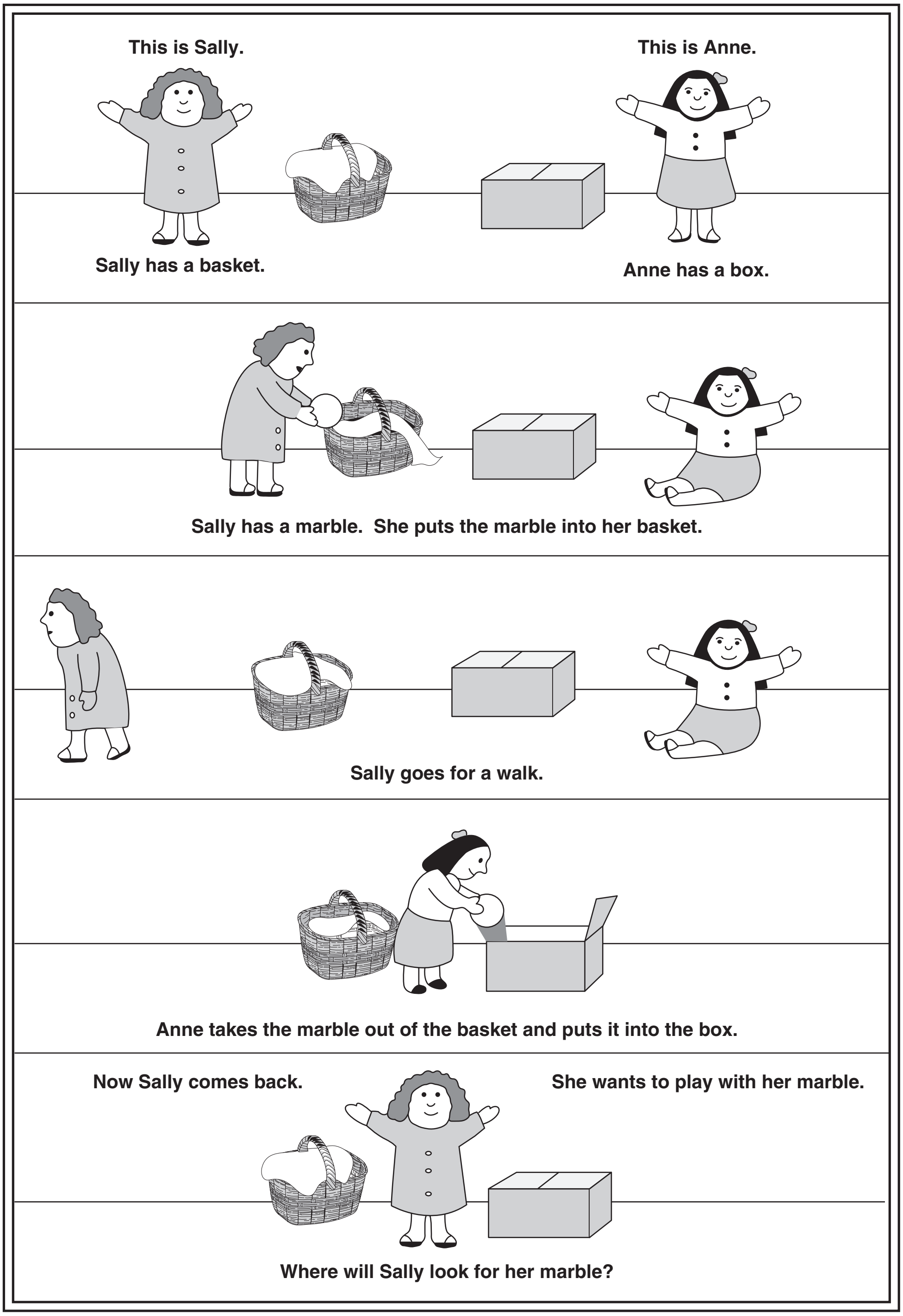


Figure 12.4 Example of a false belief task. To answer correctly, the child must realize that beliefs are mental representations that may differ from reality.

SOURCE: From *Autism: Explaining the Enigma* (p. 160) by U. Frith, 1989, Oxford: Basil Blackwell. Copyright 1989 by Axel Sheffler. Reprinted by permission.

candy box, elicitation of an initial judgment about the contents, revelation of the true contents. Now, however, the test question becomes the following: "What did you think was inside the box before I opened it?" We might expect this question to be easy; after all, all the child need do is repeat the response from a few seconds before. Three-year-olds, however, are likely to say "pencils." Understanding one's own false belief seems to be just as difficult as understanding false belief in another. The understanding that one's own mental states can change is referred to as **representational change**.

Another of the basic procedures under the theory-of-mind heading is the **appearance-reality task**. This task, too, is directed to an important form of knowledge: the realization that things can appear different from what they really are. You realize, for example, that a straw placed in a glass of water does not really bend upon entering the water—it merely looks bent. Similarly, you understand that a white cutout butterfly placed under a red filter is not really red—it just looks red.

