

strategies including heuristics or rules of thumb. More specifically, decision making can be “bounded” by constraints in the environment (e.g., information costs; time costs) or by constraints in the mind (e.g., limited attention; limited memory). What is important is the degree of fit or match between the mind and the environment. According to Simon (1990, p. 7), “Human rational behavior is shaped like a scissors whose blades are the structure of the task environment and the computational capabilities of the actor.” If we consider only one blade (i.e., the task environment or the individual’s abilities) we will have only a partial understanding of how we make decisions. In similar fashion, we couldn’t understand how scissors cut if we focused on only one blade.

How does bounded rationality work in practice? According to Simon (1978), we often use a heuristic or rule of thumb known as **satisficing**. The essence of **satisficing** (formed by combining satisfactory and sufficing) is that individuals consider various options one at a time and select the first one meeting their minimum requirements. Satisficing does not guarantee the decision is the best one, but it is especially useful when the different options become available at different points in time. For example, you might decide who to marry on the basis of the satisficing heuristic by setting a minimum acceptable level. The first person reaching (or exceeding) that level would be chosen. If you set the level too low, you may spend many years bitterly regretting having used the satisficing heuristic!

Tversky (1972) put forward a theory of decision making resembling Simon’s bounded rationality approach. According to his elimination-by-aspects theory, decision makers eliminate options by considering one relevant attribute after another. For example, someone buying a house may initially consider the attribute of location, eliminating all houses not lying within a given area. Next they may consider the attribute of price, eliminating all those properties costing above a certain figure. This process continues until only one option remains. This strategy has the advantage of being fairly undemanding. However, the option selected depends on the order in which the attributes are considered, and so frequently someone using this strategy won’t make the best choice.

Dijksterhuis and Nordgren (2006) proposed unconscious thought theory. According to this theory, there is an important distinction between conscious thought (involving attention) and unconscious thought (thinking occurring without attention). Most people assume that complex decision making is better when people use conscious thought, but Dijksterhuis and Nordgren argued precisely the opposite. Conscious thought has the disadvantage of limited processing capacity (because of the involvement of attention). In addition, when we are thinking consciously, we often attribute excessive weight or importance to only a small fraction of the information presented to us. Since unconscious thought does not possess these disadvantages, it can produce superior decision making to conscious thought when the problem is a complex one.

Findings

Payne (1976) addressed the issue of the extent to which decision makers behave as predicted by Simon and by Tversky. Participants decided which apartment to rent on the basis of information about various attributes (e.g., rent; cleanliness; noise level; and distance from campus). The number of apartments to be considered ranged between 2 and 12 and the number of attributes per apartment varied between 4 and 12.

When there were many apartments to consider, participants typically started by using a simple strategy such as satisficing or elimination-by-aspects. Here is one participant’s account of using elimination-by-aspects: “I’m going to look at landlord attitude. In H it’s fair. In D it’s poor. In B it’s fair, and in A it’s good. In L, the attitude is poor. In K it’s poor. In J it’s good, and in I it’s poor . . . So, that’s important to me . . . So, I’m not going to live any place where it’s poor” (Payne, 1976, p. 379). When only a few apartments remained to be considered, there was often a switch to a more complex strategy corresponding to the assumptions of multi-attribute utility theory.

An important factor influencing decision-making strategies is time pressure. Payne, Bettman, and Johnson (1988) found that moderate time pressure led decision makers to become slightly more selective in their processing. When the time pressure was severe, however, decision makers speeded up their processing, focused their attention on only a fraction of the available information, and changed their processing strategies.

Key Term

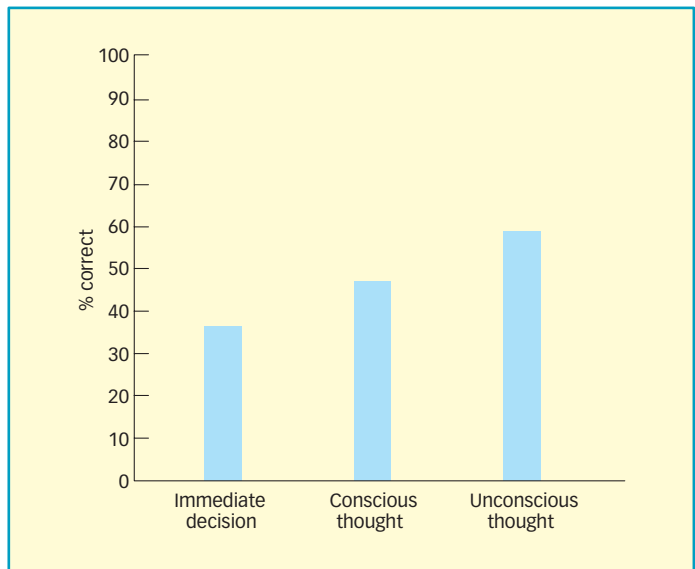
Satisficing: selection of the first choice that meets certain minimum requirements; the word is formed from the two words satisfactory and sufficing.

It seems reasonable to assume (and it was assumed within multi-attribute utility theory) that any given individual's assessment of the utility or preference (desirability \times importance) of any given attribute remains constant over time. In fact, our preferences are actually subject to change. Simons, Krawczyk, and Holyoak (2004) asked participants to decide between job offers from two department store chains, "Bonnie's Best" and "Splendor." There were four relevant attributes (salary; holiday package; commuting time; and office accommodation). Each job offer was preferable to the other on two attributes and inferior on two attributes. Participants assessed their preference for each attribute. They were then told that one of the jobs was in a much better location than the other. This often tipped the balance in favor of choosing the job in the better location. The participants then assessed their preference for each attribute again. Preferences for desirable attributes of the chosen job increased and preferences for undesirable attributes of that job decreased. Such preference changes make sense because they increase decision makers' satisfaction with their choice.

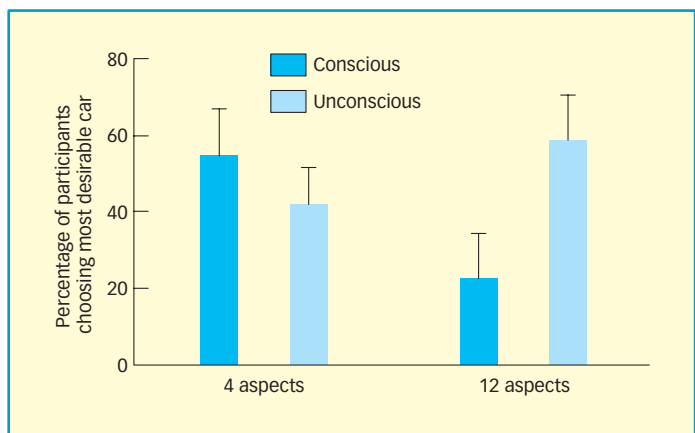
We turn now to unconscious thought theory. Some of the most convincing evidence in its support was obtained by Dijksterhuis (2004), who used three conditions in research on decision making. In the control condition, participants made immediate decisions when the various options had been presented. In the conscious thought condition, participants had a few minutes to think about their decision. In the unconscious thought condition, participants were distracted for a few minutes and then made their decision in the virtual absence of conscious thought.

Participants were presented with detailed information about four hypothetical apartments in Amsterdam. Each apartment was described in terms of 15 attributes. The most attractive apartment had 8 positive, 4 negative, and 3 neutral attributes, whereas all the others had 5 positive, 6 negative, and 4 neutral attributes. The task was to select the best apartment. Performance was best in the unconscious thought condition and worst in the control condition (see the figure on the right). The participants were then asked whether their choice was based on a global judgment or on one or two specific attributes. Far more of those in the unconscious thought condition than in the conscious thought condition indicated that they had made a global judgment (55.6% vs. 26.5%, respectively). This suggests that the relatively poor performance in the conscious thought condition occurred because participants focused too much on only a small fraction of the available information.

We have seen that unconscious thought can lead to superior decisions to conscious thought when decision making is complex. However, it is predicted by unconscious thought theory that there should be an interaction between mode of thought and complexity of decision making. Since conscious thought is precise but has limited capacity, it is well suited to simple decision making but becomes increasingly ineffective as decision making becomes more complex. Findings consistent with these prediction were reported by Dijksterhuis, Bos, Nordgren, and van Baaren (2006). Participants read information concerning four hypothetical cars. In the simple condition, each car was described in terms of 4 attributes. In the complex condition, each car was described in terms of 12 attributes. Participants either spent 4 minutes thinking about the cars (conscious thought condition) or they solved



Performance on an apartment-selection task as a function of whether the decision was immediate or delayed and based on conscious or unconscious thought. Data from Dijksterhuis (2004).



Percentage of participants choosing the most desirable car as a function of decision complexity and mode of thought (conscious vs. unconscious). From Dijksterhuis et al. (2006). Copyright © 2006 AAAS. Reproduced with permission.

anagrams for 4 minutes before choosing a car (unconscious thought condition). The findings are shown in the figure on the previous page). As predicted, more participants in the unconscious thought condition than in the conscious thought condition selected the most desirable car when the decision was complex, with the opposite pattern being found when the decision was simple.

In sum, it is clear that unconscious thought can be superior to conscious thought when people are confronted by a complex problem. However, that is only the case under certain conditions. In Dijksterhuis's experiments, participants in the conscious thought condition had to rely heavily on their memory for the information contained in the problem. It is probable that conscious thought would have proved more effective than unconscious thought if participants had been allowed access to problem information throughout the time preceding their decision.

Evaluation

- + There is much support for the assumptions that humans possess bounded rationality and frequently use the satisficing heuristic.
- + Several factors influencing decision making (e.g., number of choices remaining to be considered; extent of time pressure) have been identified.
- + Some of the strengths and limitations of conscious thought and of unconscious thought have been identified.
- Most theories implicitly assume that attribute preferences don't change over time, but there is evidence (e.g., Simons et al., 2004) that this assumption is incorrect.
- Insufficient attention has been paid to individual differences. For example, Schwartz et al. (2002) found that some individuals are satisficers (content with making reasonable decisions), whereas others are maximizers (perfectionists). Satisficers are happier than maximizers and experience less regret and self-blame.
- Superiority of unconscious thought over conscious thought with complex decision making would probably not be found if people had full access to problem information throughout the decision-making process.

Chapter Summary

Finding the solution

- The phenomenon of functional fixedness and our persistence with problem-solving strategies that are no longer appropriate reveal that past experience isn't always useful for problem solving.
- Beneficial effects of past experience on current problem solving (positive transfer) often involve making use of analogies.
- Positive transfer over long periods of time and to new situations is facilitated by metacognitive training (e.g., orienting; self-judging).
- Insight typically consists of a sudden restructuring of a problem in which an incorrect problem representation is replaced by the correct one. This can involve constraint relaxation, re-encoding, or elaboration.
- Incubation (putting a problem to one side for a while) often leads to effective problem solving, probably because it produces forgetting of incorrect strategies.
- According to Newell and Simon in their General Problem Solver, we have limited processing capacity and so make use of heuristics or rules of thumb (e.g., means-ends analysis) in problem solving.

- Problem solvers engage in progress monitoring, changing their strategy when progress towards a solution is too slow.

Decisions, decisions

- We can assess the quality of decision making by focusing either on the outcome or on the reasoning involved.
- According to prospect theory, people are much more sensitive to potential losses than to potential gains. The theory explains the framing effect, why people are unwilling to accept potential losses even though the potential gains exceed the potential losses, and why they prefer a sure gain to a risky but potentially greater gain.
- Prospect theory is limited because insufficient attention is paid to emotional, social, and moral factors, to individual differences, and to cross-cultural differences.
- Decision making (even among medical experts) is often influenced by the great importance we attach to justifying our decisions to others.
- People possess bounded rationality and often use the satisficing heuristic when engaged in complex decision making.
- Complex decision making is influenced by time pressure, by the number of options to be considered, and by changing attribute preferences.
- Conscious thought has been found to be superior to unconscious thought on simple decision-making tasks, with the opposite being the case with complex tasks.

Further Reading

- Bowden, E.M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences*, 9, 322–328.
- Eysenck, M.W. (2006). *Fundamentals of cognition*. Hove, UK: Psychology Press. The topics discussed in this chapter are dealt with more fully in Chapters 23 and 24 of this introductory textbook on cognitive psychology.
- Hardman, D., & Macchi, L. (2003). *Thinking: Psychological perspectives on reasoning, judgement and decision making*. Chichester, UK: Wiley. Several chapters in this edited book provide good reviews of decision making.
- Koehler, D.J., & Harvey, N. (2004). *Blackwell handbook of judgement and decision making*. Oxford, UK: Blackwell. This book contains interesting chapters by several well-known experts on decision making.
- Robertson, S.I. (2001). *Problem solving*. Hove, UK: Psychology Press. There is good coverage of many topics in problem solving in this introductory textbook.
- Sternberg, R.J. (2003). *Cognitive psychology* (3rd ed.). Belmont, CA: Wadsworth. This textbook provides a useful introductory account of research on problem solving.

chapter 10

Contents

Reading	237
Speech perception	243
Understanding language	247
Speech production	254

Language

10

Language is of absolutely central importance to our lives. Our social interactions with other people depend very heavily on language, and language plays a key role in most of our thinking. Imagine trying to study psychology without a good command of language! The main reason why we are much more knowledgeable than people of previous generations is that knowledge (mostly in the form of language) is passed on from one generation to the next.

What is language? According to the *Oxford Dictionary of Psychology* (2001), language is, “a conventional system of communicative sounds and sometimes (though not necessarily) written symbols.” The crucial word is “communicative”—the main purpose of language is communication. However, Crystal (1997) identified a total of *eight* language functions, of which communication was one. We also use language for thinking, to record information, to express emotion (e.g., “I love you”), to pretend to be animals (e.g., “Woof! Woof!”), to express identity with a group (e.g., singing in church), and so on.

There are four main language skills: reading; listening to speech; speaking; and writing. It is easy to assume that any given person will have generally strong or weak language skills. This assumption may sometimes (but certainly not always) be correct with respect to an individual’s first language. However, it is often incorrect with second language acquisition. After having spent 10 years at school learning French, I can just about read newspapers and easy novels in French and can write French reasonably well. However, I find it agonizingly difficult to understand rapid spoken French, and I have very limited ability to speak French.

In this chapter, we are concerned with three of the four main language skills with writing being omitted. As someone who has written over 7 million words for publication, that isn’t because I think writing is unimportant! The reason is that much less is known about the processes involved in writing than any other language skill. We start with a discussion of the basic processes involved in reading, followed by an analysis of what is involved in speech perception. After that, we focus on the comprehension processes used to understand language whether presented in the form of text or spoken words. Finally, we focus on the processes involved in speech production.

READING

Reading is an extremely important skill in most societies, and adults without effective reading skills are at a great disadvantage. It is also important to study reading because most mental activities are related to reading. Indeed, it has sometimes been described as “visually guided thinking.”

We focus on basic reading processes mainly at the level of the individual word within sentences. In addition, we consider the processes involved in working out the syntactical or grammatical structure of sentences. Processes involved in the comprehension of texts are discussed later in the chapter.

EYE MOVEMENTS

We can obtain some understanding of basic processes in reading by recording readers’ eye movements. Our eyes seem to move smoothly across the page when reading but actually move in rapid jerks (**saccades**). Reading saccades take 20–30 ms to complete and are separated by fixations lasting for 200–250 ms. The length of each saccade is about

Key Term

Saccades:
fast eye movements or jumps that cannot be altered after being initiated.



Reading is a vital skill in most societies and a lack of this skill in adults is a distinct disadvantage.

eight letters or spaces. Information is extracted from the text only during each fixation and not during the intervening saccades.

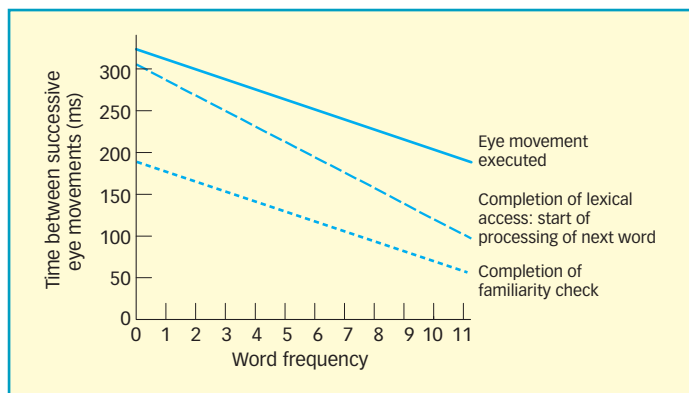
Readers typically fixate 80% of content words (nouns, verbs, and adjectives) compared to 20% of function words (articles, conjunctions, and pronouns). Words not fixated tend to be common, short, or predictable. In contrast, words fixated for longer than average are generally rare words or words unpredictable in the sentence context. Common words are usually more predictable than rare words, making it important to disentangle the effects of word frequency and word predictability. McDonald and Shillcock (2003) found that many of the apparent effects of word frequency on length of eye fixations were actually a result of word predictability. Finally, there is the **spillover effect**: The fixation time on a word is longer when preceded by a rare or unpredictable word rather than a common or

predictable one.

How can we explain these various findings? Perhaps readers fixate a word until they have processed it adequately, after which they immediately fixate the next word until it in turn has been adequately processed. Alas, there are two problems with that viewpoint. First, it takes upwards of 80 ms to execute an eye-movement program. If readers didn't start to plan the next eye movement until they had finished processing the current word, they would waste time waiting for their eyes to move to the next word. Second, it is difficult to see how readers could decide to skip words, because they would know nothing about the next word until they had started fixating it.

Reichle et al. (1998, 2003) provided an answer to the question posed in the previous paragraph in their E-Z Reader model (see the figure on the left). They argued that the next eye movement is programmed after only *part* of the processing of the currently fixated word has occurred. This assumption makes a lot of sense, because it greatly reduces the time between completion of processing on the current word and movement of the eyes to the next word. Any spare time is used to start processing the next word. There is typically less spare time available with rare or unpredictable words than with common or predictable ones, and that accounts for the spillover effects described above. If the processing of the next word is completed rapidly enough (e.g., it is highly predictable in the sentence context), it is skipped.

The E-Z Reader model indicates the processes determining *when* we move our eyes in reading and *where* we move them. However, it is limited in that it doesn't tell us much about *what* kinds of processing readers are engaged in during reading.



The effects of word frequency on eye movements according to the E-Z Reader model. Adapted from Reichle et al. (1998).

SOUND IN READING

It is easy to think of reading as being only a visual skill since all the information presented to the reader is visual. However, when reading a really difficult chapter in a book, you have probably found yourself muttering some of the words in the text to yourself. This suggests that **phonology** (the sound of words) may play an important role in reading. According to Frost's (1998) strong phonological model, phonological processing is *always* involved in reading. In fact, as we will see, phonological processing is often (but not always) used in reading.

Phonological processing is more likely to be used when reading is difficult. For example, Jared, Levy, and Rayner (1999) found using a proofreading task that the use of phonology depended on the nature of the words and on participants' reading ability. More specifically, phonological processing was used to access the meaning of low-frequency

Key Terms

Spillover effect:

any given word is fixated longer during reading when it is preceded by a rare or unpredictable word rather than a common or predictable word.

Phonology:

information about the sounds of words and parts of words.

words but not high-frequency ones. In addition, poor readers were more likely than good readers to use phonological processing to access meaning.

Hanley and McDonnell (1997) reported reasonable evidence that accessing word meaning doesn't have to involve prior phonological processing. They studied a patient (PS) who didn't seem to have access to an internal phonological representation of words, and who couldn't pronounce words accurately. In spite of this, PS understood the meaning of words while reading.

READING ALOUD

Reading aloud familiar words (and even most nonwords) seems about as easy a task as you can imagine. Indeed, I would be very surprised if you had any difficulty at all in reading out the following list of words and nonwords:

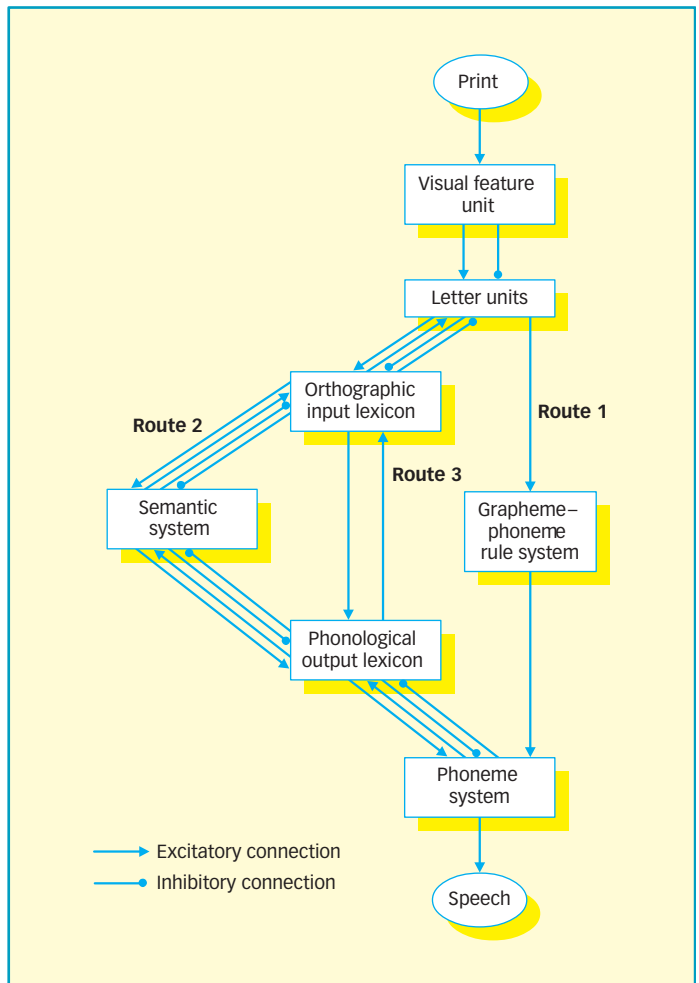
CAT FOG COMB PINT MANTINESS FASS

However, close inspection of what you have just done reveals some hidden complexities. For example, how do you know the “b” in “comb” is silent, and that “pint” doesn't rhyme with “hint”? Presumably you have specific information stored in long-term memory about how to pronounce these words. However, that can't explain your ability to pronounce nonwords such as “mantiness” and “fass” that you have never seen before. Perhaps nonwords are pronounced by analogy with real words (e.g., “fass” is pronounced to rhyme with “mass”). Alternatively, we may use rules governing the translation of letter strings into sounds to generate pronunciations for nonwords.