Young children do not at first possess this understanding. The filter task is in fact one common measure, and it generally proceeds as follows (e.g., Flavell, Flavell, & Green, 1983). After some pretraining in the language to be used, the experimenter first presents the butterfly in its nonillusory form, next places the filter over it, and then asks the following two questions: "When you look at the butterfly, does it *look* red or does it *look* white?" "For *real,* is the butterfly *really and truly* white or *really and truly* red?" Correct response requires distinguishing the appearance and the reality, and this is what young children cannot do; most 3-year-olds answer with the appearance in both cases, thus claiming that the butterfly is really, and not just apparently, red. Judging in terms of appearance to the neglect of reality is referred to as a *phenomenalism error.* In other versions of the appearance-reality task, children make the opposite, *realism error:* judging in terms of the reality, to the neglect of the appearance. An often used stimulus in this case is the sponge/rock: a foam rubber sponge that looks for all the world like a gray rock. Once children have had a chance to touch the sponge and learn its true nature, they are asked the usual two questions: What does it look like, and what is it really? Here, remarkably, 3-year-olds tend to claim not only that the object really is a sponge but also that it looks like a sponge. Once they are aware of the reality, they can no longer report the discrepant appearance.

Unlike false belief and appearance-reality, the next methodology to be discussed does not have a single, agreed-upon label. Many investigators have probed what children understand about the *origins of knowledge*—that is, their conceptions of how we form our beliefs about the world (e.g., Miller, Hardin, & Montgomery, 2003; O'Neill, Astington, & Flavell, 1992). Do children appreciate, for example, the central role of perception in belief formation? Can they judge that someone with perceptual access to a stimulus will acquire information about it, whereas someone who lacks access will not? Can they distinguish among different perceptual modalities, judging what it is that can be learned (or that they themselves have learned) through sight as opposed to hearing as opposed to touch? Do they realize that there are sources of knowledge other than perception—logical inference, for example, or communication from others? And can they make all of these judgments not just with respect to themselves but with respect to other people?

As this run-through of (some of) the many possibilities suggests, there are too many different issues here to attempt a summary of findings. I will, however, note one general conclusion from the work on origins of knowledge, a conclusion that parallels findings from the studies of false belief and appearance-reality. The conclusion is that young preschoolers have only a very limited understanding of where beliefs come from, and they often err on tasks that seem transparently obvious to an adult.

They may claim, for example, that they have always known a fact that the experimenter has just taught them (Taylor, Esbensen, & Bennett, 1994), and they may be unable, seconds after learning something, to indicate whether sight, touching, or hearing was the source of the information (O'Neill & Gopnik, 1991). We can see again that there is much to develop with respect to theory-of-mind understanding.

This sampling of tasks hardly exhausts the topics studied under the heading of theory of mind. Researchers have probed, for example, for understanding of various psychological states in addition to thoughts and beliefs—for what children know about emotions (e.g., Lagutta, 2005) or desires (e.g., Moses, Coon, & Wusinich, 2000) or fantasies (e.g., Sharon & Woolley, 2004) or intentions (e.g., Baird & Moses, 2001). They have also explored theory of mind in age periods other than the preschool years, searching for both precursors in infancy (e.g., Legerstee, 2006) and more advanced developments in later childhood (e.g., Bosacki & Astington, 1999). Further discussions of tasks and related findings are available in a number of sources, including Astington (2000), Harris (2006), and Wellman (2002).

Research Issues

Just as my coverage of tasks was selective, so must be the coverage of research issues. I will settle for discussing two questions that have been the focus of much research attention.

The first issue is the same fundamental question addressed with respect to Piaget's studies: the question of assessment. In a sense, the study of false belief has been interesting for the same reason that much of Piagetian research is interesting: It provides surprising, hitherto unsuspected information about young children's cognitive limitations. Just as it is remarkable that an infant should lack object permanence or a preschooler conservation, so is it surprising that a 3-year-old should have no conception of the possibility that a belief can be false. As with Piagetian research, a natural and appropriate response to such a counterintuitive claim is to wonder about the accuracy of the assessment. Do the standard tasks really provide an accurate measure of what young children understand about belief? It is possible that they do not, and for the same general reasons that Piaget's measures have been criticized: The tasks involve language that may confuse a young child, the context for the assessment is strange and at least somewhat unnatural, and the procedures may not optimally engage the child's interest and motivation. Chandler and Hala (1994, p. 412) summarize the criticisms as follows: "These procedures have typically turned upon matters that are often static, hypothetical, third-party, and of no immediate relevance or personal interest to the young subjects in question." Perhaps 3-year-olds would appear more competent if the assessment could somehow be made more "child-friendly."