Dual-route model

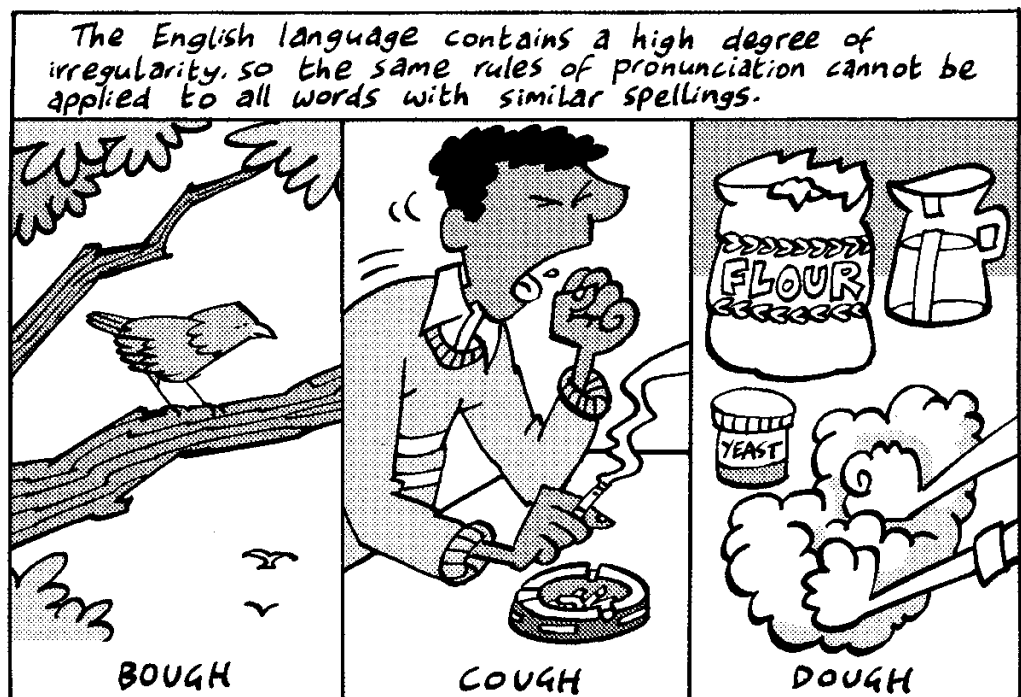
The take-home message from the above paragraph is that reading words aloud can be achieved in more than one way. This insight is at the heart of the dual-route model put forward by Coltheart et al. (2001; see the figure on the right). According to this model, there are three routes between the printed word and speech, all of which start with orthographic analysis (used for identifying and grouping letters in printed words). Why is a model with *three* routes called a *dual*-route model? The reason is that the key distinction is between a lexical look-up route based on a lexicon or mental dictionary (Routes 2 and 3) and a non-lexical route (Route 1). The nonlexical route involves converting spelling (graphemes or basic units of written language) into sound (phonemes or basic units of sound). This is known as grapheme–phoneme conversion. To oversimplify somewhat, people reading aloud will generally rely mainly on the lexical or dictionary look-up route when reading words. However, they will use the non-lexical route when reading nonwords because nonwords don't have dictionary entries. More precisely, words and nonwords activate *both* routes, and what readers say is determined by which route first provides a pronunciation.

How can we test this model? One approach involves identifying brain-damaged patients who are largely reliant on only *one* of the routes. What would happen if a patient had access only to Route 1? The use of rules for converting letters into sounds should permit accurate pronunciation of *regular* words (i.e., words in which the pronunciation is predictable



The dual-route cascaded model. From Coltheart et al. (2001). Copyright © American Psychological Association. Reprinted with permission.

Regular words	Irregular words
fantasy	chemistry
meet	hierarchy
passenger	sapphire
Swiss	Worcester



from the letters). Such words (e.g., “fine”; “save”) can be read with no special knowledge of the word. The use of grapheme–phoneme conversion rules also allows nonwords to be pronounced accurately. However, such patients should struggle with *irregular* words in which the pronunciation is *not* predictable from the spelling (e.g., “island”; “yacht”).

Patients conforming most closely to the above predictions suffer from **surface dyslexia**, a condition in which regular words can be read but irregular ones cannot. McCarthy and Warrington (1986) studied a surface dyslexic (KT) who read 100% of nonwords accurately as well as 81% of regular words. However, he was successful with only 41% of irregular words, with more than 70% of his errors involving treating irregular words as if they were regular.

What would happen if a patient had access only to Route 2, which involves the lexicon or mental dictionary and the semantic system? He/she would be able to pronounce familiar words (whether regular or irregular) and understand them. This is because Route 2 goes through the semantic or meaning system and provides access to information about the sound pattern of words in that system. However, the patient should find it difficult to pronounce unfamiliar words and nonwords, because he/she wouldn’t be able to use grapheme–phoneme conversion. Patients with **phonological dyslexia** fit this pattern, because it involves problems with reading unfamiliar words and nonwords. For example, Beauvois and Derouesné (1979) presented a phonological dyslexic (RG) with 40 words and 40 nonwords. RG read 100% of the real words accurately but only 10% of the nonwords.

What would happen if a patient had access only to Route 3, in which the lexicon is used but not the semantic system or grapheme–phoneme conversion? Such a patient would be able to pronounce regular and irregular words accurately. However, he/she wouldn’t understand the meanings of regular or irregular words, and would have great problems with nonwords. Some phonological dyslexics fit this pattern. For example, Funnell (1983) found that WB could read about 85% of words but had poor ability to make semantic judgments about these words. These findings support the notion that WB bypassed the semantic system when reading words. As predicted, WB couldn’t pronounce any nonwords.

Studies on healthy individuals have also produced findings consistent with the model. Seidenberg et al. (1984) found that irregular words took longer to name than regular ones with low-frequency words but not with high-frequency ones. According to the model, the lexical route operates more slowly with low-frequency words than with high-frequency ones. As a result, the naming of low-frequency irregular words is especially slow because

Key Terms

Surface dyslexia: a condition in which brain-damaged patients cannot read irregular words but can read regular ones.

Phonological dyslexia: a condition in which familiar words can be read but there is impaired ability to read unfamiliar words and nonwords.

conflicting information about their pronunciation is received from the lexical and nonlexical routes.

Coltheart et al. (2001) developed a computational model to test their theoretical approach. They presented the model with 7981 words, and found that 7898 (99.5%) were read accurately. When the model was presented with 7000 one-syllable nonwords, it read 98.9% of them correctly.

Evaluation

- + The dual-route model has been applied successfully to reading in healthy individuals.
- + The dual-route model provides a very good account of reading disorders such as surface dyslexia and phonological dyslexia.
- The model is very good at naming words and nonwords accurately. However, it is much less successful at accounting for naming *times* (especially for words).
- Coltheart et al. (2001, p. 236) admitted their model couldn't be applied to several major languages: "The Chinese, Japanese, and Korean writing systems are structurally so different from the English writing system that a model like the [dual-route model] would simply not be applicable: for example, monosyllabic non-words cannot even be written in the Chinese script or in Japanese kanji, so the distinction between a lexical and non-lexical route for reading aloud cannot even arise."

PARSING

It is very important in reading a sentence to identify its syntactical (grammatical) structure; this is known as **parsing**. Grammar is concerned with the way words are combined within a sentence to convey meaning. Much research on parsing has focused on ambiguous sentences. Why is that the case? Parsing typically occurs very rapidly and accurately, and this makes it difficult to study the processes involved. However, observing the problems readers have with ambiguous sentences provides revealing information about parsing processes. Such sentences can also be the source of jokes. For example, Groucho Marx in the film *Animal Crackers* came out with the line: "One morning I shot an elephant in my pyjamas. How he got into my pyjamas, I don't know."



Key Term

Parsing:
an analysis of the syntactical or grammatical structure of sentences.

A central issue concerns *when* semantic (meaning) information is used in parsing. According to one-stage models, *all* sources of information (syntactic and semantic) are used at the same time to construct a syntactic model of sentences. In contrast, the first stage of processing in two-stage models uses *only* syntactic information, with semantic information being used during the second stage. The most influential two-stage approach is the garden-path model (Frazier & Rayner, 1982) and the most influential one-stage approach is the constraint-based theory of MacDonald, Pearlmuter, and Seidenberg (1994).

Two-stage approach

Frazier and Rayner's (1982) garden-path model was given that name because readers can be misled or "led up the garden path" by ambiguous sentences. Its crucial assumption is that semantic factors don't influence the construction of the initial syntactic or grammatical structure. Support for this prediction was reported by Ferreira and Clifton (1986). Eye movements were recorded while participants read sentences such as the following:

- The defendant examined by the lawyer turned out to be unreliable.
- The evidence examined by the lawyer turned out to be unreliable.

If (contrary to the model) readers initially make use of semantic information, they will experience ambiguity for the first sentence but not for the second. The reason is that the defendant could possibly examine something but the evidence could not. In fact, readers experienced ambiguity equally for both sentences. This implies that semantic information did *not* influence the formation of the initial syntactic structure.

Trueswell, Tanenhaus, and Garnsey (1994) argued that the semantic manipulations used by Ferreira and Clifton (1986) were too weak to allow semantic information to influence the early stages of parsing. Accordingly, they used sentences with stronger semantic constraints. With their sentences, semantic information *was* used at an early stage to assist rapid identification of the correct syntactic structure.

One-stage approach

According to MacDonald et al.'s (1994) constraint-based theory, all relevant sources of information (including semantic) are available immediately to the parser. Competing analyses of the current sentence are activated at the same time, with these analyses being ranked according to the strength of their activation. The syntactic structure receiving the most support from the available information is highly activated. Readers become confused when reading ambiguous sentences if the correct syntactic structure is less activated than one or more incorrect structures.

The study by Trueswell et al. (1994; discussed above) suggested that semantic information is used early in parsing and so is in line with the model. Additional support was provided by Boland and Blodgett (2001). Their main focus was on noun/verb homographs (e.g., duck; train)—words that can be used as nouns or verbs. For example, if you read a sentence starting, "She saw her duck and . . .," you wouldn't know whether "duck" was being used as a noun (" . . . and chickens near the barn") or as a verb (" . . . and stumble near the barn"). According to the theory, readers should initially try to form a syntactic structure in which the homograph is used in its most common meaning. For example, "duck" is a verb more often than a noun, whereas "train" is more often a noun. As predicted, readers had greater problems in parsing when noun/verb homographs were used in their less common meaning.

Boland and Blodgett (2001) pointed out that if meaning is used very early in parsing then readers should take account of information from the preceding context (e.g., the previous sentence). They arranged matters so the prior context was sometimes misleading and sometimes not. Consider the following example in which the context is misleading:

*As they walked around, Kate looked at all of Jimmy's pets.
She saw her duck and stumble near the barn.*

As predicted, it took longer to read the second sentence when the context was misleading than when it was not. However, context influenced parsing *later* in processing than predicted by the theory.

Overall Evaluation

- + There is reasonable support for both one-stage and two-stage models or theories, suggesting that each approach works well in some circumstances.
- + Semantic information is sometimes used from the outset to clarify the syntactic structure of ambiguous sentences as predicted by constraint-based theory. It seems efficient that readers should use all relevant information in this way.
- Semantic information is sometimes used later in processing than assumed by constraint-based theory (e.g., Boland & Blodgett, 2001).
- It has proved difficult to obtain findings distinguishing clearly between the two approaches. As Harley (2001, p. 264) argued, “Proponents of the garden-path model argue that the effects that are claimed to support constraint-based models arise because the second stage of parsing begins very quickly, and that many experiments that are supposed to be looking at the first stage are in fact looking at the second stage of parsing.”

SPEECH PERCEPTION

We can generally understand what other people are saying to us even when they speak in a strange dialect and/or ungrammatically. Indeed, we take our ability to understand others’ speech for granted. This is not surprising in view of the enormous experience we have all had in using the English language and in listening to other people.

Can you think of any reasons why speech perception might be much more complex than it appears? Consider the problems you have probably experienced trying to understand foreigners speaking in a language you studied for several years at school. If your experience is anything like mine, what you hear is someone who seems to be talking incredibly rapidly without any pauses for breath. This illustrates two of the problems listeners have to contend with all the time. First, language is typically spoken at the rate of about 10 **phonemes** (basic speech sounds conveying meaning) per second, so we must process spoken language very rapidly. Second, there is the **segmentation problem**, which is the difficulty of separating out or distinguishing words from the pattern of speech sounds. This problem occurs because speech typically consists of a continuously changing pattern of sound with few periods of silence. This makes it difficult to know when one word ends and the next one begins.

How do we cope with the segmentation problem in everyday life? First, we use our knowledge of what is possible in the English language. For example, a stretch of speech lacking a vowel can’t form a word. Evidence that we use this knowledge was reported by Norris et al. (1997). Listeners found it very easy to detect the word “apple” in “vuffapple,” because “vuff” could conceivably be an English word. In contrast, listeners found it difficult to identify the word “apple” in “fapple” because the [f] couldn’t possibly be an English word.

Second, there is stress. In the English language, the initial **syllable** (a rhythmic unit of speech) is stressed in most content words (e.g., nouns; verbs). Suppose that listeners assume when they hear a word that its first syllable will be stressed. If so, they should find it difficult to identify strings of words in which the stress is *not* on the first syllable.



We tend to take our ability to understand the speech of others for granted, but speech perception is more complex than it might appear.

Key Terms

Phonemes:

basic speech sounds that distinguish one word from another and so convey meaning.

Segmentation problem:

the listener’s problem of dividing the almost continuous sounds of speech into separate **phonemes** and words.

Syllable:

a rhythmic unit of speech; words consist of one or more syllables.

When such strings of words (e.g., “conduct ascents uphill”) were presented very faintly, they were often misheard (Cutler & Butterfield, 1992). Note, however, that stress is mostly used as a cue to identifying word boundaries when the speech signal is faint or impoverished. It is little used when the speech signal is reasonably strong (Mattys, 2004).

Listeners also have to cope with the problem of **co-articulation**. What this means is that how a speaker produces a given phoneme depends in part on the phonemes preceding and following it. Thus, the pronunciation of any phoneme varies, and the listener has to adjust to variations in pronunciation.

It is generally known that deaf people use lip-reading to help them to understand speech. Less well known is that people whose hearing is intact also use lip-reading. Clear evidence was reported by McGurk and MacDonald (1976). They prepared a videotape of someone saying “ha” repeatedly. The sound channel then changed so there was a voice saying “ga” in synchronization with lip movements still indicating “ba.” Listeners said they heard “da,” which is based on a blending of the visual and auditory information. This combining of visual and auditory information when the sources of information are in conflict is known as the **McGurk effect**.

In what follows, we focus mainly on the processes involved in identifying words and sentences in spoken language. Issues concerning listeners’ overall comprehension or understanding of what a speaker is saying are considered later.

CONTEXT EFFECT

We have seen that listeners contend with various problems when working out what a speaker is saying. Not surprisingly, listeners struggling to identify a spoken word often make use of information from the surrounding sentence context. For example, consider a study by Warren and Warren (1970). Listeners heard a sentence in which a small portion of one word had been removed and replaced with a meaningless sound. The sentences used were as follows (the asterisk indicates a deleted portion of a word):

- It was found that the *eel was on the axle.
- It was found that the *eel was on the shoe.
- It was found that the *eel was on the table.
- It was found that the *eel was on the orange.

Listeners’ perception of the crucial element in the sentence (i.e., “*eel”) was influenced by sentence context. Those listening to the first sentence heard “wheel,” those listening to the second sentence heard “heel,” and those exposed to the third and fourth sentences heard “meal” and “peel,” respectively. The auditory stimulus (i.e., “*eel”) was always the same so *all* that differed was the contextual information. This phenomenon is known as the **phonemic restoration effect**.

What causes the phonemic restoration effect? Samuel (1997) identified two major possibilities:

1. Sentence context has a *direct* and immediate effect on speech processing (i.e., the missing phoneme is processed almost as if it were present).
2. Sentence context has an *indirect* effect involving guessing the identity of the missing phoneme after basic speech processing has occurred.

The evidence (e.g., Samuel, 1981, 1987) favors the second possibility. Samuel added noise to the crucial phoneme or replaced the missing phoneme with noise. If listeners processed the missing phoneme as if it were there, they would have heard the crucial phoneme plus noise in both conditions. However, listeners readily distinguished between the conditions, which seems inconsistent with the first possibility.

COHORT MODEL

Two of the most influential theoretical approaches to speech perception are the TRACE model (McClelland & Elman, 1986) and the cohort model originally proposed by Marslen-Wilson and Tyler (1980). We focus on the cohort model because its predictions are better supported than those of the TRACE model.

Key Terms

Co-articulation:

the finding that the production of a **phoneme** in one speech segment is influenced (and distorted by) the production of the previous sound and preparations for the next sound.

McGurk effect:

the **phoneme** perceived in speech is influenced by visual and acoustic information when the two are in conflict.

Phonemic restoration effect:

the finding that listeners are unaware that a phoneme has been deleted from an auditorily presented sentence.

The four key assumptions of the original cohort model are as follows:

1. Early in the auditory presentation of a word, all the words conforming to the sound sequence to that point become active. This is the “word-initial cohort.”
2. Words belonging to this word-initial cohort are subsequently eliminated if they don’t match additional information from the word being presented, or if they are inconsistent with the semantic or other context.
3. Various knowledge sources (e.g., lexical; syntactic; semantic) are processed in parallel (at the same time). These knowledge sources interact and combine with each other.
4. Processing of the presented word continues up until the moment at which information from the word itself and contextual information have eliminated all but one of the words in the word-initial cohort. This is the “recognition point.”

A word-initial cohort

This idea suggests that our brains process a word in sections, as it is said, gradually identifying it.

gen-	gener-	generous
gendarme	general	
gender	generate	
general	generous	
generate		
generous		
gentleman		
gentian		
gentile		

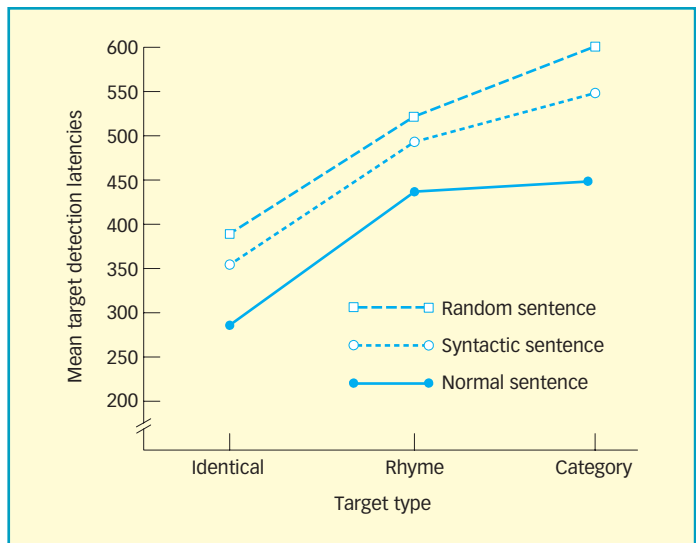
Findings

Support for this model was reported by Marslen-Wilson and Tyler (1980). Listeners had to detect specified targets (a given word; a member of a given category; or a word rhyming with a given word) presented within spoken sentences. The sentences were normal (providing useful semantic and syntactic context), syntactic (providing useful grammatical information only), or random (unrelated words lacking any useful context). Listeners used the contextual information available to reduce the time taken to detect targets, often identifying them before the entire word had been presented (see figure on the right).

O’Rourke and Holcomb (2002) tested the prediction that a spoken word is identified when the point is reached (the recognition point) at which only *one* word is consistent with the acoustic signal. Listeners heard words and nonwords, and decided whether each one was a word. The words were selected so that some had an early recognition point whereas others had a late recognition point. As predicted, listeners realized that a word had been presented shortly after its recognition point had been reached. As a result, listeners worked out faster that a word had been presented when its recognition point was early.

In the original cohort model, it was argued that the processing of the *initial* part of a word was of special importance. Indeed, it was assumed that a spoken word wouldn’t be recognized if its initial phoneme was unclear or ambiguous. However, this assumption is too strong. Accordingly, Marslen-Wilson (1990) revised the cohort model so the word-initial cohort can contain words having *similar* initial phonemes to the presented word rather than being limited to words having exactly the same initial phoneme. This revision is an improvement. For example, Allopenna, Magnuson, and Tanenhaus (1998) found a tendency for listeners to activate words rhyming with the auditory input (e.g., “beaker” activated “speaker”). Thus, some words *not* sharing an initial phoneme with the auditory input are nevertheless included in the word cohort.

According to the original version of the cohort model, context influences word recognition very early in processing. In contrast, the effects of context on word recognition are more limited in the revised version and occur only fairly late in processing. Earlier we saw that the effects of sentence context on the phonemic restoration effect occur late in processing (e.g., Samuel, 1981, 1987). However, sentence context sometimes influences the early stages of word processing. Van Petten et al. (1999) used sentence contexts allowing



Detection times for word targets presented in sentences. Adapted from Marslen-Wilson and Tyler (1980).

listeners to predict the final word before it was presented. For example, guess what the last word is going to be in this sentence: “Sir Lancelot spared the man’s life when he begged for ____.” The last word was “mercy” when it was *congruent* with the sentence context and “mermaid” when it was *incongruent* with the context. Brain activity to congruent and incongruent words differed 200 ms *before* the recognition point was reached. This suggests that context can influence spoken word processing earlier than expected within the revised version of the cohort model.

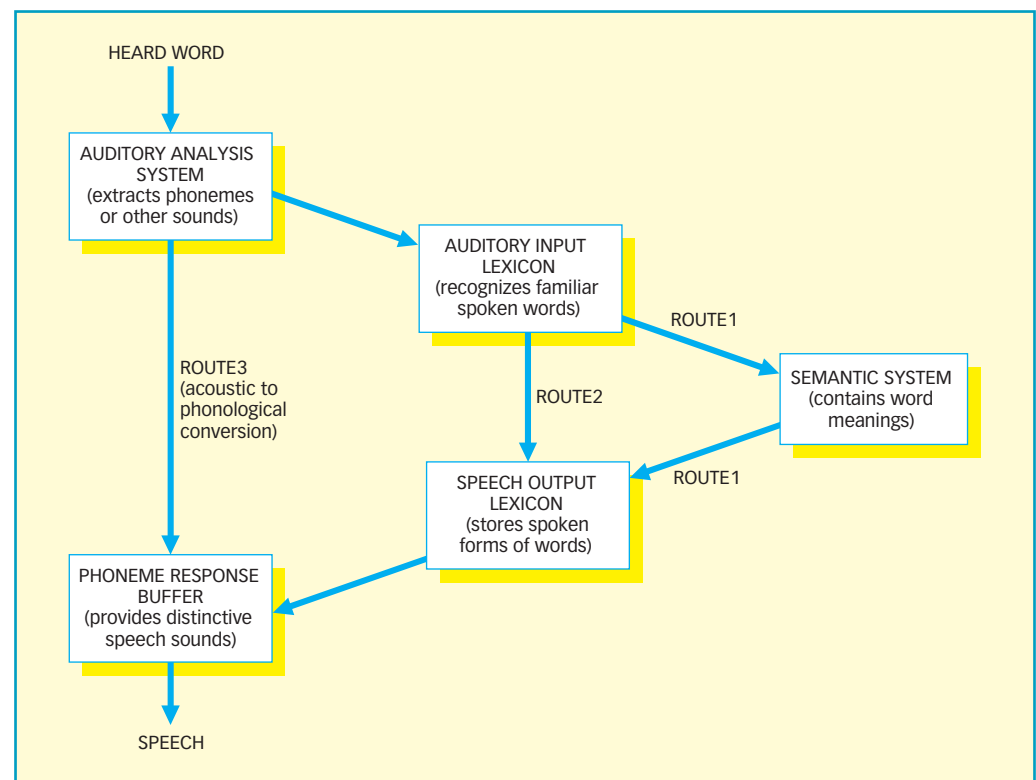
Evaluation

- + The cohort model is probably the most adequate theoretical approach to speech perception.
- + The revised version of the model accurately predicts that most contextual effects on spoken word recognition occur relatively late in processing.
- + The assumption in the revised model that membership of the word-initial cohort is flexible is correct.
- The revised model is less precise than the original version. As Massaro (1994, p. 244) pointed out, “These modifications are necessary to bring the model in line with empirical results, but they . . . make it more difficult to test against alternative models.”
- The task of identifying a word’s starting point from the acoustic signal is complex (remember our earlier discussion of the segmentation problem). However, the cohort model doesn’t make it clear how this is accomplished.