As with the conservation task, researchers have tried in various ways to probe for earlier competence. Modifications in wording have again been explored. In the locations paradigm, for example, it is possible that young children interpret the standard question as asking where the protagonist *should* search to find the object, or perhaps as asking where the protagonist will *eventually* search to find it; in either case the children would "err" by picking the true location. To test this possibility, Siegal and Beattie (1991) added a potentially clarifying "first" to the question ("Where will Jane look first for her kitten?"). Other researchers have varied the kind of response that is required of the child. It has been suggested that predicting an action based on a false belief (the usual measure) may be more difficult for young children than explaining a false-belief-based action that has already occurred. Several investigators (e.g., Bartsch & Wellman, 1989; Moses & Flavell, 1990), therefore, have contrasted the standard prediction measures with explanation measures, in which the child is asked to interpret a

story character's mistaken action (e.g., Jane searching in the wrong place for her kitten). A study by Clements and Perner (1994) reduces the response demands on the child still further. These investigators measured children's looking behavior on a locations task, the attempt being to see whether children might look first at the original, false belief location, even if they were unable to answer the standard verbal question (as indeed many young children did).

Probably the most common procedural modification has been to incorporate *deception* into the false belief assessment. The arguments for doing so are both theoretical and methodological. Theoretically, a full understanding of deception implies an understanding of false belief as well, since the purpose of deception is to implant a false belief in someone else. Methodologically, an assessment based on deception can build upon children's natural experiences with tricks and games, thus adding a degree of familiarity and personal involvement that is lacking with the standard measures. Researchers have explored children's understanding of deception in various ways. Some have added deception to the standard paradigms by enlisting the child's help in "playing a trick" on someone else; the child is then the one to switch the contents of the candy box or move the toy to a new location (e.g., Sullivan & Winner, 1993). Others have devised games in which the child has a chance to deceive a competitor to win a prize—for example, by pointing to the false location for a desired chocolate (e.g., Russell, Mauthner, Sharpe, & Tidswell, 1991). Here, clearly, the motivation to think clearly about the beliefs of another should be maximal.

What have all these modified-assessment studies shown? In answering this question, I will draw from a meta-analysis of false belief studies by Wellman, Cross, and Watson (2001). As with meta-analyses in general (recall the discussion in chapter 8), the Wellman et al. meta-analysis combines and statistically analyzes the results from dozens of different studies. Various "moderator variables," or possible contributors to false belief performance, are explored, including the paradigm used (unexpected locations or unexpected contents), the target for the question (self or other), the wording of the question, and the use of a deceptive context. Variations across studies and relatively early success are sometimes found; in particular, young children sometimes perform better when deception is made part of the assessment. The main conclusion, however, is one of similarity in development across a range of procedures and a range of samples. Wherever and however they are studied, 3-year-olds find it difficult to realize that a belief can be false. Failure to understand false belief, like failure to understand conservation, is a genuine cognitive phenomenon.

The second line of research concerns possible links between theory-of-mind knowledge and social experience. It is certainly plausible that there could be a link, and in both directions—that children's understanding of the mental world can help them interact effectively with other persons, but also that interactions with others can teach children about beliefs and desires and other mental states. This research thus speaks to both the implications and the origins of mental state understanding.

The main approach to the question has been a correlational one. A number of studies have measured both theory-of-mind development and aspects of social experience in preschool samples (e.g., Astington & Jenkins, 1995; Watson, Nixon, Wilson, & Capage, 1999). In general, the relation between the two sets of measures is a positive one, that is, children who are more advanced in theory-of-mind understanding are also more advanced in their social behavior. Astington and Jenkins found, for example, that performance on a battery of false belief tasks was positively related to the level of pretend play with peers. Watson et al. reported a positive relation between false belief understanding and teacher ratings of preschoolers' social skills.

Of course the existence of a correlation between two measures does not tell us the

direction of cause and effect. As we saw in chapter 3, one way to make causal inferences more certain is to trace relations over time through longitudinal study. A number of investigators have done so, and the overall results from this research suggest that the cause and effect can flow in both directions (Astington, 2003). In some instances developments in theory of mind appear to pave the way for later advances in social behavior. Jenkins and Astington (2000), for example, found that false belief performance at the initial measurement occasion predicted later competence in pretend play; the reverse relation, however, did not hold. In other instances it is social experience that appears to play the causal role. A program of research by Judy Dunn and associates, for example, has shown that aspects of early family interaction are related to later theory-of-mind understanding. In one study, mothers' talk about feelings when their children were 2 was related to understanding of emotions and beliefs 7 months later (Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991). In another study, a relatively high degree of family talk about feelings at age 3 predicted children's emotional understanding when they were 6 (Dunn, Brown, & Beardsall, 1991).

The correlational studies are not the only evidence for links between theory of mind and social behavior—or the only evidence that the causal relations between the two domains go in both directions. I mention two other forms of evidence here; Astington (2003) and Hughes and Leekam (2004) provide a fuller discussion.

One kind of evidence concerns the effects of family size. On the average, children with a relatively large number of siblings, perhaps especially older siblings, are faster to master false belief than are children in general (Peterson, 2000; Ruffman, Perner, Naito, Parkin, & Clements, 1998). Children who grow up in large, extended families (Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996) also show faster than average development. Presumably, the heightened social experience with a variety of social agents gives children a helpful basis for learning about the mental states of others. Whatever the specific explanation, the causal direction for the relation must be a social to cognitive one. It is not plausible that being advanced in theory of mind causes the child to have more siblings.

In contrast, the second sort of evidence indicates a cognitive to social direction. It comes from the phenomenon of *autism.* Autism is a severe disorder, almost certainly biological in origin, that is characterized by a number of abnormalities in development, prominent among which are difficulties in social interaction. From early in life, children with autism show little interest in other people and little ability to form interpersonal relationships. They also show marked deficiencies with regard to theory-of-mind understanding (Baron-Cohen, 1995; Frith, 1989). Even when other aspects of mental functioning are relatively unimpaired, children with autism typically perform very poorly on theory-of-mind tasks. As Baron-Cohen (1995) observed, these children's insensitivity to the thoughts, wishes, and feelings of others is not surprising; they may literally not know that such psychological states exist. The consequences of such "mindblindness" (Baron-Cohen, 1995) are clear testimony to the importance of theory of mind for normal social relations.

Concepts

Infancy

Let us revisit for a moment the work on infant memory. We saw that infants demonstrate recognition memory from early in life, as shown by their differential response to familiar and novel stimuli. Even babies a day or two old, in fact, can recognize stimuli that they have encountered before. Impressive as this ability is, by about 3 months of age babies begin to do something that is more impressive still.

FOCUS ON

Box 12.4. Executive Function

Much research in cognitive development is directed to what children know about particular topics or content areas. This is true for the work under the heading of theory of mind that was just discussed. It is also true for the work on concepts to which we turn next.

One fairly common finding from such research is that what children know or how they reason about one content domain (such as theory of mind) is only weakly related to their knowledge or reasoning about another content domain (such as biological concepts). Thus, development often seems to be *domain-specific*—that is, specific to a particular content area, with only limited generality across areas. Such a conclusion is important, for it stands in contrast to the *domain-general* models that long dominated the field—most notably, Piaget's theory of general stages of development.

Although domain specificity is a common message of recent research, many psychologists are reluctant to abandon the search for general processes that cut across content areas. In recent years a body of research and theorizing has developed that attempts to identify certain basic capacities that may underlie performance on a range of different tasks. The term **executive function** has been coined to refer to such general problem-solving resources. In the words of one summary, executive function "is an umbrella term for all of the complex set of cognitive processes that underlie flexible goal-directed responses to novel or difficult situations" (Hughes & Graham, 2002, p. 131).

What sorts of processes fall under this heading? Let us consider an example that will introduce some of the most often studied ones in early childhood. Figure 12.5 shows the stimuli for the *dimensional card sort task* (Zelazo & Frye, 1998). The child is given a series of cards that vary in both shape and color and is instructed to sort them by one of the two dimensions. If the task is the "color game," for example, then the instructions will be to put the red ones in the box with the red picture and the blue ones in the box with the blue picture. After several such trials the rule changes: Now the task becomes the "shape game," and the instructions are to put the cars in one box and the flowers in the other. Simple though this task seems, it is beyond the capacity of most 3-year-olds. Even when they receive the new instructions at the start of every trial, and even when they themselves succeed in verbalizing the new rule, 3-year-olds continue to sort according to the original rule.

Why should such a seemingly simple task prove so difficult? Success on the card sort task in fact requires a number of component processes, processes that turn out not to be well developed at age 3. The child must possess sufficient *short-term memory* to keep the relevant rule in mind while performing the task. The child must be able to *inhibit* a dominant response in order to make the new response when the rule changes. And the child must possess sufficient *self-awareness* to reflect upon and choose among the rules that have been learned.

As noted, the processes just identified are commonly studied ones in executive function research with children. The card sort task is just one of a variety of assessment procedures that have been developed. Inhibition, for example, has been assessed through a version of a Simon Says game in which the child must respond to commands from "Nice Bear" but refrain from responding when "Nasty Dragon" is the speaker (Carlson & Moses, 2001). Short-term memory has been assessed through tasks of backward digit span or backward word span. Hughes and Graham (2002) provide an overview of measures, including description of several standardized assessment batteries.

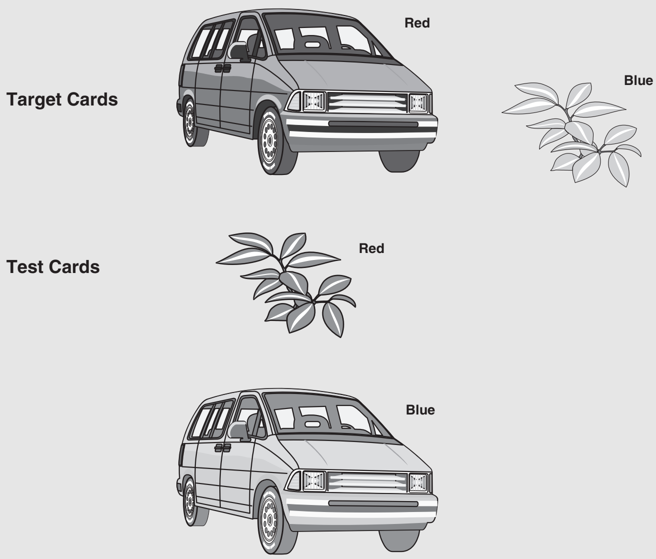


Figure 12.5 Dimensional card sort task used in Zelazo and Frye's study of early rule following.