DISORDERS OF SPEECH PERCEPTION

Consider the task of repeating a spoken word immediately after hearing it. In spite of the apparent simplicity of this task, many brain-damaged patients who aren’t deaf experience difficulties with it. These patients show different patterns of impairment, and this has proved useful in identifying the processes involved in repeating spoken words.

Processing and repetition of spoken words. Adapted from Ellis and Young (1988).



The figure on the previous page provides a framework for making sense of the findings from brain-damaged patients. The most striking feature is the notion that saying a spoken word can be achieved using *three* different routes. This is probably two more than most people would have guessed. Routes 1 and 2 are designed to be used with familiar words, whereas Route 3 is for use with unfamiliar words and nonwords. When Route 1 is used, a heard word activates relevant stored information about it, including its meaning and its spoken form. Route 2 closely resembles Route 1 except that information about the meaning of heard words is *not* accessed. Thus, someone using Route 2 can say familiar words accurately, but doesn't have access to their meaning. Finally, Route 3 involves using rules about the conversion of the acoustic information contained in heard words into the appropriate spoken forms of those words (see the figure on the right).

According to the model shown in the figures on the right, there should be brain-damaged patients who use only one or two routes when repeating heard words. Patients using Routes 1 and 2 but not Route 3 would be able to perceive and to understand spoken familiar words. However, they would be impaired at perceiving and repeating unfamiliar words and nonwords. This pattern is found in patients with **auditory phonological agnosia**. For example, a patient (ORF) repeated words much more accurately than nonwords (85% vs. 39%, respectively) (McCarthy & Warrington, 1984).

Patients who can only use Route 2 would be able to repeat familiar words but would often not understand their meaning. In addition, they should have problems in saying unfamiliar words and nonwords, because nonwords can't be processed through Route 2. Patients with **word meaning deafness** fit this description. Dr. O's ability to repeat words was dramatically better than his ability to repeat nonwords (80% vs. 7%, respectively) (Franklin et al., 1996). In addition, his ability to understand the meaning of words (especially abstract ones) was impaired.

In sum, there is support for the three-route model of the processing and repetition of spoken words. However, there would be more support for the model if patients could be found who use mainly Route 3. As yet, no clear cases have been reported in the literature.

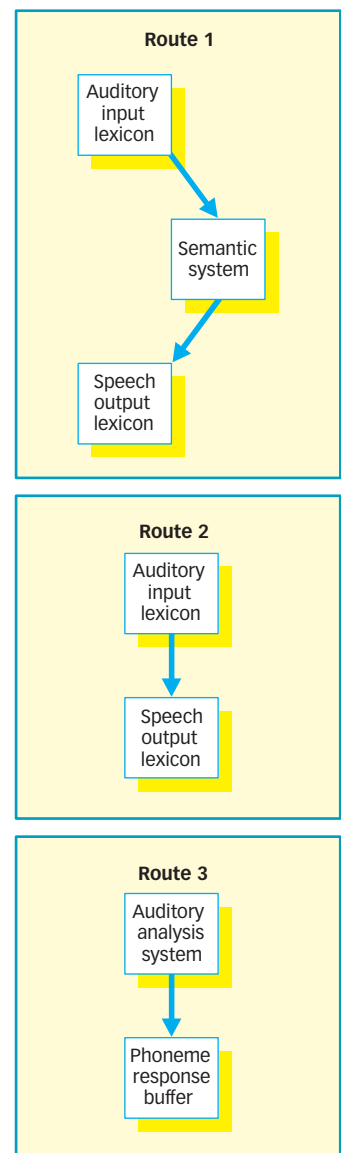
UNDERSTANDING LANGUAGE

The processes involved in understanding or comprehending language are very similar whether we are reading text or listening to speech. As a result, individuals good at understanding written texts are also generally good at understanding spoken language (e.g., Daneman & Carpenter, 1980). However, it is generally easier to understand written text than spoken language. Written text is continuously available and so we can re-read earlier parts of the text if the need arises. In contrast, spoken language is only fleetingly available and we can't listen again to what the speaker said several seconds ago.

How good are we at correctly understanding the information contained in texts and spoken language? Sometimes we aren't as good as we like to think. For example, consider the Moses illusion (Erickson & Mattson, 1981). When asked, "How many animals of each sort did Moses put on the ark," many people reply, "Two." In fact, the correct answer is, "None" (think about it!). Ferreira (2003) argued that people often assume that the subject of a sentence is responsible for some action whereas the object of the sentence is the recipient of the action. That assumption generally works well (e.g., "The mouse ate the cheese"; "The woman visited the man"). However, Ferreira found that the assumption often led to misinterpretations when sentences such as the following were presented auditorily: "The mouse was eaten by the cheese"; "The man was visited by the woman."

INFERENCE DRAWING

What happens when we read texts or listen to speech? The obvious answer is that we focus on the words and the sentences, and try to make sense of them. In terms of our conscious experience, that is certainly what we seem to be doing. In fact, however, language comprehension is more complex. What we read or hear rarely contains *all* the information needed for complete understanding, and so we have to draw inferences to fill



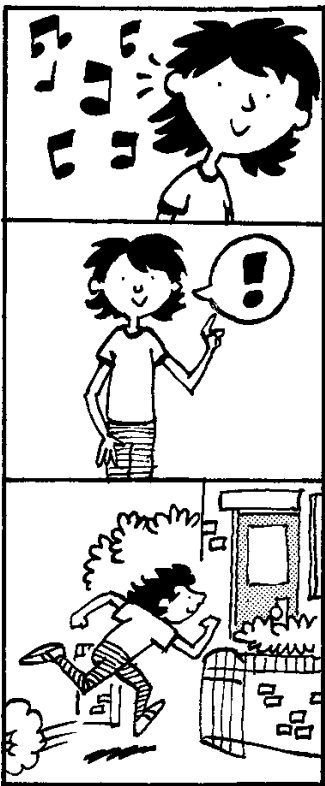
Key Terms

Auditory phonological agnosia:

a condition in which there is poor auditory perception of unfamiliar words and non-words but not of familiar words.

Word meaning deafness:

a condition in which there is selective impairment of the ability to understand spoken (but not written) words.



in the gaps in the information provided. For example, consider the following very short story taken from Rumelhart and Ortony (1977):

1. Mary heard the ice-cream van coming.
2. She remembered the pocket money.
3. She rushed into the house.

You probably made various assumptions or inferences while reading the story. For example, Mary wanted to buy some ice-cream; buying ice-cream costs money; Mary had some pocket money in the house; and Mary had only a limited amount of time to get hold of this money before the ice-cream van arrived. None of these assumptions is explicitly stated in the three sentences. It is simply so common for us to draw inferences that we are generally unaware of doing so.

There has been much controversy on the issue of precisely *which* inferences are drawn when readers try to understand a text. The focus of this controversy has been on **elaborative inferences**, which add details to the text but aren't essential for a proper understanding of what we are listening to or reading. Here is an example of an elaborative inference taken from a study by Singer (1994):

*The tooth was pulled painlessly. The patient liked the new method.
A dentist pulled a tooth (elaborative inference).*

Singer found that readers were relatively slow to draw this elaborative inference, indicating that such inferences are not drawn automatically.

According to the minimalist hypothesis put forward by McKoon and Ratcliff (1992), people generally draw only those inferences that are very important to establish the meaning of what is being read. As a result, elaborative inferences are typically not drawn. In contrast, Graesser, Singer, and Trabasso (1994) argued in their search-after-meaning theory that more inferences are generally drawn than is proposed by the minimalist hypothesis. For convenience, the predictions of the two approaches are shown in the box below.

Key Term

Elaborative inferences: inferences that add details to a text that is being read but that are not essential to understanding.

"Sam had been in pain all night. In the morning the dentist drilled and filled his tooth."

Logical inference: Sam is male.

Bridging inference: The dentist treated Sam's tooth.

Elaborative inference: Sam phoned the dentist early to get an emergency appointment. The treatment cured the pain.

The types of inferences normally drawn, together with the predictions from the search-after-meaning and minimalist perspectives. Adapted from Graesser et al. (1994).					
Type of inference	Answers query	Predicted by search-after-meaning theory	Predicted by minimalists	Normally found	
1. Referential	To what previous word does this apply? (e.g., anaphora)	✓	✓	✓	
2. Case structure role assignment	What is the role (e.g., agent, object) of this noun?	✓	✓	✓	
3. Causal antecedent	What caused this?	✓	✓	✓	
4. Supraordinate goal	What is the main goal?	✓		✓	
5. Thematic	What is the overall theme?	✓		?	
6. Character emotional reaction	How does the character feel?	✓		✓	
7. Causal consequence	What happens next?			×	
8. Instrument	What was used to do this?			×	
9. Subordinate goal-action	How was the action achieved?			×	

The findings favor the search-after-meaning theory over the minimalist hypothesis. For example, Suh and Trabasso (1993) used texts in which a character's initial goal was or wasn't satisfied. In one text, Jimmy wants a bicycle and his mother either buys him one immediately or doesn't buy him one. Later on, Jimmy has earned a lot of money and sets off for a department store. In the condition in which Jimmy hasn't satisfied his goal of having a bicycle, readers inferred that his intention was to buy one in the store. However, the minimalist hypothesis doesn't lead to the prediction that such inferences would be drawn.

Evaluation

- + Readers and listeners often draw inferences when reading text or listening to speech, respectively.
- + There is some validity in the minimalist and search-after-meaning theories. According to Graesser, Millis, and Zwaan (1997, p. 183), "The minimalist hypothesis is probably correct when the reader is very quickly reading the text, when the text lacks global coherence, and when the reader has very little background knowledge. The [search-after-meaning] theory is on the mark when the reader is attempting to comprehend the text for enjoyment or mastery at a more leisurely pace."
- Insufficient attention has been paid to individual differences. For example, individuals of greater intellectual ability draw more inferences than those of lesser ability (see Eysenck & Keane, 2005, for a review).

SCHEMAS

Our stored knowledge and experience play an important role in reading and listening because they allow us to fill in the gaps in what we read or hear. Bartlett (1932) argued that our knowledge and experience can have much more dramatic effects. His central claim was that the knowledge we possess can produce systematic *distortions* in what we remember about what we have read. According to Bartlett, this knowledge is in the form of **schemas** (organized packets of knowledge about the world, events, or people). According to Bartlett, remembering is a process of *reconstruction* in which we use schematic knowledge to organize and make sense of the information in a previously read text. Some schemas are in the form of **scripts**, which contain organized information about typical events (e.g., going to a restaurant).

Bartlett (1932) had the ingenious idea of presenting people with stories involving *conflict* between the information contained in the story and their prior schematic knowledge. His participants were mostly students from Cambridge University in the UK and the stories (e.g., "The War of the Ghosts") were often tales from the very different North American Indian folk culture. He argued that the students would read each folk tale from the standpoint of their own cultural background. Bartlett's findings supported his predictions, with most of the recall errors being in the direction of making the story read more like a conventional English story. He used the term **rationalization** for this type of error. He also found other kinds of errors, including flattening (failure to recall unfamiliar details) and sharpening (elaboration of certain details).

Bartlett (1932) made a further prediction. He argued that memory for the precise information contained in the story would be forgotten over time, whereas participants' relevant schematic knowledge would not. As a result, rationalization errors (dependent on participants' knowledge) should increase at longer retention intervals. Bartlett's findings supported this prediction.

Bartlett's (1932) experimental approach was hardly ideal. The instructions he gave his participants were deliberately vague, and he practically never used any statistical tests on his data! More worryingly, many of the recall distortions he obtained were a result of conscious guessing rather than genuine problems in memory. Perhaps the participants were trying to be helpful by guessing when they couldn't remember some aspect of the story. Convincing evidence that this is a problem was produced by Gault and Stephenson

Key Terms

Schemas:

in Bartlett's theory, organized information about the world, events, or people stored in long-term memory.

Scripts:

organized information or schemas representing typical events.

Rationalization:

in Bartlett's theory, the tendency to recall stories in distorted ways influenced by the reader's cultural expectations and **schemas**.

Key Study

Bartlett (1932): The War of the Ghosts

In one of his studies Bartlett asked his English participants to read a North American Indian folk tale called “The War of the Ghosts,” after which they tried to recall the story. Part of the story was as follows:

One night two young men from Edulac went down the river to hunt seals, and while they were there it became foggy and calm. Then they heard war-cries, and they thought: “Maybe this is a war-party.” They escaped to the shore, and hid behind a log. Now canoes came up, and they heard the noise of paddles, and saw one canoe coming up to them. There were five men in the canoe, and they said: “What do you think? We wish to take you along. We are going up the river to make war on the people.”

... one of the young men went but the other returned home...[it turns out that the five men in the boat were ghosts and after accompanying them in a fight, the young man returned to his village to tell his tale]...and said: “Behold I accompanied the ghosts, and we went to fight. Many of our fellows were killed, and many of those who attacked us were killed. They said I was hit, and I did not feel sick.”

He told it all and then he became quiet. When the sun rose he fell down. Something black came out of his mouth. His face became contorted...He was dead. (p. 65)

One of the subject’s recall of the story (two weeks later):

There were two ghosts. They were on a river. There was a canoe on the river with five men in it. There occurred a war of ghosts...They started the war and several were wounded and some killed. One ghost was wounded but did not feel sick. He went back to the village in the canoe. The next morning he was sick and something black came out of his mouth, and they cried: “He is dead.” (p. 76)

Bartlett found that the participants’ recall distorted the content and style of the original story. The story was shortened, and the phrases, and often words, were changed to be similar to the English language and concepts (e.g., “boat” instead of “canoe”). He also found other kinds of errors, including flattening (failure to recall unfamiliar details) and sharpening (elaboration of certain details).

Discussion points

1. Why do you think that Bartlett’s research has been so influential?
2. Do you think that the kinds of errors and distortions observed by Bartlett would be found with other kinds of material?

KEY STUDY EVALUATION

Bartlett’s research is important because it provided some of the first evidence that what we remember depends in an important way on our prior knowledge, in the form of schemas.

However, Bartlett’s studies are open to criticism. He did not give very specific instructions to his participants (Bartlett, 1932, p. 78: “I thought it best, for the purposes of these experiments, to try to influence the subjects’ procedure as little as possible.”) As a result, some distortions observed by Bartlett were a result of conscious guessing rather than deficient memory. Gauld and Stephenson (1967) found that instructions emphasizing the need for accurate recall eliminated almost half the errors usually obtained.

Another criticism of Bartlett's work was that his approach to research lacked objectivity. Some psychologists believe that well-controlled experiments are the only way to produce objective data. Bartlett's methods were somewhat casual. He simply asked his group of participants to recall the story at various intervals and there were no special conditions for this recall. It is possible that other factors affected their performance, such as the conditions around them at the time they were recalling the story, or it could be that the distortions were simply guesses by participants who were trying to make their recall seem coherent and complete rather than genuine distortions in recall.

On the other hand, one could argue that his research is more ecologically valid than those studies that involve the recall of syllables or lists of words. In recent years there has been an increase in the kind of research conducted by Bartlett, looking more at "everyday memory."

(1967). Clear instructions emphasizing the need for accurate recall (and so reducing deliberate guessing) eliminated many of the errors obtained using Bartlett's vague instructions. However, it should be noted that Bartlett's key findings have been replicated in well-controlled studies (e.g., Sulin & Dooling, 1974).

It is generally assumed that schema- or script-based information can be distinguished from information about individual words or concepts. If so, it seems reasonable to predict that some brain-damaged individuals will have greater problems with the former than the latter. Support for this prediction was reported by Cosentino et al. (2006), who studied patients with fronto-temporal dementia who showed attentional deficits and poor executive functioning. These patients and healthy controls were presented with various scripts. Some of these scripts contained sequencing errors (e.g., dropping fish in a bucket occurring *before* casting the fishing line) whereas others contained semantic or meaning errors (e.g., placing a flower on the hook in a story about fishing). Healthy controls detected as many sequencing errors as semantic ones, but the patients failed to detect almost twice as many sequencing errors as semantic ones. It seems that these patients had relatively intact semantic knowledge of concepts but fairly severe impairment of script-based knowledge.

Do schemas influence the process of *comprehension* when people are reading a text or do they influence the subsequent processes of *retrieval*? Bartlett (1932) argued that schemas influence retrieval, but the evidence indicates that schemas influence comprehension as well as retrieval. For example, Bransford and Johnson (1972) presented a passage of which this is the start:

The procedure is quite simple. First, you arrange items into different groups. Of course one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step; otherwise, you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run this may not seem important but complications can easily arise.

What on earth was that all about? Participants hearing the passage in the absence of a title rated it as incomprehensible and recalled very little of it. In contrast, those supplied beforehand with the title, "Washing clothes," found it easy to understand and recalled twice as much. The title indicated the nature of the underlying schema and helped comprehension of the passage rather than retrieval. We know this because participants receiving the title *after* hearing the passage (but *before* recall) had as poor recall as those never receiving the title.

Evidence that schemas influence retrieval processes was reported by Anderson and Pichert (1978). Participants initially recalled a story from the perspective of a burglar or of someone interested in buying a home. After they had recalled as much as possible, they shifted to the alternative perspective (from burglar to home buyer or vice versa) before recalling the story again. Accessing different schematic knowledge in this way at the time of retrieval led to enhanced recall.

Evaluation

- + We use schematic knowledge during comprehension (e.g., Bransford & Johnson, 1972) and retrieval (e.g., Anderson & Pichert, 1978).
- + The use of schematic knowledge creates errors and distortions when there is a conflict between that knowledge and information in the text.
- + There is evidence from brain-damaged patients supporting a distinction between knowledge of concepts and schema- or script-based knowledge (e.g., Cosentino et al., 2006).
- It has proved difficult to establish clearly the characteristics of schemas.
- As Harley (2001, p. 331) argued, “The primary accusation against schema . . . approaches is that they are nothing more than re-descriptions of the data.”

CONSTRUCTION-INTEGRATION MODEL

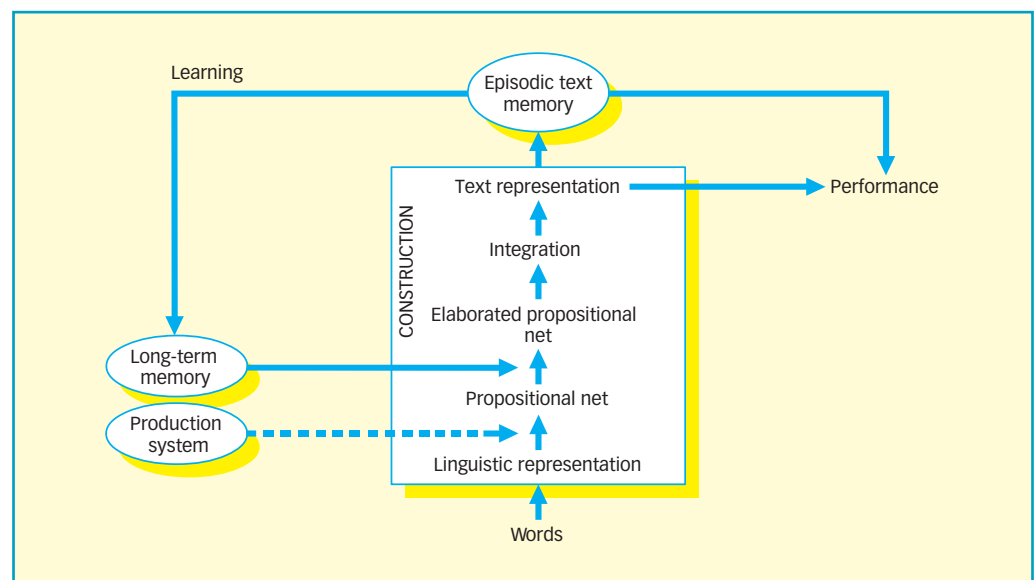
The construction–integration model (Kintsch, 1988, 1992, 1998) is the leading theoretical approach to language comprehension (see the figure below). Here are its main assumptions:

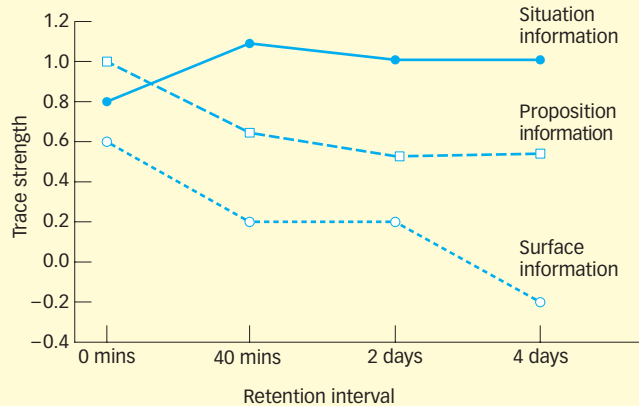
1. The sentences in the text are turned into propositions (a **proposition** is a statement making an assertion or denial; it may be true or false).
2. *Construction process*: Readers use their knowledge base to retrieve propositions relevant to the text propositions (e.g., inferences). These propositions and text propositions together form an *elaborated propositional net*. This seems inefficient, because many propositions in this net are of only marginal relevance to the text’s theme.
3. *Integration process*: Contextual information provided by the previous parts of the text is used to select from the elaborated propositional net only those propositions most relevant to the theme of the text.
4. *Three levels of text representation are constructed*:
 - (i) Surface representation (the text itself); this representation lasts for the shortest time.
 - (ii) Propositional representation (propositions formed from the text).
 - (iii) Situational representation (a mental model describing the situation referred to in the text). Schemas (packets of knowledge) can be used as the building blocks for the construction of situational representations or models. This representation lasts for the longest time.

Key Term

Proposition:
a statement that makes an assertion or denial and which can be true or false.

The construction–integration model. Adapted from Kintsch (1994).





Forgetting functions for situation, proposition, and surface information over a 4-day period. Adapted from Kintsch et al. (1990).

In sum, the initial construction process uses semantic knowledge to form a large number of propositions. Then the integration process uses contextual information from the text to weed out those propositions of little relevance while retaining those most related to the theme of the text.

Findings

A striking assumption of the construction–integration model is that no fewer than *three* representations are formed when we read a text. This assumption was tested by Kintsch et al. (1990). Participants were given brief descriptions of very stereotyped situations (e.g., going to see a movie) and then their recognition memory was tested immediately or at times ranging up to 4 days later.

The forgetting functions for the surface, propositional, and situational representations differed considerably (see the figure above). There was rapid and complete forgetting of the surface representation, whereas information from the situational representation showed no forgetting over 4 days. Propositional information differed from situational information in that there was only partial forgetting. Thus, the findings were as predicted by the model.