SOURCE: From "Cognitive Complexity and Control: II. The Development of Executive Function in Childhood," by P. D. Zelazo and D. Frye, 1998, *Current Directions in Psychological Science*, 7, 122. Copyright © 1998 by Blackwell Publishing. Reprinted with permission.

What has research on executive function in childhood shown? Two general sets of conclusions have emerged (Hughes, 2002; Hughes, Graham, & Grayson, 2004). The first concerns executive function processes themselves. As the findings from the card sort task suggest, capacities such as short-term memory and inhibition are far from perfectly developed in early childhood. All of the executive function processes that have been studied in childhood improve as children grow older. In addition to developmental differences, there are individual differences—that is, children at any given age vary in capacities such as short-term memory or inhibition. At the extreme, these individual differences extend to the problems that are evident in certain clinical syndromes; in particular, both Attention Deficit Hyperactivity Disorder (ADHD) and autism are characterized by marked deficits in executive function (Hughes, 2002). Indeed, one reason for interest in executive function—in adulthood as well as childhood—has always been the clinical implications of such work.

The second set of conclusions concerns relations of executive function to other cognitive measures. This issue has been examined most fully with respect to performance on theory-of-mind tasks. A large research literature makes clear that there *is* a relation: Children who are more advanced in various areas of executive function are also more advanced in theory-of-mind understanding. This, in fact, is a general finding across a range of cognitive outcomes. Exactly why such relations occur is a matter of dispute among researchers (Moses, 2001; Perner & Lang, 1999). Most, it is probably fair to say, subscribe to some version of a necessary but not sufficient position. By this view, certain executive function processes may be necessary for concepts such as false belief to emerge and be expressed, but domain-specific experiences and knowledge are also necessary. Development, in other words, is *both* domain-general and domain-specific.

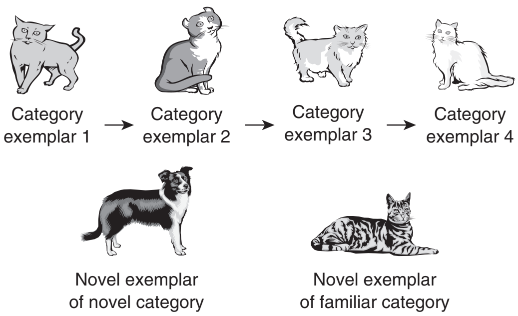


Figure 12.6 Examples of stimuli used in the study of concept formation in infancy

SOURCE: From "Early Category Representation and Concepts" (p. 29), by P. C. Quinn and J. Oates. In J. Oates & A. Grayson (Eds.), *Cognitive and Language Development in Children* (pp. 21–60), Oxford, UK: Blackwell Publishing. Reprinted by permission of Blackwell Publishing, Oxford, UK.

Imagine that we first show the infant a series of cat pictures, such as those in the top part of Figure 12.6. We present one picture at a time, with each picture different from the preceding one. After this initial *familiarization phase,* we run a test using the Fantz preference method. The picture in one window is of a cat that has not been presented before; the picture in the other window is of a dog. Both stimuli, therefore, are new to the baby. The 3-month-old, however, does not treat them as equally new; rather the baby looks more at the dog than at the cat.

What findings such as this demonstrate is that even a young infant is not limited to encoding and remembering only stimuli that have been directly experienced. If this were the case, then the cat in the bottom part of the figure would seem no more familiar than the dog. Rather, from early in life infants go beyond their specific experiences to form *concepts* for familiar stimuli and events. The term concept refers to a mental grouping of different items into a single category on the basis of some underlying similarity—something that makes them all, in a sense, the "same thing." Concepts are a basic way in which we organize and make sense of the world; indeed, without concepts virtually everything we encounter would be experienced as new. Thus, we treat different cats as in some sense the same thing, and so, apparently, do 3-month-olds.

Of course, it may have occurred to you that the cat versus dog comparison is not necessarily clear evidence for the existence of concepts. Perhaps dogs are just more interesting to babies at this age. Or perhaps the specific pictures used provide bases for response other than catness or dogness. These possibilities have also occurred to the investigators who do such research, and they have been careful to build in controls to rule out other explanations for the results. Thus there seems little doubt that 3-month-olds can distinguish cats from dogs. They also can distinguish cats from birds and horses (not, however, from female lions—Quinn, 1999). Moreover, the distinctions that they can make are not limited to biological categories but extend also to human-produced artifacts. Three-month-olds can distinguish animals from furniture, and they can make some differentiations within the category of furniture—can distinguish chairs from couches, beds, and tables, for example (Behl-Chadha, 1996).

The preference method is just one of several techniques that can be used to study concept formation in infancy (Flavell et al., 2002). All are based on the same rationale: See whether the infant responds more similarly to members of a category than to nonmembers, and thus seems in some way to be treating the category members as instances of the same thing. The habituation-dishabituation procedure is another possible

source of evidence. We might present a series of cat pictures until interest declines and then test for dishabituation to one of two stimuli: another cat or a dog. Operant conditioning has also been employed. We might use the Rovee-Collier paradigm (described earlier in the chapter) to condition response to a particular mobile, perhaps one covered with cat pictures, then test for generalization to new mobiles that vary in their similarity to the original.

As infants grow older, a wider variety of response measures becomes available. By about a year of age the *sequential touching procedure* can be used. In this case we present an array of objects from different categories (e.g., toy animals, toy vehicles, toy furniture) and record the order in which the infant touches the objects (e.g., Mandler & Bauer, 1988). If the infant shows an above-chance tendency to touch items from the same category in succession—which is, in fact, a typical finding—then we have evidence that the infant considers them as instances of the same thing.

Finally, with slightly older infants it is possible to examine the *inferences* or *inductions* that the infant can draw from knowledge of category membership. In one study, for example, 14-month-olds first imitated various actions by an adult model, such as giving a dog a drink and turning a key in a car door. They were then given a chance to reproduce the actions with new members of the animal and vehicle categories. Even when the new instances were unfamiliar (e.g., an armadillo as the animal), the infants generalized appropriately, performing animal-appropriate activities only with animals and vehicle-appropriate activities only with vehicles (Mandler & McDonough, 1996). The ability to understand and respond appropriately to new instances of a category is a main function that concepts serve. Such findings indicate that children begin to use their concepts adaptively from very early in development.

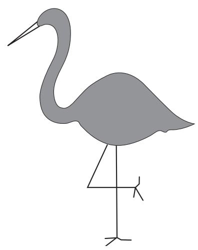
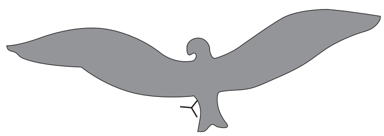
Older Children

As is always the case, the response measures available to us expand when we move from infants to older children. In particular, it is now possible to convey instructions through words and to receive responses in words. Rather than rely on the indirect measures used with infants, we now can simply ask children to tell us or show us which items in some collection go together and to explain why.

Studies of how children sort items into groups date back more than a century, including some influential work by Piaget (Inhelder & Piaget, 1964). A general conclusion from such studies (a fairly safe conclusion for any aspect of cognitive development) was of early limitations that were gradually overcome with development. In particular, young children often appear perceptually oriented, concentrating on surface features to the neglect of more basic underlying properties. Presented, for example, with a fire engine, car, and apple, most 3- and 4-year-olds group together fire engine and apple, the two red things, rather than fire engine and car, the two vehicles (Tversky, 1985). Note that the perception-oriented nature of the young child is also evident in various findings discussed earlier in the chapter—Piaget's nonconservation phenomenon, for example, or the preschooler's failure to distinguish appearance and reality.

While not negating these early limitations, more recent research suggests a more positive picture of preschoolers' abilities. Table 12.3 shows one example. The problems in the table are from a program of research by Gelman, Markman, and associates (Gelman, 2000; Gelman & Markman, 1986, 1987). The question is how the child will generalize from what is already known to something new, and two possible bases for response are contrasted. If perceptual similarity is taken to be critical, then the new item should be judged to be like the one that it most resembles. This means, for example,

Table 12.3 Sample Items From Gelman and Markman's Studies of Children's Concepts

This bird's legs get cold at night. (picture of flamingo)  This bat's legs stay warm at night. (picture of black bat)  See this bird. Do its legs get cold at night, like this bird, or do its legs stay warm at night, like this bat? (picture of blackbird, looks like the bat) (see pictures) 
This fish stays under water to breathe. (picture of tropical fish) This dolphin pops above the water to breathe. (picture of dolphin) See this fish. Does it breathe under water, like this fish, or does it pop above the water to breathe, like this dolphin? (picture of shark, looks like the dolphin)
This puppy hides bones in the ground. (picture of brown dachshund) This fox hides food in the ground. (picture of red fox) See this puppy. Does it hide bones in the ground, like this puppy, or does it hide food in the ground, like this fox? (picture of red dog, looks like the fox)

SOURCE: From "Categories and Induction in Young Children," by S. A. Gelman and E. M. Markman, 1986, *Cognition, 23*, 183–209. Copyright © 1986. Reprinted by permission of Elsevier.

that the blackbird would be expected to have warm legs at night, just like the similar-looking bat. In contrast, if category membership is deemed more important, then the legs would be expected to be cold, just like those of the other bird. The contrast in these tasks is a basic one in studies of children's concepts: between surface similarity and underlying essence as the basis for judging that things are the same.