Subsequent research has indicated that readers don't *always* construct all three forms of text representation. In a study by Zwaan and van Oostendorp (1993), participants read an edited mystery novel describing the details of a murder scene, including the locations of the body and various clues. Most readers did *not* construct a situational representation when they read normally, something of which a legion of detective story writers since Agatha Christie have taken advantage. However, situational representations were formed (at the cost of a substantial increase in reading time) when the initial instructions emphasized the importance of constructing a spatial representation. Thus, limited processing capacity may often restrict the formation of situational representations.

Kaup and Zwaan (2003) found that there are circumstances in which readers don't even form propositional representations. They pointed out that nearly all previous research involved the use of texts in which all the objects referred to each sentence were present. According to construction–integration theory, propositional representations should contain information in which *all* the objects referred to in each sentence even if they were absent. However, Kaup and Zwaan wondered whether people would bother to construct propositional representations referring to absent objects. Accordingly, they presented stories containing sentences such as, “Sam was relieved that Laura was not wearing her pink dress.” Information about absent objects was less accessible than that about present objects only 1.5 seconds after a sentence had been read. This suggests that

information about absent objects was generally not included in the propositional representation.

According to the model, it is only late on in language comprehension that contextual information from the text is used. However, Cook and Myers (2004) found that contextual information can be used at a very early stage of processing. For example, suppose you read a short story about the making of a movie. The story contains the contextual information that the budget was so low that everyone had to take on extra jobs. When this information was provided, readers very rapidly understood why it was the actress rather than the director who said, “Action!”

Evaluation

- + There is reasonable support for the notion that text comprehension involves successive stages of construction and integration.
- + There is evidence for the three levels of representation (surface; propositional; and situational) specified in the model.
- The three representations are not always formed. Sometimes the situational representation isn’t constructed (Zwaan & van Oostendorp, 1993) and sometimes the propositional representation isn’t formed (Kaup & Zwaan, 2003).
- Contextual information can be used much earlier in language comprehension than assumed by the model.
- There may be other levels of representation ignored by Kintsch. For example, Graesser, Millis, and Zwaan (1997) argued there is the *text genre* level, which is concerned with the nature of the text (e.g., narration; description; jokes). How the information is to be interpreted varies greatly as a function of the type or genre of text.



Sports commentators use preformulation extensively. Repeating stock phrases and jargon enables them to speak quickly about the sporting event as it happens.

Key Terms

Preformulation:

this is used in speech production to reduce processing costs by using phrases often used previously.

Underspecification:

a strategy used to reduce processing costs in speech production by producing simplified expressions.

SPEECH PRODUCTION

Speech production seems almost effortless as we chat with friends or acquaintances about the topics of the day. Indeed, we often seem to speak without much preparation or planning, which can cause us to say things we immediately regret saying! Speech production is sufficiently easy that we typically speak at three words a second or almost 200 words a minute. That simply wouldn’t be possible if we had to think everything out carefully in advance.

The impression created in the previous paragraph is deliberately misleading. Speech production is actually much more complex than might be imagined. As M. Smith (2000) pointed out, most of us resort to various strategies when talking to reduce the processing demands on us. Here are three of those strategies. First, there is **preformulation**, which involves reducing processing costs by producing phrases used before. About 70% of our speech consists of word combinations we use repeatedly (Altenberg, 1990). Sports commentators (who often need to speak very rapidly) make extensive use of preformulation (e.g., “They’re on their way”; “They’re off and racing now”) (Kuiper, 1996).

Second, there is **underspecification**, which involves using simplified expressions. Smith (2000) illustrated

underspecification as follows: “Wash and core six cooking apples. Put them in an oven.” In the second sentence, the word “them” underspecifies the phrase, “six cooking apples.”

Third, when we have a chat with someone, we often use the strategy of copying phrases and even sentences produced by the other person (Pickering & Garrod, 2004). In addition, we typically make extensive use of the ideas communicated by the other person. Both of these processing strategies reduce the processing burden associated with speech production.

SPEECH PLANNING

How much do we plan what we are going to say before saying it? The answer varies from person to person. You probably know some motormouths who don’t seem to plan at all. In fact, nearly all speakers plan what they are about to say to some extent, but there is controversy concerning the extent of such planning.

One possibility is that speech planning extends over an entire **clause**, a part of a sentence containing a subject and a verb. There is support for this view from speech errors (e.g., Garrett, 1982). For example, word-exchange errors involve two words exchanging places. The words exchanged often come from different phrases but belong to the same clause (e.g., “My chair seems empty without my room”). The fact that word-exchange errors rarely involve words from different clauses strengthens the argument that the clause is a key unit in speech planning. Additional evidence was reported by Holmes (1988). Speakers often had hesitations and pauses before the start of a clause, suggesting they were planning the forthcoming clause.

Other evidence suggests that speech planning is at the level of the **phrase**, a group of words expressing a single idea and smaller in scope than a clause. Martin, Miller, and Vu (2004) asked people to produce sentences to describe moving pictures. The sentences either had a simple initial phrase (e.g., “The ball moves above the tree and the finger”) or they had a complex initial phrase (e.g., “The ball and the tree move above the finger”). Speakers took longer to initiate speech when using complex initial phrases than when using simpler ones. Thus, participants planned the initial phrase before starting to speak.

What probably happens is that the amount of planning that precedes speech is *flexible*, varying as a function of the speaker’s skills and situational demands. We can draw an analogy with diving from the high board in a swimming pool. If time permits, there will be careful planning before the diver launches himself/herself into the dive. However, a skilled diver can dive reasonably well with little or no planning.

Evidence that the diving analogy applies to speech production was reported by Ferreira and Swets (2002). Participants answered mathematical problems varying in difficulty level and spoke their answers. In their first experiment, task difficulty affected the time taken to start speaking but not the time spent speaking. These findings suggest that participants fully planned their responses before speaking. However, the findings differed in a second experiment. In this experiment, participants had to start producing their answers very rapidly for them to count. What happened here was that some planning occurred before participants started to speak, with additional planning occurring while they were speaking. Thus, speakers are flexible. They do only as much planning as is feasible in the time available before they need to start talking.

SPEECH PRODUCTION LEVELS

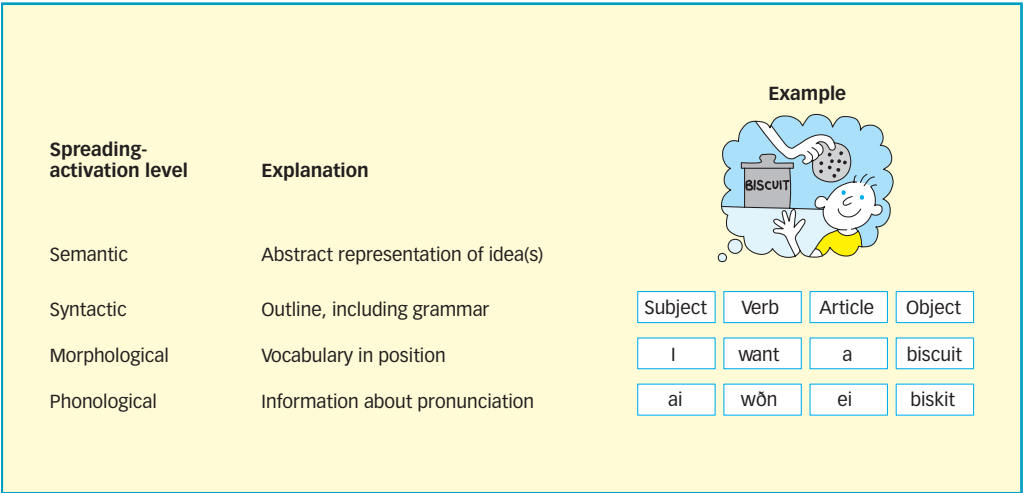
Most theorists assume that several levels or stages of processing are involved in speech production. So far we have focused on the initial level or stage of processing during which the meaning of what the speaker wants to say is worked out. The main levels or stages were identified by Dell (1986; see the figure overleaf). It is assumed that speakers start with an overall idea of what they want to

Key Terms

Clause:
part of a sentence that contains a subject and a verb.

Phrase:
a group of words expressing a single idea; it is shorter than a clause.

The sentence “I want a biscuit” broken down into spreading-activation levels identified by Dell (1986).



say, and this gradually becomes more detailed and specific through successive stages of processing.

Initial planning of the message to be communicated is considered at the semantic level. At the syntactic level, the grammatical structure of the words in the planned utterance is decided. At the morphological level, the **morphemes** (basic units of meaning or word forms) in the planned sentence are worked out. At the phonological level, the *phonemes* or basic units of sound within the sentence are added.

The figure above implies that the processes involved in speech and word production proceed in an orderly way from the semantic level through the syntactic and morphological levels down to the phonological level. There has been much disagreement concerning the extent to which the processes involved do actually occur in this neat sequential way (discussed later).

There are several theories of speech production. However, we will focus on the spreading-activation theory put forward by Dell (1986) and by Dell and O’Seaghdha (1991). Let’s start with the key notion of **spreading activation**. It is assumed that the nodes within a network (many corresponding to words) vary in their activation or energy. When a node or word is activated, activation or energy spreads from it to other related nodes. For example, strong activation of the node corresponding to “tree” may cause some activation of the nodes corresponding to “plant” and to “oak.”

A key theoretical assumption is that processing can occur at the same time at all four levels (semantic; syntactic; morphological; and phonological). However, it is typically more advanced at the semantic and syntactic levels than the morphological and phonological ones. Thus, the system operates flexibly. It also operates in parallel, meaning that several processes can all occur together. What is the advantage of having such a speech production system? According to Dell, this flexibility makes it easier for us to produce novel sentences in our speech.

The notion that speech production involves several processes at the same time may sound like a recipe for chaos. However, two kinds of rules help to prevent that from happening:

Key Terms

Morphemes:
the smallest units of meaning within words.

Spreading activation:
the notion that activation or energy spreads from an activated node (e.g., word) to other related nodes.

1. *Categorical rules:* These rules impose constraints on the categories of items produced at each level. As a result, selected items will belong to the appropriate category (e.g., noun or verb at the syntactic level).
2. *Insertion rules:* These rules select the items to be included in the representations at each level according to the following criterion: the most highly activated node or word belonging to the appropriate category (e.g., noun) is selected.

After an item has been selected, its activation level immediately reduces to zero. This ensures it will not be selected repeatedly.

As we will see, much of the support for spreading-activation theory comes from the study of speech errors. Most of this evidence comes from errors personally heard by the researcher concerned (e.g., Stemberger, 1982). You might imagine this would produce distorted data, since some errors are easier to detect than others. However, the types and proportions of speech errors obtained in this way are very similar to those obtained from analyzing tape-recorded conversations (Garnham, Oakhill, & Johnson-Laird, 1982).

Findings

According to the theory, speech errors will occur whenever an incorrect item is more activated than the correct one. Several findings from speech errors support the theory. First, the existence of categorical rules means that errors should typically belong to the appropriate category (e.g., an incorrect noun replacing the correct one). Relevant evidence comes from *semantic substitution errors*, in which the correct word is replaced by a word of similar meaning. For example, someone might say, “Where is my tennis bat?” instead of “Where is my tennis racket?” In 99% of cases, the substituted word is of the same form class as the correct word (e.g., nouns substitute for nouns) (Hotopf, 1980).

Second, the theory assumes that many sounds and words are all activated at the same time. As a result, there will be anticipation errors, in which a sound or word is spoken earlier in the sentence than intended. Anticipation errors are found with *word-exchange errors*, in which two words in a sentence switch places. For example, someone says, “I must let the house out of the cat” instead of “I must let the cat out of the house.” *Sound-exchange errors* are another type of anticipation error. They involve two sounds exchanging places (e.g., “barn door” instead of “darn bore”).

Third, the theory accounts for the **mixed error effect**. This happens when an incorrect spoken word is semantically *and* phonemically related to the correct word. Dell (1986) quoted the example of someone saying, “Let’s stop” instead of, “Let’s start.” The spoken word and the correct word are generally both more similar in sound (as well as meaning) than would be expected by chance alone (Dell & O’Seaghdha, 1991). According to the theory, activation often occurs at the semantic and phonological levels at the same time, and this helps to produce the mixed error effect.

There is a major issue on which the predictions from spreading-activation theory differ from those of the influential WEAVER++ model (Levelt, Roelofs, & Meyer, 1999). According to this model, the processing involved in speech production is typically neat and tidy. For example, we decide on the meaning of a word we want to use in a spoken sentence before we access phonological (sound) information about the word in question. In other words, the process of deciding on word meaning is complete before the sound of the word is accessed. Support for this view comes from the **tip-of-the-tongue phenomenon**, in which we have a concept or idea in mind but search in vain for the right word to describe it. People in the tip-of-the-tongue state “appear to be in a mild torment, something like the brink of a sneeze” (Brown & McNeill, 1966, p. 325).

It follows from WEAVER++ that we can only access the sound of *one* word at a time. In contrast, it is assumed within spreading-activation theory that processing is flexible, and so phonological or sound processing of two or more words can occur at the same time. Morsella and Miozzo (2002) compared the two theories. Participants were presented with two colored pictures with one superimposed on the other. They had to name the picture in a given color (target picture) while ignoring the picture in a different color (distractor picture). Some distractor pictures were related in sound to the target pictures (e.g., *bell* as a distractor presented with *bed*), but others weren’t related in sound (e.g., *bat* as a distractor presented with *bed*). Target pictures were named faster when accompanied by phonologically related distractors rather than by unrelated ones. This means that the sounds of both words were activated at the same time, precisely as predicted by spreading-activation theory.

Key Terms

Mixed error effect:

speech errors that are semantically and phonologically related to the intended word.

Tip-of-the-tongue phenomenon:

the experience of having a specific concept in mind but being unable to access the correct word to describe it.

According to spreading-activation theory, speech production often involves parallel processing in which irrelevant words are activated as well as relevant ones. That suggests that there should be numerous errors when incorrect words are readily available. This prediction was tested by Glaser (1992), who studied the time taken to name pictures (e.g., a table). Theoretically, there should have been a large increase in errors when each picture was accompanied by a semantically related distractor word (e.g., chair). In fact, there was only a modest increase in the error rate. Thus, there seems to be more limited processing of distractor words than expected on Dell's theory.

Evaluation

- + The four levels assumed to be involved in speech production are consistent with the evidence.
- + The notion that speech production involves flexible, parallel processing is basically correct.
- + The theory makes precise (and accurate) predictions about the kinds of errors occurring most often in speech production.
- The theory focuses on individual words rather than on broader issues relating to message construction.
- The theory predicts the nature and number of errors produced in speech. However, it can't predict the *time* taken to produce spoken words, because there are no theoretical assumptions about the timing of speech processes.
- The theory seems to predict too many errors when irrelevant words are activated at the same time as relevant ones (e.g., Glaser, 1992).

SPEECH DISORDERS

As we have seen in several chapters of this book, we can often learn much about human cognition by studying brain-damaged patients. It has been claimed that some speech-disordered patients speak reasonably grammatically but have great difficulties in accessing content words (e.g., nouns; verbs), whereas other patients have the opposite pattern. If that is the case, it would provide some support for those theories (e.g., spreading-activation theory; WEAVER++) assuming that speech production involves separate stages of syntactic processing and word finding.

Patients suffering from **jargon aphasia** apparently speak fairly grammatically but cannot find content words. They substitute one word for another, and often produce **neologisms** (made-up words). Ellis, Miller, and Sin (1983) studied a jargon aphasic, RD. Here is his description of a picture of a scout camp with the words he seemed to be looking for in brackets: "A b-boy is swi'ing (SWINGING) on the bank with his hand (FEET) in the stringt (STREAM). A table with orstrum (SAUCEPAN?) and . . . I don't know . . . and a three-legged stroe (STOOL) and a strane (PAIL)—table, table . . . near the water." RD, in common with most jargon aphasics, produced more neologisms or invented words when the word he wanted wasn't a common one.

Robson, Pring, Marshall, and Chiat (2003) studied a jargon aphasic, LT, whose speech consisted almost entirely of neologisms. He often produced consonants common in the English language regardless of whether or not they were correct. This finding can be explained on Dell's spreading-activation theory (discussed earlier). In essence, the resting activation level of frequently used consonants is greater than that of rarely used consonants. This increases the probability of producing frequently used consonants correctly and incorrectly.

Key Terms

Jargon aphasia:

a brain-damaged condition in which speech is reasonably correct grammatically, but there are great problems in finding the right words.

Neologisms:

made-up words that are often found in the speech of patients with **jargon aphasia**.

Case Study: Expressive Language Disorder

Three percent to 5% of children have expressive language disorder, which develops at about age 4. It produces ongoing difficulties in understanding and using language, for example having for their developmental stage a markedly limited vocabulary, making errors in tense, having difficulty recalling words or producing sentences of appropriate length and complexity. Sufferers from this disorder show problems with educational or occupational achievement, and social difficulties.

A 7-year-old girl with the disorder was described at school as not an independent worker but this was because she did not have the skills needed to work independently. She found following instructions, e.g., in science lessons, difficult, as was the writing of stories, and mathematics. Extra help such as discussing the use of silent (i.e., sub-vocal) speech is a technique that helps some young people with the disorder.

Can you explain how the difficulties of this disorder contributed to this girl's reduced educational achievement?

Patients with **agrammatism** can generally find the right words but can't order them grammatically. They typically produce short sentences containing content words (e.g., nouns; verbs) but lacking function words (e.g., the; in; and) and word endings. In addition, it has often been assumed that patients with agrammatism have problems in understanding syntactically complex sentences. Guasti and Luzzati (2002) found that agrammatic patients had very impaired syntactic processing especially as revealed in their inappropriate use of verbs. They often failed to adjust the form of verbs to take account of person or number. Most of the time they used only the present tense of verbs, and omitted many verbs altogether.

In sum, research on patients with jargon aphasia and agrammatism has provided general support for the notion that there are separate stages of syntactic processing and word finding in speech production. However, the pattern of intact and impaired language abilities in any given patient is typically rather complex. Second, and related to the first point, it is a serious oversimplification to assume that all patients assigned to the same category (e.g., agrammatism) have the same impairments. For example, Berndt, Mitchum, and Haendiges (1996) discussed the findings from several studies on comprehension of active and passive sentences by patients with agrammatism. Some patients performed at chance level on both kinds of sentences, whereas other patients had reasonable comprehension performance on active and passive sentences, especially active ones.

Chapter Summary

Reading

- Readers fixate 80% of content words and 20% of function words, with nonfixated words tending to be common, short, and predictable.
- Reading is more likely to involve phonological processing when reading is difficult.
- There are two main routes to reading words aloud. One is a lexical look-up route based on a mental dictionary and the other is a nonlexical route based on converting spelling into sound via grapheme–phoneme conversion.
- Patients with surface dyslexia mainly use the nonlexical route, whereas those with phonological dyslexia mainly use the lexical route.
- According to two-stage models of parsing, syntactic information is used in the first stage and semantic information is used in the second stage.

Key Term

Agrammatism: a condition in brain-damaged patients in which speech lacks grammatical structure and many function words and word endings are omitted.

- According to one-stage models of parsing, all relevant information is available immediately to the parser. There is little clear evidence favoring one type of model over the other.

Speech perception

- Listeners have to cope with various problems including the segmentation problem, co-articulation, and the fact that speakers typically produce 10 phonemes per second.
- Listeners use contextual information from the surrounding sentence context when identifying individual words. This is shown in the phonemic restoration effect, in which listeners guess the identity of a missing phoneme.
- According to the cohort model, processing of a spoken word continues until one word is consistent with all of the available information. Contextual information is used mainly at a fairly late stage of processing.
- Saying a spoken word can be achieved using three different routes. Two of the routes are used with familiar words and the other route is used with unfamiliar words and nonwords.
- Support for the three-route model comes from patients with auditory phonological agnosia who repeat words more accurately than nonwords. It also comes from patients with word meaning deafness who don't understand the meaning of words they repeat and who find it harder to repeat nonwords than words.

Understanding language

- Readers draw inferences when reading texts or listening to speech.
- According to the minimalist hypothesis, fewer nonessential inferences are drawn than is assumed within search-after-meaning theory. The evidence generally favors search-after-meaning theory, but the number of inferences drawn depends on the reader's goals.
- Recall of text is systematically distorted when there is a conflict between readers' schematic knowledge and the information in the text.
- Schemas influence processes occurring during comprehension as well as those occurring during retrieval.
- According to the construction-integration model, a construction process is used to retrieve propositions relevant to the text. The integration process then selects those propositions of most relevance to the theme of the text.
- According to the construction-integration model, surface, propositional, and situational representations are formed. In fact, readers often don't construct all three representations.

Speech production

- Speakers use preformulation, underspecification, and copying what someone else has just said to simplify the task of speech production.
- The amount of planning that precedes speech varies as a function of the speaker's skills and the demands of the situation.
- Speech production involves four levels of processing: semantic; syntactic; morphological; and phonological.
- According to Dell's spreading-activation theory, processes at different levels of speech production take place in parallel. The assumption that whichever word is most activated will be produced accounts for many speech errors. However, the theory doesn't consider broader issues relating to message construction.
- Patients with jargon aphasia speak fairly grammatically but can't find content words, whereas those with agrammatism can generally find the correct words but can't order them grammatically. These patterns of impairments are consistent with the notion that there are separate stages of syntactic processing and word finding.

Further Reading

- Butcher, K.R., & Kintsch, W. (2003). Text comprehension and discourse processing. In R.W. Proctor & A.F. Healy (Eds.), *Handbook of psychology: Experimental psychology* (Vol. 4). New York: Wiley. This chapter discusses what is known about language comprehension.
- Eysenck, M.W. (2006). *Fundamentals of cognition*. Hove, UK: Psychology Press. Theory and research on language are discussed in Chapters 17–21 of this introduction to cognitive psychology.
- Harley, T.A. (2008). *The psychology of language: From data to theory* (3rd ed.). Hove, UK: Psychology Press. Several chapters in this excellent textbook provide clear accounts of the topics discussed in this chapter.
- Jay, T.B. (2003). *The psychology of language*. Upper Saddle River, NJ: Prentice Hall. This textbook has good basic coverage of research and theory on language.



- **Intelligence** *pp.267–285*

Introduction

Intelligence testing

Theories of intelligence

Heredity and environment

Size does matter!

- **What does personality look like?** *pp.287–307*

Introduction

Theory of psychosexual development

Personality assessment

Trait approach to personality

Social cognitive theory

INTRODUCTION TO Individual Differences

Most of psychology is devoted to the search for broad generalizations and laws of behavior that are applicable to nearly everyone. For example, consider most of the research on cognitive psychology (covered in Chapters 6–10). Most cognitive psychologists assume that everyone (apart from brain-damaged patients) makes use of the same attentional and perceptual processes, has a working memory system, forgets information over time, makes certain inferences when reading text, and so on. The situation is similar in social psychology (Chapters 17–20). It is assumed that people tend to be obedient to authority, to conform to the views of other group members, to fall in love with people who are similar to them, to show prejudice in certain circumstances, and so on.