As noted, much of the research prior to Gelman and Markman's had indicated that young children rely on surface similarity. This was not what Gelman and Markman found. Despite the compelling perceptual cues, most 4-year-olds opted for category membership as the relevant basis for inference. Thus, they judged that the bird's legs would get cold, that the shark would breathe under water, and so forth. A subsequent study using simplified procedures showed that even 2-year-olds had some ability to overlook perceptual appearance in favor of category membership (Gelman & Coley, 1990).

Why the more positive picture of preschoolers' competence in recent research? Two factors are probably important. One is the methods used. As noted, many of the older studies based their conclusions on children's response to explicit instructions to sort items into categories (e.g., "Show me which ones go together"). The Gelman and Markman procedure, in contrast, is tied to a natural, everyday use of concepts—drawing inferences about new instances from what is already known. The grounding of the response measure in this natural function of concepts may be one explanation for the impressive performance in recent research.

The type of concept at issue is probably important as well. Some studies have used arbitrary concepts created on the spot for the purposes of research—for example, the category of blue circles in a study of sorting behavior. There is nothing arbitrary, however, about the concepts that children naturally form—concepts of people, of animals, of vehicles, of food. Such concepts reflect important commonalities among real-life experiences that children extract as they attempt to make sense of the world. The focus on familiar and interesting material may also contribute to the good performance in recent research.

Further evidence for this conclusion comes from perhaps the most active current area in the study of conceptual development: research directed to children's concepts of biology. Again, the main message from the older research literature was one of confusions and limitations in young children's understanding. The best known example of such confusions is Piaget's (1929) concept of **animism**: the tendency to attribute properties of life to nonliving things. The young child who indicates that the sun shines "because it wants to" is engaging in animistic thinking, as is the child who is concerned that a piece of paper will be hurt by being cut. Obviously, such thinking reflects a fundamental gap in biological understanding.

As with concepts in general, contemporary research suggests a more positive picture of young children's competence. This understanding is, to be sure, far from complete, and there are disagreements about exactly how much knowledge to attribute to young children (Astuti, Solomon, & Carey, 2004). Still, when questioned in an optimal fashion, even preschoolers do seem to appreciate some of the basic properties of life. Here are a few examples.

One fundamental property of life is growth. Children as young as 3 or 4 understand some of the basic facts about growth. They realize, for example, that only living things grow, that growth is inevitable (for example, you can't keep a baby pet small and cute just because you want to), and that growth is directional—that is, people, plants, and animals get bigger, not smaller, as they age (Inagaki & Hatano, 1996, 2002; Rosengren, Gelman, Kalish, & McCormick, 1991). They also have some appreciation of the fact that only animals *re*grow—that a cat with a scratch, for example, will eventually heal, but that a car with a scratch will require human intervention (Backschneider, Shatz, & Gelman, 1993).

Whereas growth is characteristic of all life, only animals are capable of self-produced movement. Many other things, of course, do move (cars, bikes, clouds, leaves in the wind, etc.), and young children, as Piaget's studies of animism indicate, are sometimes confused by this fact. Still, even 3-year-olds show some success in judging which things can move by themselves and which things (e.g., statues, plants) cannot (Massey & Gelman, 1988).

One final property of life that has been the subject of much study is inheritance. Living things come from other living things, and they inherit properties both of the species in general and of their own parents in particular. Again, even preschoolers demonstrate some biological knowledge, in this case of origins and kinship. They realize that dogs produce baby dogs, not cats, and that offspring generally resemble their parents (Springer, 1996; Springer & Keil, 1991). By the grade-school years children can also make sensible judgments about which properties are likely to be biological in origin (e.g., physical appearance, dietary preferences) and which are likely to be environmental (e.g., personality traits—Astuti et al., 2004; Springer, 1996).

Summary

This chapter considers five topics that fall under the heading of cognitive development in childhood.

The chapter begins with the historically influential Piagetian approach. Just as is true for infancy, Piaget's research on later childhood is distinguished by its ability to identify interesting and developmentally basic forms of knowledge, all studied through the flexible *clinical method.* Among these forms of knowledge, conservation—the knowledge that quantities remain invariant in the face of perceptual change—has proved especially intriguing. To the discussion of conservation are added descriptions of four other Piagetian concepts: *class inclusion, transitivity, formal-operational reasoning,* and *perspective taking.*

Piaget's work has inspired a host of follow-up studies. A basic question is that of *assessment:* How accurately do our experimental procedures assess the child's abilities? Piaget has long been charged with underestimating the child's ability, primarily because of the heavy verbal emphasis in many Piagetian tasks. That there is some validity to this charge is suggested by a discussion of two sorts of studies: those that have simplified the language involved in Piagetian assessment and those that have increased the naturalness of the assessment situation.

The intelligence test or IQ approach is in many ways quite different from the Piagetian approach. An overview of the differences serves as a lead-in to a discussion of the types of items that appear on two leading IQ tests: the Stanford-Binet and the WISC. These tests are highly standardized instruments whose purpose is to identify individual differences among children. Their content is oriented to the kinds of academic-verbal skills that are important in school. Indeed, it is the predictive power of the tests, including correlations with school performance, that has always served as their chief validity index.

The description of IQ tests is followed by a consideration of some of the theoretical issues that have been of interest in the study of IQ. One is the question of *stability:* Is a child's IQ constant across development, or can the value go up or down? Answering this question requires a longitudinal approach; hence this section of the chapter serves as a reminder of some of the points about longitudinal designs that were made in chapter 3. Another basic question concerns the *determinants of differences* in IQ—specifically, the extent to which such differences are genetic or environmental in origin. Two approaches to this question are reviewed. Twin studies capitalize on the fact that identical twins are more genetically alike than are fraternal twins; greater similarity in IQ for identicals is then taken as evidence for the importance of genetic factors. Adopted child studies offer the opportunity to disentangle the genetic and environmental factors that are confounded in the normal parent-child correlation. The adopted child's IQ can be compared with those of the biological parents and with those of the adoptive parents. Influential though the twin and adopted child studies have been, they do have their limitations, and these limitations are discussed along with the findings.

A topic of longstanding interest in developmental psychology is children's memory. Memory is a basic cognitive process that is present from birth. Memory in early infancy seems to consist mainly of *recognition memory,* which has typically been studied through either the habituation-dishabituation paradigm or the Fantz preference method. Work on deferred imitation suggests that some capacity for *recall memory* emerges by a few weeks of age.

Studies of older children have concentrated on recall memory and on the attempt to explain the developmental improvement in recall across the childhood years. Research on strategies attempts to measure the existence and effects of various *mnemonic strategies* that can be used to aid recall. In some cases it is possible to observe the spontaneous occurrence of the strategy; in other cases it may be necessary either to infer the use of the strategy from overt memory performance or to induce its use experimentally. Research on *constructive memory* examines the effects of the general knowledge system, and of developmental changes in this system, on memory. Several examples of studies of constructive memory are discussed, including the effects of scripts and of expertise on memory.

The term *theory of mind* refers to children's thoughts and beliefs about the mental world. Of special interest has been the concept of *false belief:* the realization that it is possible for people (both others and oneself) to hold beliefs that are false. Another important development is mastery of the *appearance-reality* distinction: the ability to distinguish appearance and reality when the two are in conflict. Several specific methods of study are described not only for false belief and appearance-reality but also for a third focus of theory-of-mind research: children's understanding of the *origins of knowledge.* As is true with Piagetian research, issues of *assessment* are important in studies of theory of mind, and several alternative approaches to the study of false belief are therefore described. Another important issue concerns possible links between theory-of-mind understanding and social experience. A variety of kinds of evidence indicate that there are in fact links and that the causal relations run in both directions.

The chapter concludes with research directed to the formation of *concepts*—that is, mental groupings of items on the basis of some underlying similarity. Infants begin to form simple concepts from early in life, as shown by differential response to members versus nonmembers of particular categories (such as cats and dogs). The specific procedures used to demonstrate such response include the preference method, habituation-dishabituation, sequential touching, and generalization. Such nonverbal procedures are supplanted by verbal techniques in studies with older children, but the rationale remains the same: look for differential response to members of different categories. Recent research indicates that even preschoolers sometimes weight category membership more heavily than perceptual similarity when drawing inferences from what they know. Recent research also indicates that basic forms of biological understanding begin to emerge early in development.

Exercises

1. One of the attractive features of both Piagetian and theory-of-mind tasks is that they are simple to administer and hence lend themselves readily to demonstration projects. The following exercise assumes that you have access to at least one (and preferably several) children with whom you can do simple testing. If the children are between 3 and 5 years old, you should concentrate on theory-of-mind measures; if they are older, you should concentrate on Piagetian tasks. In either case, read both the descriptions in the

textand some of the original sources that are cited, try out some of the tasks yourself, and compare your results with those in the literature. Note: Be sure to obtain the informed consent of both the children and their parents before beginning testing.

2. Follow-up studies of Piaget's final period of development indicate that many adults show less than perfect performance on tasks intended to assess formal-operational thinking. Obtain a copy of the Inhelder and Piaget book cited in the text, reproduce as many of the tasks as you are able, and administer the test battery to a sample of your peers. Do you believe that tasks such as these provide a sufficient measure of the capacity for hypothetical-deductive reasoning? If not, how would you modify or add to the assessment?

3. One criticism of IQ tests has been that they measure only certain kinds of intellectual skills to the neglect of other forms of cognitive competence. One often cited area of neglect is the domain of social or interpersonal intelligence. Imagine that you had the task of devising a test of social intelligence for grade-school children. How would you select items for your measure? How would you validate your test?

4. Think back on your own mnemonic activities of the last week or so. To what extent do they reflect the sorts of processes discussed in the text? Suppose that a researcher wished to study your everyday memory activities. How could the researcher go about collecting evidence?