The above approach has proved very successful, and numerous very important generalizations have emerged as a result. However, an approach that focuses on universally applicable generalizations is necessarily limited. What it misses is the extraordinary diversity of human behavior, which is obvious to us nearly all the time in the course of our everyday lives. For example, some people are much better than others at remembering information over long periods of time, some people can control their attention much better than others, some people are more willing than others to obey the orders of authority figures, and some people are bigoted and biased in their views of minority groups whereas others are not.

Individual differences (which form the subject matter of Chapters 11 and 12) have a great influence on our behavior. For example, if you need advice when writing a coursework essay, you probably find it easier to approach some people rather than others, perhaps because they are friendlier or more knowledgeable. In similar fashion, we take full account of individual differences when choosing friends. Thus, we are more likely to become friends with other people who seem similar to us, who are warm and sociable, and who can be trusted with secrets.

Cronbach (1957) argued that there are two scientific disciplines within psychology. One is devoted to the search for general laws, and the other is devoted to the study of individual differences. Cronbach's most important point was that what was needed for psychology to realize its potential were wholehearted attempts to combine these two scientific disciplines. According to him, it should be possible to consider general laws and individual differences within a single approach. It is disappointing (and surprising) that relatively little has been done along the lines suggested by Cronbach over 50 years ago although there are recent very encouraging signs.



Human behavior is extraordinarily diverse, and so approaches that do not take into account individual differences but rely solely on universally applicable generalizations are limited.

ASPECTS OF INDIVIDUAL DIFFERENCES

Which aspects of individual differences are most important? This is not an easy question to answer given that people differ from each other in a bewildering number of ways. Indeed, if you sat down to draw up a list, you could probably identify literally dozens of interesting ways in which the people you know differ from each other. However, academic and occupational psychologists have (rightly or wrongly) focused mainly on individual differences in intelligence and personality, although some research on other aspects of individual differences (e.g., social attitudes) has also been carried out.

There has been much emphasis on individual differences within occupational psychology. It is common for the intelligence and personality of job applicants to be assessed in personnel selection. There are sound reasons for focusing on these aspects of individual differences. For example, highly intelligent individuals on average have superior work performance and career development than those who are less intelligent, and this is especially the case when the job concerned is relatively complex (Gottfredson, 1997). This is one of the many issues discussed in Chapter 11.

You may feel that it is fairly obvious that intelligence predicts the ability to perform complex tasks and jobs. However, individual differences in intelligence are also relevant to many other life outcomes. Individuals who are highly intelligent have a much smaller probability than those who are unintelligent of being divorced within 5 years of marriage (9% vs. 21% based on American data; Gottfredson, 1997). Highly intelligent women have only one-quarter the probability of unintelligent women of having an illegitimate child and one-seventh the probability of finding themselves in prison (Gottfredson, 1997).

Individual differences in personality are also important in predicting individuals' behavior in numerous real-world situations. The personality we have helps to determine how happy we are and how many friends we have. In addition, there are interesting associations between certain types of personality on the one hand and various mental disorders on the other hand. These associations suggest that your personality influences the probability that you will develop any given mental disorder.

When we think of individual differences, we generally focus on personality. We know that Kate is always cheerful and friendly, whereas Nancy is neurotic and emotional. Why are their personalities so different? Did Kate have a happier and more secure childhood than Nancy? Do the differences lie in the genes? Is some kind of mixture or interaction of genetic factors and environmental ones responsible for the personality differences between Kate and Nancy? Tentative answers to these questions are provided in Chapter 12.

Three issues are of central importance to researchers who study individual differences. First, we need to understand the *nature* of individual differences, including the structure of intelligence and personality. More specifically, what are the main

components of human intelligence and what are the main dimensions of personality? Second, we need to establish the *origins* of these individual differences. Are some individuals more intelligent or more extraverted because of their genetic make-up, because of their experiences in life, or because of a combination of the two? Third, we need to identify the underlying *mechanisms* (e.g., physiological; cognitive) responsible for individual differences in behavior and personality. The first issue is concerned with description and the second and third issues are concerned with explanation. Not surprisingly, it has proved easier to provide answers at the descriptive level than at the explanatory level.

Chapter 11

Contents

Introduction	267
Intelligence testing	271
Theories of intelligence	272
Heredity and environment	276
Size does matter!	283

Intelligence

11

INTRODUCTION

What do we mean by “intelligence”? According to Sternberg (2004, p. 472), it involves “the capacity to learn from experience and adaptation to one’s environment.” An important implication of that definition is that we need to pay attention to cultural differences—what is needed to adapt successfully in one environment may be very different from what is required in another environment. We can distinguish between individualistic cultures (e.g., the United States; the United Kingdom) with a focus on individuals accepting responsibility for their own behavior, and collectivistic cultures (e.g., many Asian and African cultures) with an emphasis on the group rather than the individual (see Chapter 1). We would expect social considerations to loom larger in definitions of “intelligence” in collectivistic cultures (see Sternberg and Kaufman, 1998) than in individualistic ones. For example, the word for intelligence in Zimbabwe is *ngware*, meaning to be careful and prudent in social relationships. In similar fashion, the Taiwanese Chinese people emphasize interpersonal intelligence, that is, the ability to understand (and to get on well with) other people.

There has been more controversy about intelligence than almost any other area within psychology. Some experts argue that individual differences in intelligence are of great importance in understanding why people vary so much in their behavior. In contrast, others argue that intelligence is an almost valueless concept. Some researchers (e.g., H.J. Eysenck, 1981) believe that individual differences in intelligence are almost entirely a result of heredity, whereas others (e.g., Kamin, 1981) claim that only environmental factors matter. Both of these issues are considered in detail later in the chapter.

EMOTIONAL INTELLIGENCE

Traditionally, most intelligence tests focused mainly on individuals’ ability to think and to reason effectively. In recent years, however, there have been various attempts in the United States and Europe to develop tests to assess the more social and interpersonal aspects of intelligence emphasized in many non-Western cultures. Most of these attempts have focused on **emotional intelligence**, which has been defined as “the ability to monitor one’s own and others’ emotions, to discriminate among them, and to use the information to guide one’s thinking and actions” (Salovey & Mayer, 1990, p. 189).

Most early research on emotional intelligence made use of self-report questionnaire measures. For example, Bar-On (1997) developed the Emotional Quotient Inventory to assess five dimensions of emotional intelligence: intrapersonal (e.g., emotional self-awareness); interpersonal (e.g., empathy); stress management (e.g., impulse control); adaptability (e.g., flexibility); and general mood (e.g., happiness). The overall score on the Emotional Quotient Inventory (EQ-i) correlated $-.72$ with neuroticism, $+.56$ with extraversion, and $+.43$ with agreeableness (Geher, 2004). However, EQ-i scores are generally uncorrelated with IQ and have only a small relationship with academic achievement at university (see Geher, 2004). Thus, emotional intelligence as assessed by the Emotional Quotient Inventory mainly involves re-packaging well-established personality dimensions and has little resemblance to intelligence as conventionally defined.

Key Term

Emotional intelligence: the ability to understand one’s own emotions as well as those of others.

Davies, Stankov, and Roberts (1998) carried out several studies to work out exactly what is being measured by various questionnaire measures of emotional intelligence and arrived at the same conclusions. They found that measures of emotional intelligence were unrelated to intelligence assessed by IQ tests, so casting doubt on the notion that “emotional intelligence” is a form of intelligence at all. They also found that high levels of emotional intelligence were associated with high levels of extraversion (i.e., sociability) and low levels of neuroticism (i.e., experience of negative emotional states).

In recent years, various ability-based measures of emotional intelligence have been developed. Of particular importance is the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) (Mayer, Salovey, & Caruso, 2002), which is a development of the Multi-Factor Emotional Intelligence Scale (Mayer, Caruso, & Salovey, 1999). The MSCEIT is based on the notion that *four* main abilities underlie emotional intelligence:

1. *Perceiving emotions*: Identification of emotion in oneself and in others (e.g., identifying facial emotions).
2. *Using emotions*: Facilitating thought and action through experiencing the optimal emotion. For example, being in a positive mood can enhance creative thinking (e.g., Isen, Johnson, Mertz, & Robinson, 1985).
3. *Understanding emotions*: The ability to comprehend the language of emotion and to make sense of complicated relationships among emotions.
4. *Managing emotions*: Regulation of emotion in oneself and in others (e.g., evaluating various courses of action in emotional circumstances).

The MSCEIT provides separate assessments of the four abilities described above. It is assumed that individuals high in emotional intelligence are in tune with social norms, and so the MSCEIT is scored in a consensual way. In other words, high scores indicate that an individual has responded to the items in a similar way to a very large sample of

respondents across the world. It is also possible to calculate a score based on similarity between an individual’s responses and those of experts (21 emotion researchers). Both methods produce very similar results, with correlations between them often exceeding $+ .90$ (Mayer, Salovey, Caruso, & Sitarenios, 2003; Palmer, Gignac, Manocha, & Stough, 2005). Both scoring methods indicate that females score significantly higher than males (e.g., Palmer et al., 2005). It has sometimes proved difficult to find evidence for the ability of using emotion to facilitate thought with the MSCEIT (Palmer et al., 2005).

The MSCEIT has small or moderate correlations with most intelligence and personality factors (see Mayer, Salovey, & Caruso, 2004, for a review). The ability factor understanding emotions correlates about $+ .30$ with IQ, and total scores on the MSCEIT correlate approximately $+ .20$ with openness to experience and agreeableness and even less with the other Big Five factors.

What has been found using the MSCEIT? First, emotional intelligence as assessed by the MSCEIT predicted deviant behavior in male adolescents even when controlling for the effects of the Big Five personality factors and analytic intelligence (Brackett, Mayer, & Warner, 2004). Second, students with high scores on the MSCEIT were rated more positively for personal qualities than those with low scores (Lopes et al., 2004). Third, heterosexual couples in which both individuals had high MSCEIT scores were much happier than couples in which both had low scores (Brackett et al., 2004). Fourth, employees with higher scores on the MSCEIT were rated by colleagues as easier to

Test your own EQ

It’s fairly difficult to test any sort of intelligence, but emotional intelligence includes whether a person is basically optimistic or pessimistic in outlook. It is suggested that how we respond to hassles, setbacks, and obstacles is a clue to this aspect of EQ.

Test yourself! Look at these five sample statements and choose either A or B as your response to each.

1. You put on weight on holiday and now can’t lose it
A. I’ll never be thin
B. The latest fad diet isn’t right for me
2. You’ve had a nasty fall playing sport
A. I’ll never be any good at sport
B. The ground was very slippery
3. You’ve lost your temper with a friend
A. We always end up arguing
B. Something must have upset her/him
4. You are feeling really run-down and exhausted
A. I never get the chance to relax
B. I’ve been unusually busy this month
5. You’ve forgotten your best friend’s birthday
A. I’m just bad at remembering birthdays
B. I have had so much on my mind this week

More As: You tend to take setbacks personally—you are naturally pessimistic.

More Bs: You believe that life’s obstacles can be overcome—you are naturally optimistic.

deal with, as more interpersonally sensitive, more tolerant of stress, more sociable, and having greater potential for leadership than those with lower scores (Lopes et al., 2004). In addition, higher scorers on the MSCEIT earned more and had more promotions.

On the negative side, findings showing that the MSCEIT has some ability to predict job performance may be less impressive than they appear. One reason is that the MSCEIT in part assesses aspects of ability and personality that aren't new or distinctive but merely duplicate those assessed by previous tests. Another reason is that the measures of job performance sometimes used are inadequate. For example, Janovics and Christiansen (2002) found that the MSCEIT correlated $+0.22$ with job performance assessed by supervisors' ratings even when controlling for general ability and conscientiousness. However, supervisors' ratings are influenced by employees' personality as well as by their job performance, and no attempt was made to control for likely relevant personality factors (e.g., neuroticism; extraversion).

Evaluation

- + Emotional intelligence is of real importance and deserves to be the focus of research.
- + Traditional approaches to intelligence are somewhat narrow, and an emphasis on emotional intelligence serves to broaden intelligence research.
- + The MSCEIT is a reasonably promising measure of emotional intelligence and seems to assess interpersonal sensitivity and related abilities.
- Most self-report questionnaire measures of emotional intelligence are seriously deficient, and assess mainly well-established personality dimensions.
- There is little evidence that any measures of emotional intelligence predict job performance or success over and above that predicted by pre-existing ability and personality measures (Matthews, Roberts, & Zeidner, 2004; Zeidner, Matthews, & Roberts, 2004).
- More research is needed to establish that emotional intelligence is actually an important type of intelligence.

PRACTICAL IMPORTANCE OF INTELLIGENCE

Of what practical usefulness is it to assess individual differences in intelligence? There is convincing evidence that intelligence is very important in everyday life. For example, job performance and academic achievement among students are both moderately well predicted by intelligence or IQ (Mackintosh, 1998). The fact that IQ predicts real-world success indicates that it is assessing something valuable.

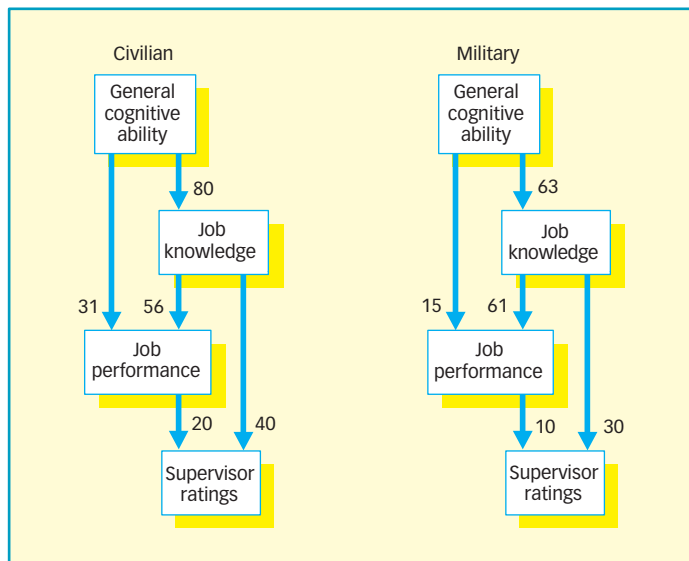
Let's consider the impact of intelligence on job performance in more detail. Hunter and Hunter (1984) considered over 32,000 workers performing 515 different jobs. They identified five levels of job complexity: professional-managerial jobs; high-level complexity technical jobs; medium-complexity jobs; semi-skilled jobs; and completely unskilled jobs. The average correlations between intelligence and job performance were as follows: $+0.58$ for professional jobs, $+0.56$ for complex technical jobs, $+0.51$ for medium-complexity jobs, $+0.40$ for semi-skilled jobs, and $+0.23$ for unskilled jobs.

It may seem surprising that something as general and broad as intelligence predicts job performance so well across such a wide range of jobs. For example, it is plausible to assume that a musician needs highly specific musical skills to perform his/her job effectively. However, the evidence indicates that specific ability or aptitude tests typically contribute surprisingly little to predicting job performance over and above that provided by general measures of intelligence (e.g., Ree et al., 1994). The most convincing evidence

was reported by Hunter (1983). He studied four very large samples of military personnel undergoing job training programs. In all four samples, intelligence strongly predicted training performance and also strongly predicted specific aptitude or ability scores. There were *no* direct effects of specific abilities on training performance because the apparent effects of specific abilities were caused by their relationship with general intelligence.

Why is there a strong association between intelligence and job performance? We will discuss two possible answers. The first starts with the finding that there is a moderate correlation between intelligence and socioeconomic status (Mackintosh, 1998). As a result, at least some of the apparent effects of intelligence on job performance might actually be due to socioeconomic status and related factors such as school quality or neighborhood. This possibility was apparently disproved by Murray (1998). He used a sample of male full biological siblings in intact families, thus controlling for socioeconomic status, schools, neighborhood, and so on. The siblings with higher intelligence had more prestigious jobs and higher income. When they were in their late 20s, a person with average intelligence earned on average nearly \$18,000 (£9000) less per annum than his sibling with an IQ of at least 120, but over \$9000 (£4500) more than his sibling with an IQ of 80 or less.

The second answer to the question above was provided by Hunter and Schmidt (e.g., 1996). They argued that the ability to learn rapidly is of crucial importance in most jobs, and learning ability is determined by intelligence. Successful job performance also sometimes requires that workers respond in an innovative or adaptive fashion, and more intelligent workers can respond more adaptively than less intelligent ones. Relevant evidence was provided by Hunter (1986), who combined the data from 14 studies on civilian and military groups (discussed in detail by Gottfredson, 1997). Three findings supported the theory. First, there was a high correlation between intelligence and job knowledge with both types of jobs (see figure on the left). Second, learning in the form of job knowledge was strongly associated with job performance. Third, there was a direct influence of intelligence on job performance that didn't depend on job knowledge. This may reflect the greater ability of intelligent workers to respond adaptively.



A path analysis of findings from studies on civilian and military groups showing direct and indirect (via job knowledge) effects of intelligence (general cognitive ability) on job performance. The numbers indicate the strength of the relationship between factors. From Hunter (1986) Copyright © 1986 Elsevier. Reproduced with permission.

Health and longevity

Individual differences in intelligence also predict health and longevity (Gottfredson & Deary, 2004). For example, Whalley and Deary (2001) found that individuals at a 15-point disadvantage in IQ relative to other individuals were only 79% as likely to live to age 76. Among the women, the less intelligent ones had a 40% increase in cancer deaths compared to the more intelligent ones, and the comparable figure for men was 27%.

How can we explain the effects of intelligence on health? It is difficult to answer that question because numerous factors could be involved. For example, part of the answer is that more intelligent individuals on average have more money and so can afford better food and living conditions. However, there is another important factor. If you are to maximize your chances of having a long life, you need to interpret your bodily symptoms, take account of information on health issues, adhere to instructions for taking medicine, avoid destructive habits such as smoking, and so on. In other words, you need good **health literacy**. More intelligent individuals have greater health literacy than less intelligent ones (see Gottfredson & Deary, 2004). The problems posed by low levels of health literacy were considered by Williams et al. (1995). Of those with “inadequate” health literacy, 65% didn’t understand directions for taking medicine on an empty stomach and 40% couldn’t work out from an appointment slip when their next appointment was due.

Key Term

Health literacy: the ability to understand health-related issues (e.g., taking medications appropriately).

INTELLIGENCE TESTING

The first proper intelligence test was produced in 1905 by two French psychologists, Alfred Binet and Théodore Simon. It measured comprehension, memory, and various other psychological processes. Among the well-known tests that followed are the Stanford-Binet test produced at Stanford University in the United States, the Wechsler Intelligence Scale for Children, and the British Ability Scales. These (and other) tests are designed to measure several aspects of intelligence. They often contain mathematical items, and many contain vocabulary tests in which individuals are asked to define the meanings of various words. Many tests contain problems based on analogies (e.g., “Hat is to head as shoe is to ____”), and items relating to spatial ability (e.g., “If I start walking northwards, then turn left, and then turn left again, what direction will I be facing?”).



The French psychologist, Alfred Binet (1857–1911).

CALCULATING IQ

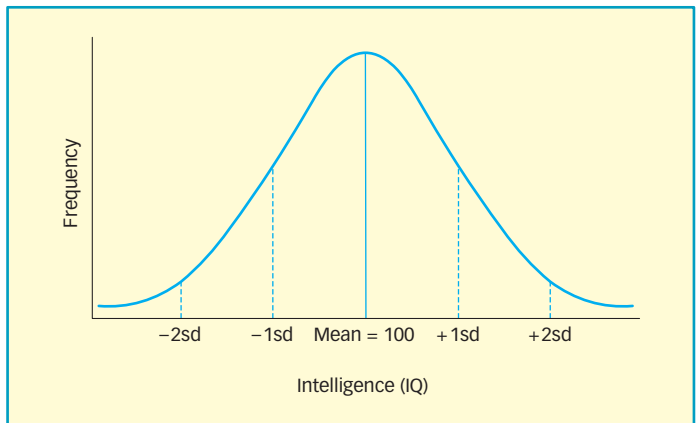
Suppose you complete an intelligence test containing 150 items, and you obtain a score of 79. Have you done well or badly? The only way to answer that question is by comparing your performance against that of other people. That can only be done effectively by using a **standardized test**, that is, one that has been given to a large, representative sample of the population. When your score has been compared to those of others, it is possible to calculate your IQ or **intelligence quotient**, an overall measure of intellectual ability.

How is IQ calculated? Most intelligence tests are devised so that the IQs from the general population are normally distributed. The normal distribution is a bell-shaped curve in which there are as many scores above the mean as below it (see the figure below). Most of the scores are close to the mean, and there are fewer and fewer scores as we move away from the mean in either direction. The spread of scores in a normal distribution is usually indicated by the **standard deviation**. In a normal distribution, 68% of the scores fall within one standard deviation of the mean or average, 95% fall within two standard deviations, and 99.73% are within three standard deviations. Intelligence tests are designed to produce a mean IQ of 100 and a standard deviation of about 16. Thus, if you have an IQ of 116, your IQ is greater than that of 84% of the population. That is the case because 50% fall below 100, and a further 34% between the mean and one standard deviation above it. Note that intelligence-test designers take steps to ensure that their test has a normal distribution of IQs—it is very difficult to know whether this corresponds to the “true” distribution.

Those with high IQs don’t usually perform well on all the tests on an intelligence-test battery, and nor do those with low IQs perform poorly on every test. As a result, tests are generally constructed to obtain measures of various abilities (e.g., numerical; spatial; reasoning; perceptual speed). We can obtain a more accurate assessment of an individual’s intellectual ability by considering the profile of his/her performance across these abilities rather than by focusing only on IQ.

RELIABILITY AND VALIDITY

We have seen that intelligence tests need to be standardized in order to be useful. Two other essential features are reliability and validity. **Reliability** refers to the extent to which a test provides consistent or replicable findings, whereas **validity** refers to the extent to which a test measures what it is supposed to be measuring. The most common way



Key Terms

Standardized test:

a test given to large representative samples so that an individual's scores can be compared to those of the population.

Intelligence quotient:

a measure of general intellectual ability; the mean IQ is 100.

Standard deviation:

a measure of the spread of scores in a bell-shaped or normal distribution. It is the square root of the variance, takes account of every score, and is a sensitive measure of dispersion or variation.

Reliability:

the extent to which a test gives consistent findings on separate occasions.

Validity:

the extent to which a test measures what it claims to be measuring.

to assess reliability is by the test–retest method. A group of people take the same test on two separate occasions, and their two scores are correlated. The higher the correlation (a measure of the relationship between the two scores), the greater is the reliability of the test. The highest possible correlation is +1.00, which would indicate perfect agreement or reliability. In contrast, a correlation of .00 would indicate no reliability. In fact, reliability correlations tend to be about +.85, which indicates high reliability.

What about validity? The most direct approach is to relate IQ to some external criterion or standard. This is known as **concurrent validity** when IQ is correlated with currently available information about a criterion, and as **predictive validity** when the criterion measure is obtained after the intelligence test has been administered. For example, we would expect more intelligent people to achieve greater academic success than less intelligent ones, and to move on to have more successful careers. However, academic and career success obviously also depend on several other factors such as motivation, support of family and friends, the quality of teaching or training provided, and so on. In spite of that, IQ typically correlates about +.50 with academic achievement and with occupational status (see Mackintosh, 1998, for a review). As discussed above, IQ predicts job performance to a greater extent with complex jobs than with relatively simple ones (Hunter & Hunter, 1984).

THEORIES OF INTELLIGENCE

So far we have focused mainly on IQ, which is a very general measure of intelligence. You may well be thinking that many people are much more intelligent in some ways than others, but this information is lost in a general measure such as IQ. The evidence indicates strongly that we need to identify general *and* specific aspects of intelligence. The most common approach that is used is **factor analysis**, which is a statistical technique used to decide on the number and nature of the factors underlying intelligence as measured by a test. The first step in factor analysis is to give a large number of items to a large number of individuals, and to obtain scores from each individual on each item. The correlations between these items are then calculated. If two items correlate highly with each other, this means that those who perform well on one item tend to perform well on the other item. The key assumption is that two items correlating highly with each other are assessing the same factor of intelligence. It is also assumed that two items correlating weakly or not at all with each other are *not* assessing the same factor of intelligence. Thus, the pattern of correlations is used to identify the main aspects of intelligence (or factors as they are known).

	Test 1	Test 2	Test 3	Test 4
Test 1	–	+.85	+.12	+.10
Test 2	+.85	–	+.08	+.11
Test 3	+.12	+.08	–	+.87
Test 4	+.10	+.11	+.87	–

We can see what happens in simplified fashion by considering the box on the left. How many factors should we extract from this correlation matrix? The answer is two. Tests 1 and 2 correlate highly with each other, and so are measures of the same factor. Tests 3 and 4 correlate highly with each other (but not with test 1 or test 2), and so they form a different, second factor.

FACTOR THEORIES

The first factor theory of intelligence was put forward by the British psychologist Charles Spearman (1923). In his two-factor theory, there is a general factor of intelligence that he called “g.” He argued that there is a general factor because practically all of the tests contained within an intelligence-test battery correlate positively with each other. However, most of these positive correlations are fairly low, so we can’t account for all the data in terms of a general factor. Accordingly, Spearman argued that there are specific factors associated with each test.

Thurstone (1938) wasn’t convinced of the need to assume that there is a general factor of intelligence. Instead,

Key Terms

Concurrent validity:
assessing **validity** by correlating scores on a test with some currently available relevant criterion (e.g., academic achievement in the case of an intelligence test).

Predictive validity:
assessing **validity** by correlating scores on a test with some future criterion (e.g., career success in the case of an intelligence test).

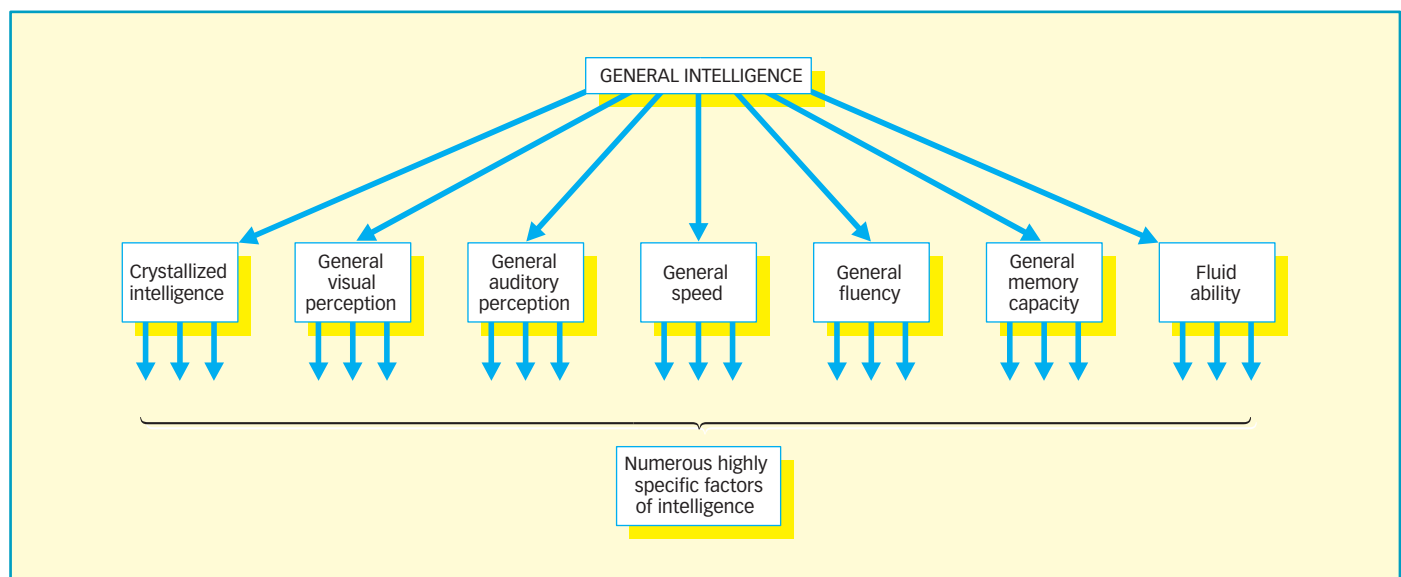
Factor analysis:
a statistical technique applied to intelligence tests to find out the number and nature of the aspects of intelligence they are measuring.

he identified seven factors that he referred to as primary mental abilities. These primary mental abilities were as follows: inductive reasoning; verbal meaning; numerical ability; spatial ability; perceptual speed; memory; and verbal fluency. There is one major problem with this approach. All seven primary abilities correlate positively with each other. As a result, factor analysis of Thurstone's seven factors produces the general factor ignored by Thurstone (Sternberg, 1985).

Hierarchical approach

Nowadays there is fairly general agreement that intelligence is organized hierarchically. One of the most influential hierarchical theories is that of Carroll (1993), who argued that we should combine the insights of Spearman and of Thurstone. He discussed evidence based on factor analysis of over 460 data sets obtained over a 60-year period from more than 130,000 people. His theory (shown in the figure below) identifies *three* levels in the hierarchy:

- At the top level, there is the general factor of intelligence (often referred to as “g”); this is the main ingredient in IQ.
- At the middle level, there are various fairly general factors, including crystallized intelligence (involving the use of acquired knowledge) and fluid ability (used when coping with novel problems and situations).
- At the bottom level, there are very specific factors associated with only one item or a small number of tests.

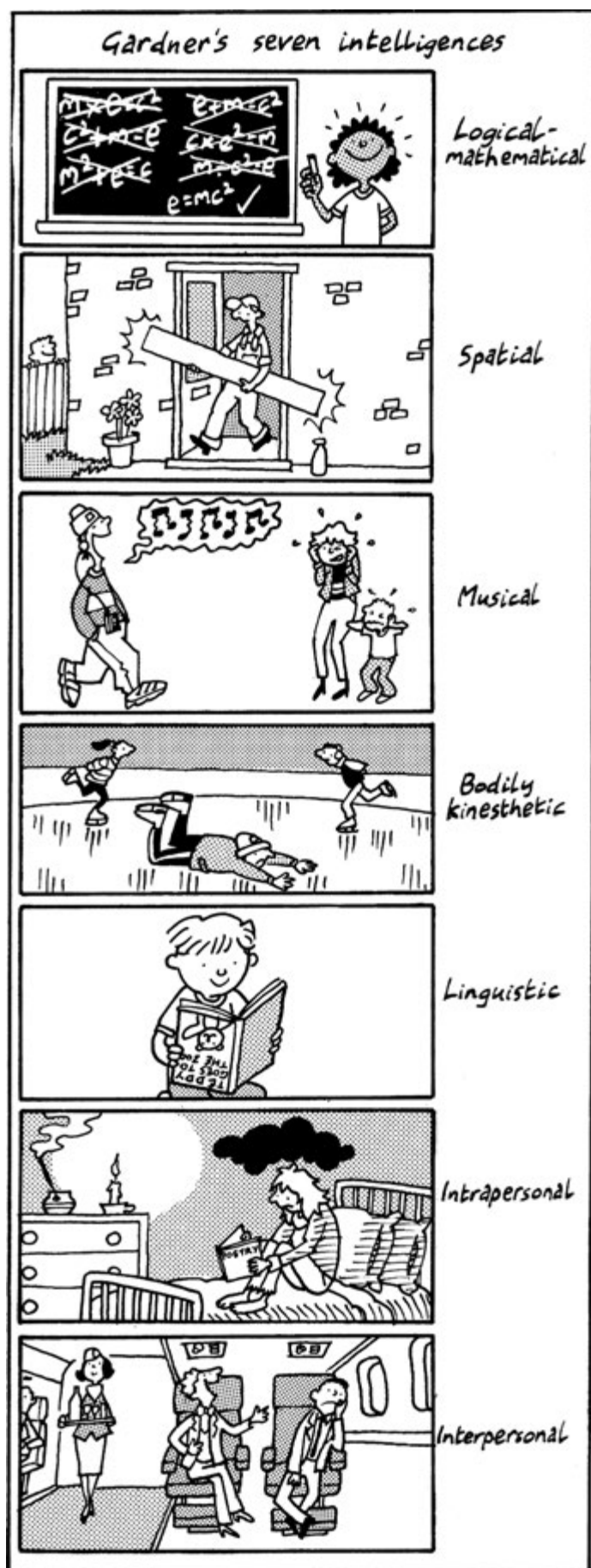


Carroll's (1986) three-level hierarchical model of intelligence.

MULTIPLE INTELLIGENCES

Howard Gardner (1983) argued strongly that most intelligence tests and most theories of intelligence are based on an excessively narrow view of the nature of intelligence. According to Gardner, there are seven separate intelligences and no single general factor. Here are his seven intelligences:

1. *Logical-mathematical intelligence*: This is of special value in handling abstract problems of a logical or mathematical nature.
2. *Spatial intelligence*: This is used when deciding how to go from one place to another, how to arrange suitcases in the trunk of a car, and so on.
3. *Musical intelligence*: This is used both for active musical processes (e.g., playing an instrument; singing) and for more passive processes (e.g., appreciating music).



4. *Bodily kinesthetic intelligence*: This is involved in the fine control of bodily movements in activities such as sport and dancing.
5. *Linguistic intelligence*: This is involved in language activities (e.g., reading; writing; speaking).
6. *Intrapersonal intelligence*: This intelligence “depends on core processes that enable people to distinguish among their own feelings” (Gardner, Kornhaber, & Wake, 1996, p. 211).
7. *Interpersonal intelligence*: This intelligence “makes use of core capabilities to recognize and make distinctions among others’ feeling, beliefs, and intentions” (Gardner et al., 1996, p. 211).

Gardner (1998) proposed adding naturalistic intelligence to the seven intelligences he had previously identified. Naturalistic intelligence is shown by individuals who can perceive patterns in nature. Charles Darwin is an example of a famous person having outstanding naturalistic intelligence. Gardner also suggested that there might be two additional intelligences: spiritual intelligence and existential intelligence. Spiritual intelligence is based on a concern for cosmic issues, and with the achievement of the spiritual as a state of being. Existential intelligence is based on concerns about ultimate issues of existence.

How did Gardner come up with his original seven intelligences? He argued that we should only identify an intelligence when it satisfies various criteria. First, it should depend on identifiable brain structures. Second, studies of brain-damaged patients should indicate that it can be impaired without disrupting any other intelligence. Third, it should involve the use of distinct cognitive processes. Fourth, there should be exceptional individuals showing a remarkable ability (or deficit) with respect to the intelligence. Fifth, in evolutionary terms the development of the intelligence should have improved humans’ adaptation to the environment. Sixth, it should be measurable by appropriate intelligence tests.

In spite of its popularity, there is surprisingly little direct evidence to support Gardner’s theory of multiple intelligences. An exception is the work of Gardner (1993), who used the theory to study creativity. He chose seven individuals who showed outstanding creativity during the early part of the twentieth century with respect to one of the seven intelligences. Albert Einstein was the representative of logical-mathematical intelligence. The other outstanding figures were Pablo Picasso (spatial intelligence), Igor Stravinsky (musical intelligence), Martha Graham (bodily kinesthetic intelligence), T.S. Eliot (linguistic intelligence), Sigmund Freud (intrapersonal intelligence), and Mahatma Gandhi (interpersonal intelligence). Nearly all of them were brought up in families forcing them to meet standards of excellence. They all had childlike qualities, showing signs of behaving like a “wonder-filled child.” They were all very ambitious, which led them to sacrifice other aspects of their lives and to cause suffering to their families.



Three of the individuals selected by Gardner (1993) to demonstrate his theory of multiple intelligences: Igor Stravinsky (for musical intelligence), Pablo Picasso (for spatial intelligence), and Mahatma Gandhi (for interpersonal intelligence).

This genius-based approach to identifying intelligences may be flawed. As Jensen (in Miele, 2002, p. 58) pointed out sarcastically, the logic of this approach is that we could claim that “Al Capone displayed the highest level of ‘Criminal Intelligence,’ or that Casanova was ‘blessed’ with exceptional ‘Sexual Intelligence.’” In other words, Gardner’s criteria for what constitutes an intelligence are so lenient that almost anything goes. For example, face recognition and the ability to learn foreign languages appear to meet his criteria for an intelligence (Mackintosh, 1998).

Evaluation

- + Gardner’s approach to intelligence is broader in scope than most others.
- + There is some supporting evidence (e.g., from geniuses; from brain-damaged patients) for all seven intelligences originally proposed by Gardner.
- The seven intelligences correlate positively with each other, whereas Gardner (1983) assumed they were independent. That means that Gardner was wrong to disregard the general factor of intelligence. As Jensen (in Miele, 2002, p. 59) pointed out, the evidence indicates that, “A level of *g* [general factor] beyond the 90th percentile [IQ over 120] is probably necessary for recognized achievement in politics, the military, business, finance, or industry.”
- Musical and bodily kinesthetic intelligences are less important than the other intelligences in Western cultures, with many very successful people being tone-deaf and poorly coordinated.
- The criteria for an intelligence are too lenient.
- The theory is descriptive rather than explanatory—it fails to explain *how* each intelligence works.

HEREDITY AND ENVIRONMENT

Why are some children and adults more intelligent than others? At the most general level, only two factors could be responsible: heredity or environment. Heredity consists of a person's genetic endowment, and environment consists of the situations and experiences encountered by people in the course of their lives. The only sensible view is that both heredity and environment contribute to individual differences in intelligence. However, what is much more complex is to decide on their relative importance.

It is perhaps natural to think of heredity and environment as having entirely separate or *independent* effects on intelligence. However, the reality is very different, because our genetic make-up influences the types of environmental experiences we have. We can see how implausible it is to assume that heredity and environment have entirely separate effects by applying that dubious assumption to basketball performance (Dickens & Flynn, 2001):

Good coaching, practising, preoccupation with basketball, and all other environmental factors that influence performance must be unrelated to whether genes contribute to someone being tall, slim, and well coordinated. For this to be true, players must be selected at random for the varsity basketball team and get the benefits of professional coaching and intense practice, without regard to build, quickness, and degree of interest.

In fact, of course, individuals whose genes predispose them to be outstanding at basketball are more likely to put themselves into an environment supporting excellent basketball performance than are those whose genes do not.

We can apply precisely the same logic directly to intelligence. No one believes that the probability that a given individual will choose to read hundreds of books or to go to university has nothing to do with his/her level of genetic ability. In fact, individuals with the highest levels of genetic ability are much more likely to seek out intellectually demanding situations than are those with less genetic ability. In the words of Dickens and Flynn (2001, p. 347), "Higher IQ leads one into better environments causing still higher IQ, and so on." Thus, individual differences in intelligence depend to a fairly large extent on the effects of heredity on environment.

Nicotine as nurture

We are all familiar with the fact that nicotine has an effect on the brain, but most of the media publicity relates to its physical ill effects such as increasing the risks of developing lung cancer and cardiovascular disease. However, there are other, more subtle behavioral effects. Jacobsen's (2007) team at Yale University studied groups of teenagers. One group were smokers; another group had mothers who had smoked when pregnant; a third group were nonsmokers; the fourth group had mothers who did not smoke during pregnancy. The first two groups were the experimental groups (a natural experiment, obviously) and the second two groups were the controls, with the groups being comparable in age, educational attainment, IQ, parental education, and symptoms of inattention. A consistent finding was that the first two groups had changes in their brain structure in the brain connections in the areas which process auditory information. What effect did these changes have? Those groups of teenagers had more difficulty in focusing their auditory attention when

there were distractions going on. The adverse effect of nicotine both antenatally and in adolescence is to disrupt the development of anterior cortical white matter in the auditory cortex. This seems to lead to a deficiency in attentional processing so that there is too much "noise" reaching that part of the cortex, thus reducing the efficiency of selecting what is important. This means that not only did indirect exposure to nicotine before birth seem to cause deficits in brain development, but also that in the teenage years when many neural pathways are maturing nicotine can also have an adverse effect. Further research to investigate whether these effects of nurture on nature are permanent is already being done. Overall, the results will not be known for some time.

Overall, the research shows that though brain structure is controlled by genes, and so is an example of nature, environmental factors at key times in life can alter or modify this structure and so such factors are a real example of nurture—and nature and nurture interact.

Plomin (1990) identified three types of interdependence between genetic endowment and environment:

1. *Active covariation*: This occurs when children of differing genetic ability look for situations reinforcing their genetic differences (e.g., children of high genetic ability reading numerous books).
2. *Passive covariation*: This occurs when parents of high genetic ability provide a more stimulating environment than parents of lower genetic ability.
3. *Reactive covariation*: This occurs when an individual's genetically influenced behavior helps to determine how he/she is treated by other people.

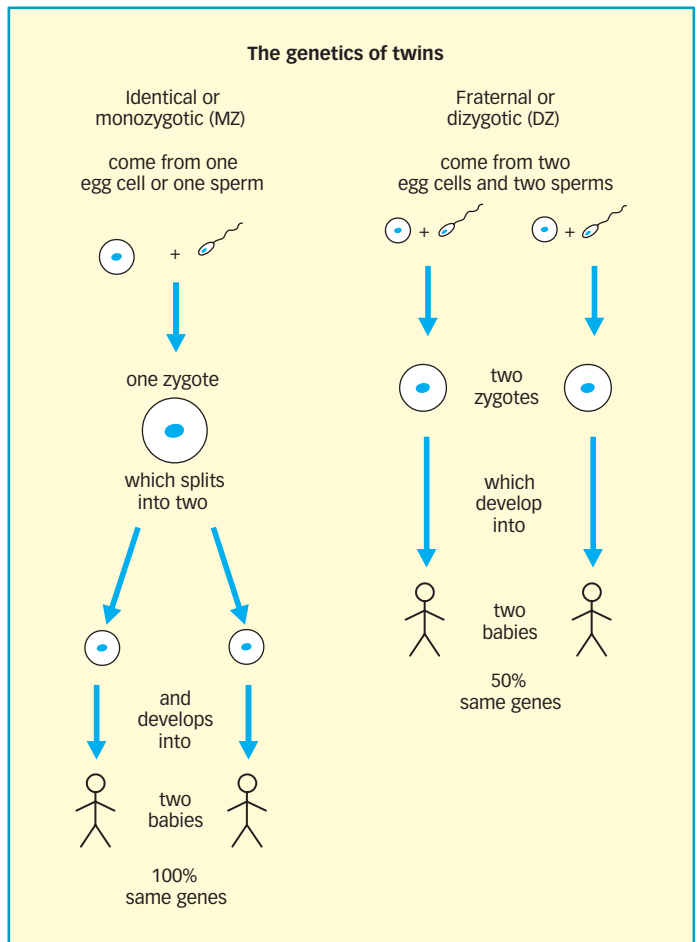
When we consider the role of heredity in determining individual differences in intelligence, we are interested in an individual's genetic potential (known as the **genotype**). However, we cannot assess the genotype directly. All we can measure directly are an individual's observable characteristics (the **phenotype**). How psychologists have grappled with this problem is discussed below.

TWIN STUDIES

The most valuable approach to assessing the roles of heredity and environment in determining individual differences in intelligence involves studying twins. Identical or **monozygotic twins** derive from the same fertilized ovum, and so have essentially identical genotypes. In contrast, fraternal or **dizygotic twins** derive from two different fertilized ova. As a result, their genotypes are no more similar than those of two ordinary siblings, that is, they share on average 50% of their genes (see the figure on the right). If heredity influences individual differences in intelligence, then identical twins should be more alike in intelligence than fraternal twins. In contrast, if environmental factors are all-important, then identical twins should be no more alike than fraternal twins.

The degree of similarity in intelligence shown by pairs of twins is usually reported in the form of correlations. A correlation of +1.00 would mean that both twins in a pair have very similar or the same IQs, whereas a correlation of 0.00 would mean that there is no relationship between the IQs of twins. Bouchard and McGue (1981) reviewed 111 studies, and reported that the mean correlation for identical twins was +.86 compared to +.60 for fraternal twins. McCartney, Harris, and Bernieri (1990) reported similar findings from a later analysis of numerous studies: the mean correlation for identical twins was +.81 compared to +.59 for fraternal twins.

On the face of it, the above findings indicate clearly that individual differences in intelligence depend to a fair extent on genetic factors. However, there are complicating factors. For example, there is the prenatal environment. All fraternal twins have separate placentas in the womb, whereas two-thirds of identical twins share a placenta. As a result, the prenatal environment of most identical twins is more similar than that of fraternal twins. Phelps, Davis, and Schwartz (1997) reviewed several studies in which it was found that identical twins



Key Terms

Genotype:

an individual's potential in the form of genes.

Phenotype:

an individual's observable characteristics based on his/her genotype plus life experiences.

Monozygotic twins:

identical twins derived from a single fertilized ovum and sharing 100% of their genes.

Dizygotic twins:

fraternal twins derived from two fertilized ova and sharing 50% of their genes.

sharing a single placenta were more similar in intelligence than those having separate placentas. Thus, the greater similarity in IQ between identical twins than fraternal ones may depend in part on the greater similarity of the prenatal environment for most identical twins.

Another potential problem is that identical twins are treated in a more similar way than fraternal twins in terms of parental treatment, playing together, and being taught by the same teachers (Loehlin & Nichols, 1976). However, these differences are relatively small, and parents may treat identical twins more similarly than fraternal twins because they are responding to the greater genetically influenced similarities in the behaviour of identical twins (i.e., reactive covariation).

Identical twins are relatively rare, and identical twins brought up in separate families are obviously even rarer (see the Case Study below). However, they are of particular value in assessing the roles of heredity and environment in determining individual differences in intelligence. Identical twins brought up apart should be very similar to each other in IQ if genetic factors are very important. In contrast, those favoring an environmentalist position would argue that placing twins in different environments should ensure they are not similar in intelligence.

On the face of it, the findings support the hereditarian position. Bouchard et al. (1990) studied more than 40 adult identical twin pairs separated at a mean age of 5 months. In spite of the fact that these twin pairs had been separated in infancy, their IQs correlated $+0.75$. However, there is a real limitation in this and other studies

Case Study: *Separated Twins*

Once upon a time in New York a pair of twin girls were born and, after being put separately into foster care, were adopted by two different couples. Neither the girls, nor the couples, were aware that they were twins separated at birth.

Elyse grew up, didn't marry, went to Paris and worked as an independent movie maker. Paula remained in New York, and made a career in journalism specializing in film. She had a partner and a baby daughter. Paula was not interested in her biological past but Elyse was, and returned to New York to investigate her heritage. And one day, 35 years after their separation, the phone rang in Elyse's rented New York apartment and for the first ever time the identical twin sisters spoke.

Physically the identical twin sisters are, as one would expect, very similar. Other similarities include a tendency to introspection and depression and the same mannerisms, allergies, political views, and taste for obscure, independent movies. As children they sucked their fingers the same way, and had eating disorders in their teens. They both write, and at high school edited their student newspapers. They both went to university to do movie studies. They even wear the same shade of lipstick! These are behavioral and cognitive similarities, and the question is whether they are caused by the identical genes or by upbringing—nature or nurture.

It is a fact that identical twins have identical genotypes. However these two women had rather different childhood environments. Paula had a secure and affluent home, whereas Elyse's mother died when the girl was only 6 years old, and her adopted brother developed schizophrenia.

So while nature was identical, nurture was not. And yet both women have a caring, nurturing approach, and both prefer the same sorts of people though they are clearly not the same person. For example, Elyse seems quieter, more thoughtful, and takes more time before answering, compared to Paula.

A separate but critically important question is why the twins were split up in the first place, and the answer is uncomfortable. The psychiatrist at the adoption agency, together with a prominent psychologist, decided to use the twins and four other similar twin pairs plus one set of triplets as a secret experiment, a study of nature versus nurture on human development. These siblings were deliberately split up, separated at birth and adopted separately by couples who did not know the baby was one of two or three. This was seen as a great research opportunity as fewer than 300 twin pairs who were separated at birth have been identified, and most of these had contact with each other or were brought up by relatives who knew they were twins. By today's standards the ethical issues with this experiment are huge and very disturbing. Once the researchers realized that their actions were likely to be heavily criticized the files on the study were sealed until 2066, continuing the secrecy until the twins, if still alive, will be nearly 100 years old.

What do you feel this case history tells us about the nature–nurture debate?

What would you identify as the most important ethical issues in this experiment?

focusing on identical twins brought up apart. More than half of such twin pairs were brought up in different branches of the same family. As a result, their environments may well have been rather similar, which could explain at least some of the similarity in their IQs.

What conclusions can we draw from the evidence? Psychologists trying to be as precise as possible about the contributions of heredity and environment to intelligence generally assess heritability. The technical definition of **heritability** is the ratio of genetically caused variation to total variation (genetic + environmental variation) within any given population. Thus, heritability tells us something about the role of genetic factors in producing individual differences in intelligence. It is important to note that heritability is a *population* measure, and so varies considerably from one population to another. The more similar the environmental factors experienced by those living in a given culture, the greater will be the effect of genetic factors in determining individual differences in intelligence. For example, Brace (1996) found that the heritability of intelligence was much higher among people living in affluent white American suburbs than among people living in American urban ghettos. This occurred because the great majority of those living in affluent white suburbs enjoy a favorable environment throughout childhood.

There is another important point to be made about the heritability measure. In essence, it combines two kinds of genetic influence. First, there is a *direct* genetic influence on intelligence. Second, there is an *indirect* genetic influence on intelligence in which genetic factors affect the environments in which an individual finds himself/herself and the environment then affects his/her intelligence. For example, individuals of high genetic ability are far more likely to go to university than those of low genetic ability, and the educational environment provided by a university enhances the intelligence of those who go there. These indirect effects in which genetic factors influence the environment, which then influences intelligence, are counted as genetic effects when heritability is calculated—you may feel that is somewhat unfair given that environmental effects are strongly involved.

Mackintosh (1998) reviewed the evidence based on heritability measures. He concluded that between 30% and 75% of individual differences in intelligence in modern industrialized societies are a result of genetic factors. However, large differences in heritability have been found as a function of age. According to Plomin (1988, p. 420), the genetic influence on individual differences in IQ “increases from infancy (20%) to childhood (40%) to adulthood (60%).” It is not known for certain why the heritability of intelligence increases during development. However, adolescents and adults select and control their own environment to a greater extent than children, and this reduces the impact of the environment on intelligence. Thus, the increase in heritability of intelligence during the course of development may reflect what Plomin (1990) called active covariation (discussed earlier).

ADOPTION STUDIES

Adoption studies provide another way of assessing the relative importance of heredity and environment in determining individual differences in intelligence. If heredity is more important than environment, adopted children’s IQs will be more similar to those of their biological parents than those of their adoptive parents. The opposite pattern will be found if environment is more important. In fact, the IQs of adopted children typically resemble those of their biological parents more than those of their adoptive parents, suggesting that genetic factors are important.

It is difficult to interpret the findings from most adoption studies. There is the issue of **selective placement**, meaning that adopted children tend to be placed in homes resembling those of their biological parents in social and educational terms. Thus, some of the similarity in IQ between adopted children and their biological parents may occur because they are living in an environment resembling the one their mother would have provided.

Capron and Duyne (1989) carried out an important study on adopted children, in which there was little evidence of selective placement. They used four very different

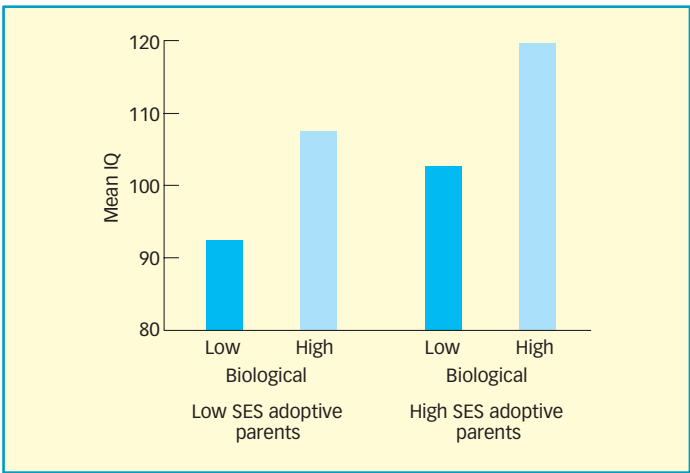
Key Terms

Heritability:

the ratio of genetically caused variation to total variation (a combination of genetic and environmental variation) within a given population.

Selective placement:

placing adopted children in homes with similar educational and social backgrounds to those of their biological parents.



Mean IQs of adopted children as a function of socioeconomic status (SES) of their biological parents (low vs. high) and their adoptive parents (low vs. high). Data from Capron and Duyne (1989).

groups of adopted children. These groups consisted of all four combinations of high and low socioeconomic status biological parents and high and low socioeconomic status adoptive parents.

What do you think happened in the study by Capron and Duyne (1989)? If genetic factors are of more importance, the intelligence of the adopted children should have depended mainly on the socioeconomic status of the biological parents. In contrast, if environmental factors are more important, the adopted children’s intelligence should have depended mainly on the socioeconomic status of the adoptive parents. In fact, the effects of the socioeconomic status of the biological and of the adoptive parents were comparable (see the figure on the left). Thus, genetic and environmental factors were equally important in determining the adopted children’s intelligence.

ENVIRONMENTAL STUDIES

The evidence from twin and adoption studies indicates clearly that genetic factors are only partly responsible for individual differences in intelligence. That means that environmental factors are also important. We can use data from twin and adoption studies to assess two kinds of environmental influence: shared environment and non-shared environment. **Shared environment** refers to all the common influences within a family that make children resemble each other. **Nonshared environment** refers to all those influences unique to any given child (e.g., experiences with peers; instruction from particular teachers).

Twin and adoption studies suggest that about 20% of individual differences in intelligence are a result of nonshared environment. The impact of shared environment changes over time. Shared environment accounts for 25% of individual differences in intelligence in childhood but very little after adolescence. This happens because adolescents spend much less time than children being exposed to the influence of parents and home.

Some of the strongest evidence that environmental factors can have a substantial effect on intelligence was reported by Flynn (1987, 1994). He obtained evidence

Key Terms

Shared environment:
environmental influences common to the children within a given family.

Nonshared environment:
environmental influences that are unique to a given individual.

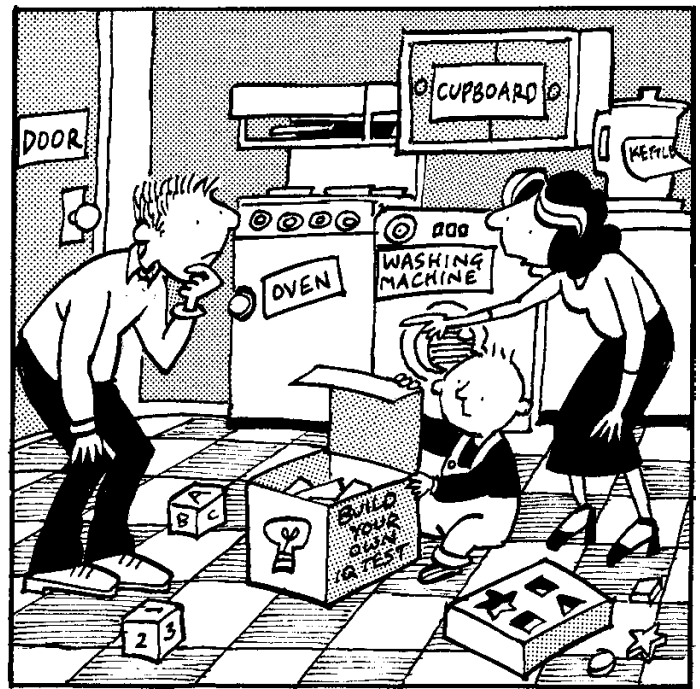
Flynn effect:
the rapid rise in average IQ in several Western countries in recent decades.

General criticisms of IQ tests, adoption studies, and twin studies		
IQ tests	Adoption studies	Twin studies
Debatable whether IQ is an adequate measurement of intelligence	Selective placement makes it hard to determine the effects of heredity and environment	Environmental similarity often occurs
Cultural differences not always considered	Heredity is less well controlled than in twin studies	Twins raised separately are often raised by different branches of the same family
		Twins have often spent some years together before being separated

from 20 Western countries, each of which showed the **Flynn effect**: a rapid rise in mean IQ in most Western countries in recent decades. Flynn (1987) reported that there had been an increase of 2.9 points per decade on nonverbal IQ and of 3.7 points on verbal IQ. This large and rapid increase could not possibly be a result of genetic factors. Embarrassingly for psychologists, we don't know precisely *why* the Flynn effect does occur. However, several environmental factors probably contribute to the effect, and here are some likely contenders:

- Increases in the number of years of education.
- Greater access to information (e.g., internet; television).
- Increased cognitive complexity of the average job now compared to the past.
- A large increase in the percentage of middle-class families.

As a footnote, there is some evidence that the Flynn effect is coming to an end. For example, Sundet, Barlaug, and Torjussen (2004) found in Norway that there had been no general increase in IQ since the mid-1990s.



Key Study

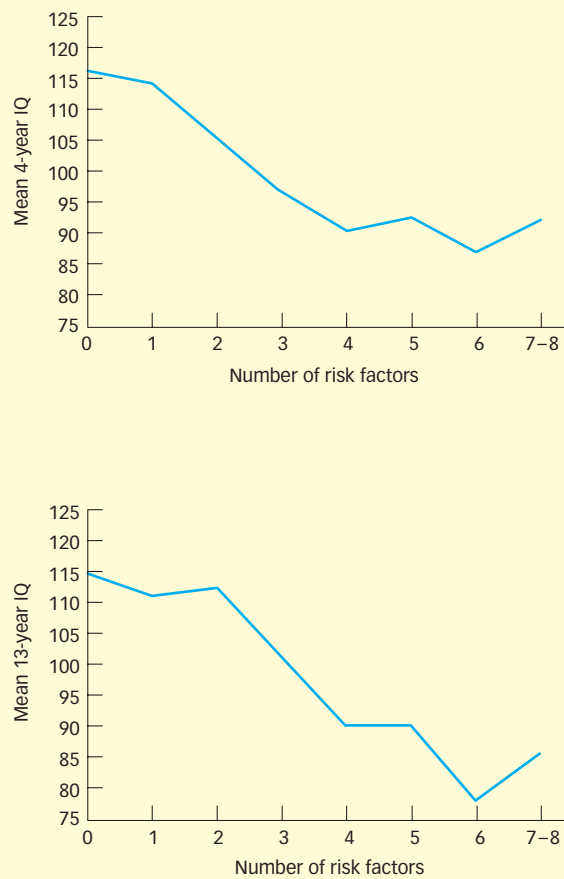
Sameroff et al. (1987, 1993): The Rochester Longitudinal Study

Sameroff et al. (1987, 1993) carried out a longitudinal study in New York State to investigate the environmental factors that might be linked to intellectual delay in young children. They selected pregnant women to be part of their study and followed 215 children, testing their IQs at ages 4 and 13. The families represented a range of socioeconomic backgrounds, maternal age groups, and number of other siblings.

Sameroff et al. (1993) identified 10 family risk factors related to lower IQ:

- Mother has a history of mental illness.
- Mother did not go to high school.
- Mother has severe anxiety.
- Mother has rigid attitudes and values concerning her child's development.
- Few positive interactions between mother and child during infancy.
- Head of household has a semi-skilled job.
- Father does not live with the family.
- Child belongs to a minority group.
- Family suffered 20 or more stressful events during the child's first 4 years of life.

There was a clear negative association between the number of risk factors associated with a child and his/her IQ (see the figure on the following page). At age 4, this correlation was $-.58$, and it was $-.61$ at age 13. Thus, there was a moderately strong association between risk factors and IQ. At the age of 4, high-risk children were 24 times more likely than low-risk children to have IQs below 85. On average, each risk factor reduced the child's IQ by 4 points.



These graphs show the negative association between IQ and number of environmental risk factors. The top graph presents data for mean 4-year IQ scores, and the bottom graph presents data for mean 13-year IQ scores.

Discussion points

1. Select one of the risk factors and suggest how it might affect intellectual development.
2. What are the political implications of this study?

KEY STUDY EVALUATION

It is difficult to interpret the above findings. For example, Sameroff et al. (1993) reported that the mother not going to high school and the head of the household having a semi-skilled job were associated with low IQ in the children. It is entirely possible that genetic factors played a part in producing these environmental factors and in leading to low IQ in children. Thus, we can't be confident that the environmental factors themselves actually produced a negative effect on the children's IQ.

SIZE DOES MATTER!

There is a popular view that people with big heads (and so presumably large brains) are more intelligent than those with small heads. However, when psychologists first studied the link between brain size and intelligence, they found practically no relationship. However, these early studies were very limited because the findings were based on imprecise estimates of brain size. One technique was based on skull size, but that provides a very inaccurate estimate. Another equally flawed technique was to measure the sizes of the shrunken brains of people who had recently died. Nowadays we can obtain good measures of brain size in living people by using brain-imaging techniques such as magnetic resonance imaging (MRI). Since male brains are slightly larger than female brains on average, what is often done is to use MRI to correlate brain size with intelligence in each sex separately.

McDaniel (2005) reported a meta-analysis based on 37 different samples. The average correlation between brain size or volume and intelligence was $+.33$. This indicates that people with larger brains do tend to be more intelligent. McDaniel also considered each sex separately. The average correlation between brain volume and intelligence was $+.40$ for females and $+.34$ for males. All of these correlations are impressive in view of the fact that brain volume is a crude measure.

It is not clear exactly *why* brain size is associated with intelligence. However, there is evidence from studies on children that nutritionally enhanced diets produce increases in IQ (Benton, 2001). As Peters (1995) pointed out, the relationship between brain size and IQ may be influenced by nutritional factors—good nutrition benefits brain-size development and intelligence.

Males vs. females

We have seen that females have smaller brains than males and that brain size is positively associated with intelligence. The possible inference that men are more intelligent than women does *not* follow—in fact, the two sexes have essentially the same mean IQ—(Mackintosh, 1998). However, the pattern of abilities differs somewhat between the sexes. Females tend to have greater verbal (language-based) abilities than males, and males have greater spatial abilities (e.g., solving jigsaw puzzles; reading maps) (Anderson, 2004). Halpern (2004) discussed interesting findings obtained from numerous Western cultures. Reading literacy was higher in girls than boys in all 25 cultures for which information was available. However, boys outperformed girls in 31 out of 32 cultures in science achievement and in 27 out of 32 cultures in mathematics achievement.

There is recent evidence that there may be important sex differences in brain structures underlying intelligence. For example, Haier et al. (2005) considered two kinds of nerve tissue in the brain: gray matter and white matter. Women had more white matter and fewer gray matter brain areas related to intelligence than did men. We don't know why this difference exists. However, it indicates the value of moving beyond simply measuring brain size towards a focus on the detailed make-up of individual brains.

In sum, the fact that even primitive measures such as brain volume predict individual differences in intelligence moderately well indicates that we can enhance our understanding of intelligence by focusing on the brain. It is probable that future research will indicate the value of focusing on the precise make-up of brain volume (e.g., percentage gray matter; percentage white matter) rather than crude measures such as brain volume.

Chapter Summary

Introduction

- Social considerations are more important in definitions of intelligence in collectivistic cultures than in individualistic ones.
- Most self-report questionnaire measures of emotional intelligence mainly assess established personality dimensions such as extraversion and neuroticism.

- Ability-based measures of emotional intelligence are promising. For example, high scorers on the MSCEIT are rated more positively for personal qualities, interpersonal sensitivity, sociability, and potential for leadership than low scorers.
- Intelligence is a good predictor of job performance, especially with complex jobs. The main reason is that more intelligent people learn job-related knowledge faster than less intelligent ones.
- Individual differences in intelligence predict health and longevity. This is in part a result of the fact that more intelligent people on average have more money and better living conditions. However, it is also relevant that more intelligent individuals have superior health literacy to less intelligent ones.

Intelligence testing

- Intelligence tests have a mean of 100 and a standard deviation of about 16.
- Useful intelligence tests are standardized and have high reliability and validity.

Theories of intelligence

- Factor theorists use factor analysis to decide on the number and nature of the factors underlying intelligence.
- Spearman proposed a two-factor theory based on a general factor and numerous test-specific factors.
- Thurstone identified seven primary mental abilities and didn't include a general factor.
- The consensual view is that the structure of intelligence is hierarchical, with the general factor at the top, fairly general factors resembling primary mental abilities at the intermediate level, and specific factors at the lowest level.
- Gardner originally proposed seven intelligences (subsequently increased to ten). These intelligences were claimed to be independent of each other, but are probably correlated to some extent.

Heredity and environment

- The effects of heredity and environment on intelligence are interdependent.
- Three types of interdependence are active covariation, passive covariation, and reactive covariation.
- On average, identical twins brought up together resemble each other in IQ more than do fraternal twins. In addition, identical twins brought up apart resemble each other nearly as much as those brought up together. These findings suggest that individual differences in intelligence depend in part on genetic factors.
- The heritability of intelligence increases with age. Adolescents and adults select their own environment to a greater extent than children, and so the impact of shared environment on intelligence decreases.
- The Flynn effect shows a strong impact of environmental factors on intelligence in many Western countries. This effect may now be decreasing.
- Children who are exposed to various risk factors in early childhood can show reductions in IQ as a consequence.

Size does matter!

- When brain size is estimated from skull size, there is generally no association with intelligence.
- When brain size is measured accurately using MRI, there is a moderate positive correlation between brain size and intelligence.
- On average males have larger brains than females, but there is no gender difference in IQ. Of potential importance, women have more white matter and fewer gray matter areas related to intelligence than men.

Further Reading

- Gottfredson, L.S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79–132. This article discusses evidence showing that individual differences in intelligence influence many aspects of everyday life.
- Mackintosh, N.J. (1998). *IQ and human intelligence*. Oxford, UK: Oxford University Press. This is an excellent book by a leading British psychologist. It stands out for providing a balanced and insightful account of human intelligence.
- Matthews, G., Zeidner, M., & Roberts, R.D. (Eds.) (2007). *The science of emotional intelligence: Knowns and unknowns. Series in affective science*. New York: Oxford University Press. This edited book provides an up-to-date account of the strengths and weaknesses of research on emotional intelligence by leading experts.
- Miele, F. (2002). *Intelligence, race, and genetics: Conversations with Arthur R. Jensen*. Boulder, CO: Westview Press. The importance of genetic factors is explored in a very readable way through discussions between Jensen (a major supporter of the hereditarian position) and the journalist Frank Miele.
- Salovey, P., & Grewal, D. (2005). The science of emotional intelligence. *Current Directions in Psychological Science*, 14, 281–285. This short article gives a good overview of development in research on emotional intelligence.
- Sternberg, R.J., & Pretz, J.E. (Eds.) (2004). *Cognition and intelligence: Identifying the mechanisms of the mind*. Cambridge, UK: Cambridge University Press. Individual differences in intelligence are considered from the perspective of cognitive psychology in this edited book with contributions by leading experts.

chapter 12

Contents

Introduction	287
Theory of psychosexual development	288
Personality assessment	291
Trait approach to personality	293
Social cognitive theory	302

What does personality look like?

12

INTRODUCTION

One of the things that gives life much of its interest and excitement (and sometimes misery and grief!) is the fact that people differ from each other in almost endless ways. Some are nearly always cheerful and friendly, others are unfriendly and depressed, and still others are aggressive and hostile. Much of this variety reflects individual differences in personality. What do we mean by “personality”? According to Child (1968, p. 83), personality consists of “the more or less stable, internal factors that make one person’s behavior consistent from one time to another, and different from the behavior other people would manifest in comparable situations.” There are *four* key words in that definition:

1. *Stable*: Personality remains fairly constant and unchanging over time.
2. *Internal*: Personality lies within us, but how we behave is determined in part by our personality.
3. *Consistent*: If personality remains constant over time, and if personality determines behavior, then we would expect people to behave reasonably consistently.
4. *Different*: When we talk of personality, we assume there are considerable individual differences leading different people to behave differently in similar situations.

How can we describe human personality? It is generally assumed that personality consists of various **traits**, which are “broad, enduring, relatively stable characteristics used to assess and explain behavior” (Hirschberg, 1978, p. 45). For example, suppose you notice that someone is talkative, smiles a lot, participates fully in social events, and has many friends. One way of explaining their behavior would be to assume they possess a high level of the trait of sociability.

Imagine you are a psychologist, and you have been given the task of designing a questionnaire to assess human personality. How many different personality traits would you put into your questionnaire? You might find yourself thinking how very different the people are from each other. This might lead you to assume there must be upwards of 20 or 30 personality traits that should be included in your questionnaire. We will be discussing this issue shortly. You will probably be surprised to discover that personality researchers’ most popular answer to the question of how many personality traits humans possess is . . . *five*!

Why do individual differences in personality traits exist? Perhaps the most obvious answer (favored by many theorists including Freud and Bandura) is that they depend in large measure on each individual’s experiences in life. For example, someone growing up in a very secure and supportive family environment will have a less anxious personality than someone growing up in a family characterized by conflicts. In fact, family

Key Term

Traits:
stable aspects of a person that influence his/her personality.

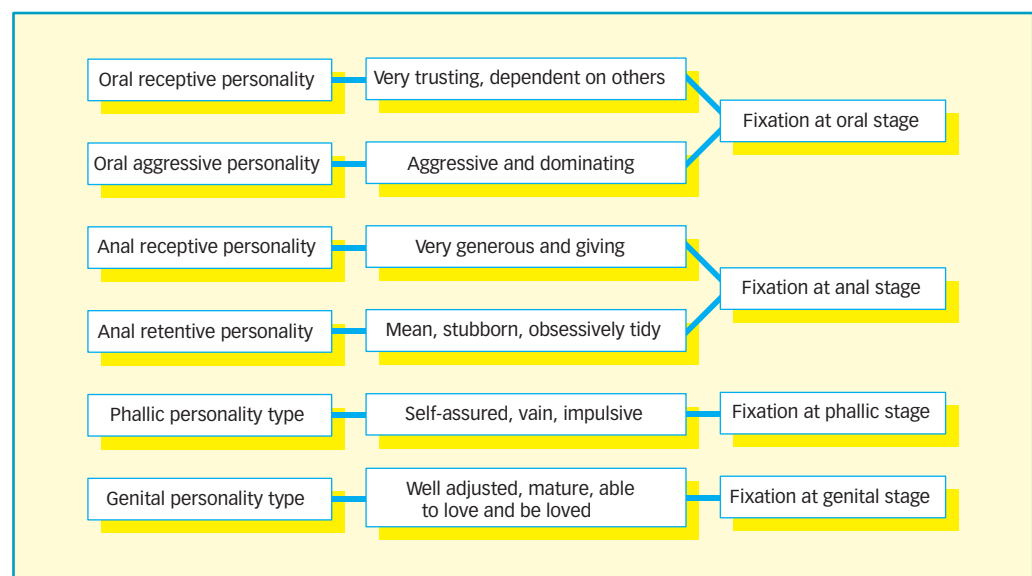
environment seems to have surprisingly little impact on adult personality. We can use data from twin studies to distinguish between shared environment (environmental factors having a similar effect on twins or siblings) and nonshared environment (environmental factors unique to a given individual). Nonshared environment accounts for about 60% of individual differences in personality, but shared environment contributes practically nothing (Plomin et al., 1997). As we will see, genetic factors (which are easily overlooked) turn out to have a fairly strong impact on individual differences in personality.

There is one final point before we embark on a discussion of theories of personality. In many countries (including the UK and the US), there are relatively few researchers who specialize in the area of personality. This helps to explain why there is a greater historical emphasis in this chapter than in most of the other chapters in this book.

THEORY OF PSYCHOSEXUAL DEVELOPMENT

According to Sigmund Freud (1856–1939), the experiences children have during the first 5 years of life are very important. Their personalities develop during that period, and adult personality depends very much on the experiences of early childhood. Freud assumed that there are several personality types arising from childhood experiences (see the figure below). According to his theory of psychosexual development, children pass through five stages:

1. *Oral stage*: This lasts up to the age of 18 months. Infants in this stage enjoy various activities involving their mouth, lips, and tongue. Children may experience problems at this stage of development (e.g., because of rapid weaning). These problems can produce adults with an oral receptive character (very dependent on other people) or an oral aggressive character (hostile and domineering).
2. *Anal stage*: This lasts between 18 and 36 months, and involves the anal area as the main source of satisfaction. This is the stage at which toilet training occurs. Children experiencing problems at this stage may become adults with an anal retentive character (mean, stubborn, and orderly) or they may become very generous and giving (anal receptive character).
3. *Phallic stage*: This stage lasts between the ages of 3 and 6. During this stage, the penis or clitoris becomes the main source of satisfaction. At about the age of 5, boys acquire the **Oedipus complex**, in which they have sexual desires for their mother and consequent fear of their father. This complex is resolved by identification with the father. A similar process in girls is based on the Electra complex (not a term used by



Key Term

Oedipus complex: the Freudian notion that boys at the age of 5 desire their mother and so become frightened of their father.

Freud), in which they desire their father. Those experiencing problems at this stage develop a phallic character. Men with a phallic character are vain and self-assured, whereas women with a phallic character fight hard for superiority over men.

4. *Latency stage*: This lasts from the age of 6 until puberty. During this stage, boys and girls experience relatively few sexual feelings and ignore each other.
5. *Genital stage*: This lasts from puberty onwards. In this stage, the main source of sexual pleasure is in the genitals. The key difference from the phallic stage is that the focus in the genital stage is on sexual pleasure with another person. Children avoiding problems during the earlier stages develop a genital character in adulthood. People with this character are mature, and can love and be loved.

Freud's stages of psychosexual development

Stage	Approximate age	Summary
Oral	0–18 months	Satisfaction from eating, sucking, etc.
Anal	18–36 months	Interest in and satisfaction from anal region
Phallic	3–6 years	Genitals become source of satisfaction
Latency	6 years old–puberty	Boys and girls spend little time together
Genital	From onset of puberty	Genitals main source of sexual pleasure

According to Freud, there can be serious consequences if a child is frustrated by receiving insufficient gratification at any psychosexual stage. There can also be serious consequences if a child receives excessive gratification at a given stage. What consequences are likely to follow? Freud argued that frustration or excessive gratification leads to **fixation**, in which basic sexual energy remains attached to that stage during adulthood. When an adult experiences major problems, he/she will show **regression**, with their behavior resembling how they behaved when they were children. Adults typically regress to the psychosexual stage at which they fixated as children. An important implication of this part of the theory is that adult mental disorders typically have their origins in childhood problems and difficulties (see Chapter 21). An account of Freud's overall theoretical approach is given in Chapter 2.

Useful mnemonic

To help you remember Freud's stages of psychosexual development, the following mnemonic is made from the initial letter of each stage: Old Age Pensioners Love Greens!

FINDINGS

There is reasonable support for the *general* approach taken by Freud, in that childhood experiences clearly influence the development of human personality (see Westen, 1998, for a review). However, this is *not* the same as arguing that there is support for Freud's specific theoretical assumptions. Mickelson, Kessler, and Shaver (1997) considered a random sample of more than 5000 adults. Parental loss or separation in childhood was associated with low attachment security and high ratings of insecure attachment in adulthood. Those adults who had experienced serious traumas in childhood (e.g., sexual abuse; severe neglect) were more likely than other adults to have anxious attachments to other people.

Franz et al. (1996) also reported evidence suggesting that childhood experiences have a long-term impact. Adult levels of depression at the age of 41 were predicted well by parental coldness when they were only 5 years old. In addition, an overall measure of difficult childhood experiences (e.g., divorce; frequent moves; loss) predicted depression in middle age.

There is a real problem with interpreting most findings showing a relationship between childhood experiences and adult problems. What we have is only correlational evidence of an association, and correlations can't be used to prove causes. For example, suppose an adult with serious emotional problems had parents who made her childhood miserable. It is possible that the unhappy childhood experiences helped to create the

Key Terms

Fixation:
in Freudian theory, spending a long time at a given stage of psychosexual development.

Regression:
in Freudian theory, returning to an earlier stage of psychosexual development when someone is highly stressed.

Case Study: *Little Hans*

When Hans was 3 years old, he began to exhibit a lot of interest in his penis, which caused his mother to threaten to cut it off. When he was about 5, he saw a horse-drawn van tip over on its side. This caused him to develop a fear of horses, and to refer to “black things around horses’ mouths and the things in front of their eyes.” As a result of his fear of horses, Hans was unwilling to leave the safety of his own home.

Freud’s diagnosis was that Little Hans suffered from an Oedipus complex. According to Freud, he was sexually attracted to his mother, but was very frightened his father would punish him for this. The fear that Hans had for his father turned into a fear of horses, with Freud arguing that horses’ black muzzles and blinkers resembled the mustache and glasses of Hans’ father. The fact that Hans’ fear of horses caused him to remain at home had the significant advantage that he could spend more time with his mother, to whom allegedly he was sexually attracted.

emotional problems in adulthood. However, it is also possible that genetic factors common to the parents and to their child are responsible. Kendler et al. (1992) carried out a twin study, and tried to remove statistically the influence of genetic factors from their findings. They still found that parental loss through separation (e.g., divorce) was associated with depression in adult life, whereas parental loss through death was not. In a similar twin study, Kendler, Neale, et al. (1996) found that childhood parental loss through separation was associated with a tendency to alcoholism. This association was still present when the impact of genetic factors was removed.

There is support for some of the personality types identified by Freud. For example, Kline and Storey (1977) discovered that the three main characteristics of the anal retentive character (stubbornness, meanness, orderliness) are often found together in the same person. There is also evidence that some people possess all the features of the oral aggressive character (Kline, 1981). However, evidence for the existence of some of the personality types suggested by Freud does *not* show the theory is correct. Freud assumed that these personality types depend on

specific childhood experiences, but there is practically no convincing evidence that that is the case.

The evidence that Freud himself reported to support his own theoretical approach was very flimsy. It consisted mainly of about a dozen case studies, most of which are not at all convincing. For example, consider the famous case of Little Hans (see the Case Study above), which was claimed to illustrate the development of the Oedipus complex.

Most people (including the author) are totally unconvinced by Freud’s account of Little Hans. Indeed, it seems to say more about Freud than about Little Hans! There is very little evidence that Hans desired sexual contact with his mother or that he was very frightened of his father. In addition, the idea that Hans’ great sexual excitement somehow turned into a state of high anxiety is fanciful.

Evaluation

- + Freud put forward the first systematic theory of personality.
- + Adult personality depends in part on the experiences of early childhood.
- + At least some of the personality types suggested by Freud do seem to exist.
- + The notion that individuals with certain types of personality are more vulnerable than others to the development of mental disorders is both powerful and convincing (see Chapter 21).
- + It is very hard to test the hypothesis that early childhood experiences have actually determined adult personality many years later.
- Adult personality depends more on heredity and on the experiences of adolescence and adulthood than assumed by Freud (discussed later in the chapter).
- Freud’s stage-based theory implies that personality development occurs in a more orderly way than is actually the case.
- “Psychoanalysts seriously shot themselves in the foot by never evolving from case study methods [based on individual patients] as their primary mode of knowledge generation and hypothesis testing” (Westen & Gabbard, 1999).

PERSONALITY ASSESSMENT

How can we assess personality? Four major kinds of personality tests have been developed:

1. self-report questionnaires
2. ratings
3. objective tests
4. projective tests.

We will shortly consider all four kinds of personality tests. As with intelligence tests (see Chapter 11), useful personality tests possess three characteristics:

1. **Reliability:** This means that the test produces consistent results. A test's reliability can be assessed by giving it on two occasions to the same individuals.
2. **Validity:** This means that the test measures what it is supposed to be measuring. The most important type of validity is **concurrent validity**, in which test scores are correlated with a relevant external criterion or measure of the underlying construct the test is supposed to be measuring. For example, suppose we have devised a measure of trait anxiety, a personality characteristic relating to the experience of anxiety. The scores of patients with anxiety disorders could be used as an external criterion—such patients should score much higher than most healthy individuals on the test. If we adopt the same approach, but assess the criterion at some point in time *after* the assessment of personality, this is known as **predictive validity**.
3. **Standardization:** This involves giving the test to large, representative samples of people so the significance of an individual's score on the test can be evaluated. For example, a score of 19 on a test of extraversion is meaningless on its own. However, it becomes meaningful if we know that only 10% of the population have such a high score.

QUESTIONNAIRES

Almost everyone in our society has filled in several personality questionnaires. What happens is that you have to decide whether various statements about your thoughts, feelings, and behavior are true. Here are some sample items: Do you tend to be moody? Do you have many friends? Do you like to be involved in numerous social activities?

The questionnaire-based approach is the most popular way of assessing personality. One of its advantages is that it is easy to use. Another advantage is that an individual knows more about himself/herself than do other people.

The most obvious problem with questionnaires is that individuals may fake their responses. Such faking typically takes the form of **social desirability bias**. This is the tendency to respond to questionnaire items in the socially desirable (but inaccurate) way. This bias is especially likely to be present in personnel selection. As Cook (1993, p. 144) pointed out, “No one applying for a sales job is likely to say true to ‘I don’t much like talking to strangers,’ nor is someone trying to join the police force likely to agree that he/she has pretty undesirable acquaintances.”

How can we deal with social desirability bias? The most common method is to use a Lie scale consisting of items where the socially desirable answer is very unlikely to be the honest answer (e.g., “Do you ever gossip?”; “Do you always keep your promises?”). If someone answers most of the questions on the Lie scale in the socially desirable direction, it is assumed they are faking their responses. Of course, this is unfair on the small minority of genuinely saintly people in the population!

Key Terms

Concurrent validity: assessing **validity** by correlating scores on a test with some currently available relevant criterion.

Predictive validity: assessing **validity** by correlating scores on a test with some future criterion.

Social desirability bias: the tendency to provide socially desirable rather than honest answers on questionnaires and in interviews.



The social desirability bias is the desire to give false answers in order to be deemed socially desirable. This could have ramifications in terms of the accuracy of a job application, for example!

In spite of the possibility of social desirability bias, most well-known personality questionnaires possess high reliability and low to moderate validity. Why is validity typically much lower than reliability? Remember that validity is typically assessed by correlating questionnaire scores with some external criterion (e.g., extraversion with number of friends). The problem in a nutshell is that there is no such thing as a perfect external criterion. For example, most people would agree that extraverts tend to have more friends than introverts. However, no one seriously believes that an individual's level of extraversion is the *only* factor determining how many friends he/she has. The imperfections of all external criteria serve to reduce the validity of personality questionnaires, as do inadequacies within the questionnaires themselves.

What is the best way of assessing validity? The answer is consensual validity. **Consensual validity** involves comparing two kinds of information:

1. Self-report questionnaire scores obtained from participants.
2. Ratings of those participants by those who know them well (e.g., friends; relatives) for the same aspect(s) of personality.

McCrae and Costa (1990) obtained promising findings based on the assessment of consensual validity. Self-report questionnaire data on five major personality factors were obtained, and ratings on the same five factors were collected from spouses and peers. The average correlation between self-report scores and spouse rating scores was $+.56$, and it was $+.50$ between self-report scores and peer rating scores. These correlations are reasonably high and indicate fairly good consensual validity.

RATINGS

As we have just discussed, ratings involve observers providing information about other people's behavior. There are various ways this can be done. For example, raters can simply be given a personality questionnaire and asked to complete it as they think a friend of theirs would have done. Alternatively, they can be given a list of different kinds of behavior (e.g., "initiates conversation"), rating their ratees (those being rated) on those aspects of behavior.

Ratings have some advantages over self-report questionnaires. In particular, the problem that people filling in a questionnaire may distort their responses to present a favorable impression (social desirability bias) doesn't apply to observers' ratings. However, the rating approach poses problems of its own. First, the items of behavior to be rated may be interpreted somewhat differently by different raters. For example, an item such as, "behaves in a friendly way towards others," might be interpreted to mean much more interaction with other people by a very sociable rater than by an unsociable one. Second, most raters have observed other people in only some of the situations in which they find themselves in everyday life. Someone who appears distant and aloof at work may relax and be very friendly outside the work environment. The partial view a rater has of his/her ratees may lead to inaccurate assessment of the ratees' personalities.

In spite of the limitations of using ratings, they typically possess high reliability. What about validity? That has been addressed by means of consensual validity, in which rating data are correlated with self-report data. As we saw earlier, consensual validity is moderately high (McCrae & Costa, 1990).

Key Terms

Consensual validity: the extent to which scores on a self-report questionnaire correlate with scores obtained from ratings provided by other people.

Objective tests: a method of assessing personality under laboratory conditions in an unobtrusive way.

OBJECTIVE TESTS

Once upon a time, there was much interest in objective tests. **Objective tests** involve measuring behavior under laboratory conditions with the participants not knowing what the experimenter is looking for. For example, asking people to blow up a balloon until it bursts is a measure of timidity, and the extent to which people sway when standing on tiptoe is a measure of anxiety.

There was enthusiasm for objective tests several decades ago because they appeared free from the problems of deliberate distortion that can influence the responses on self-report questionnaires. The reason is that the participants aren't aware that their personality is being assessed, and so have no particular motivation to respond one way

rather than another. However, it is often difficult to know what any given objective test is actually measuring, and the results are often much influenced by apparently minor changes in procedure. Most objective tests have low reliability and validity, and so are of very limited value (Cooper, 2002).

PROJECTIVE TESTS

The final form of personality assessment is by **projective tests**. People are given an unstructured task to perform (e.g., devising a story to fit a picture; describing what can be seen in an inkblot). The underlying notion is that people confronted by such ambiguous stimuli and unstructured tasks will reveal their innermost selves. Most users of projective tests favor the psychodynamic approach pioneered by Sigmund Freud, and such tests are used mainly for clinical purposes.

One of the best-known projective tests is the Rorschach Inkblot Test introduced by the Swiss psychologist Hermann Rorschach in 1921. The standard form of the Rorschach test involves presenting 10 inkblots. Participants suggest what each inkblot might represent, and indicate which part of the inkblot formed the basis of their response. Another well-known projective test is the Thematic Apperception Test developed by Henry Murray (Morgan & Murray, 1935). People are presented with various pictures. For each picture, they have to say what is happening, what led up to the situation depicted, and what will happen next. The stories produced are interpreted in the attempt to identify the individual's underlying motivational conflicts.

Projective tests are generally low in both reliability and validity (Eysenck, 1994). There are two main reasons for this state of affairs. First, the unstructured nature of the tests means the participants' responses are determined by their current moods or concerns rather than by deep-rooted characteristics. Second, the very subjective nature of the interpretation of responses on projective tests means that much depends on the expertise of the person carrying out the interpretation. Nonexpert interpretation reduces the validity of the tests, and the subjectivity of interpretation reduces their reliability.

We will consider the Rorschach Inkblot Test in more detail to see some of its limitations. Three aspects of a participant's responses are used to interpret their meaning: content; location; and determinants. Content refers to the nature of what is seen by the participant, location refers to the part of the inkblot used to produce the response, and, finally, the determinants are the inkblot's characteristics (e.g., color; form) influencing the choice of response. Most Rorschach experts argue that location and determinants are more informative than content. Alas, the evidence suggests that content possesses more validity than location or determinants (Eysenck, 1994)!



An example of a Rorschach inkblot.

TRAIT APPROACH TO PERSONALITY

For many years, there was a major controversy between two groups of personality researchers. In one group (associated especially with Cattell), it was assumed that many personality traits resemble each other, because there is no reason to assume that personality is tidily organized into unrelated traits. According to members of this group, personality traits often resemble other traits, and we need to identify numerous personality traits to provide a comprehensive account of personality. In the other group (associated especially with H.J. Eysenck), it was assumed that major personality traits are unrelated or independent of each other. In more technical terms, traits should *not* correlate or be associated with each other. For good statistical reasons, those making these latter assumptions identify far fewer traits than those belonging to the former group. For example, Cattell claimed that there were 16 personality traits whereas H.J. Eysenck argued that there were only three! Ask yourself which group your money is on before reading further . . .

CATTELL'S APPROACH

Perhaps the greatest problem faced by trait theorists trying to devise a personality questionnaire is to try to ensure they include *all* of the important personality traits. How

Key Term

Projective tests: a method of assessing personality in which people are given an unstructured test to perform (e.g., describing inkblots).

can this goal be achieved? Raymond Cattell came up with the ingenious answer of using the **fundamental lexical hypothesis**, according to which each language contains words describing all of the main personality traits.

Cattell’s use of the fundamental lexical hypothesis led him to the work of Allport and Odbert (1936), who had identified 4500 words used to describe personality. These 4500 words were reduced to 160 trait words, in part by eliminating words having the same meaning and by removing unfamiliar words. Cattell (1946) then added 11 traits from the personality literature in psychology, producing a total of 171 trait names, which were claimed to cover almost everything of importance in the area of personality.

Cattell still found himself with an unwieldy number of personality traits. As a result, he used findings from several previous rating studies to identify traits that resembled each other. This allowed him to eliminate some more traits, leaving him with 35 traits. He called them **surface traits**, because they were easily observable. Cattell carried out rating studies in which raters assessed people they knew well. These studies suggested that there are about 16 **source traits**, which are basic traits underlying the surface traits.

Cattell argued correctly that any given method of assessing personality has limitations. Since different methods have different limitations, we can in principle best work out the main personality traits by combining information from various methods. Accordingly, Cattell made extensive use of three methods of assessing personality:

- 1. *Life (L) data*: Observers’ ratings of other people’s behavior.
- 2. *Questionnaire (Q) data*: Self-report questionnaires.
- 3. *Objective test (T) data*: Careful assessment of personality under controlled conditions (e.g., measuring anxiety by seeing how much people sway when standing on tiptoe).

Cattell found that there was reasonable similarity in the personality traits identified using the questionnaire and life or rating methods. However, the findings with the objective test method were very different. This method has largely been abandoned because it produces very inconsistent findings. For example, you would probably sway much more when standing on tiptoe shortly after a visit to the pub than when completely sober!

16PF

Cattell’s work suggesting that there are 16 source traits led him to develop the Sixteen Personality Factor Questionnaire, generally known as the 16PF. As the name indicates, it was intended to assess 16 personality factors, some relating to intelligence and social attitudes rather than to personality in a narrow sense. The construction of the 16PF was based on Cattell’s assumption that personality traits are often associated or correlated with each other, so many of his 16 factors are fairly closely related to each other (see the box on the left).

In spite of the 16PF’s massive popularity, it provides an inadequate assessment of personality. All systematic factor analyses of this test have shown that it doesn’t measure anything like 16 different personality traits. For example, Barrett and Kline (1982) gave the 16PF to almost 500 people. They then carried out five different factor analyses on their data, some precisely in line with Cattell’s recommendations. Barrett and Kline obtained between seven and nine factors in each factor analysis, and these factors generally didn’t relate closely to those proposed by Cattell.

What is going on here? A key problem is that several of Cattell’s factors are very similar to each other. For example, the following factors assessing anxiety all correlate very highly with each other: placid–apprehensive; relaxed–tense; affected by feelings–emotionally stable. Any thorough factor analysis reveals that these three factors are simply too similar to each other to be distinguished.

Key Terms

Fundamental lexical hypothesis:
the assumption that dictionaries contain words describing all of the main personality traits.

Surface traits:
personality traits that are readily observable and are related to underlying source traits.

Source traits:
personality traits underlying the more superficial surface traits.

The factors of Cattell’s 16PF

Remember that each pair represents a continuum.

Reserved	Outgoing
Less intelligent	More intelligent
Affected by feelings	More emotionally stable
Humble	Assertive
Sober	Happy-go-lucky
Expedient	Conscientious
Shy	Venturesome
Tough-minded	Tender-minded
Trusting	Suspicious
Practical	Imaginative
Forthright	Shrewd
Placid	Apprehensive
Conservative	Experimenting
Group-dependent	Self-sufficient
Casual	Controlled
Relaxed	Tense

Evaluation

- + Cattell's notion of using the fundamental lexical hypothesis to assist in uncovering all the main personality traits is a valuable one.
- + Cattell's attempt to combine information from several methods (questionnaires, ratings, objective tests) was thorough and systematic.
- There are only about eight different personality traits in the 16PF, and so Cattell's main questionnaire is badly flawed.
- Cattell's approach was not very theoretical or explanatory. As Cattell (1957, p. 50) admitted, "I have always felt justifiably suspicious of theory built much ahead of data."

H.J. EYSENCK'S APPROACH

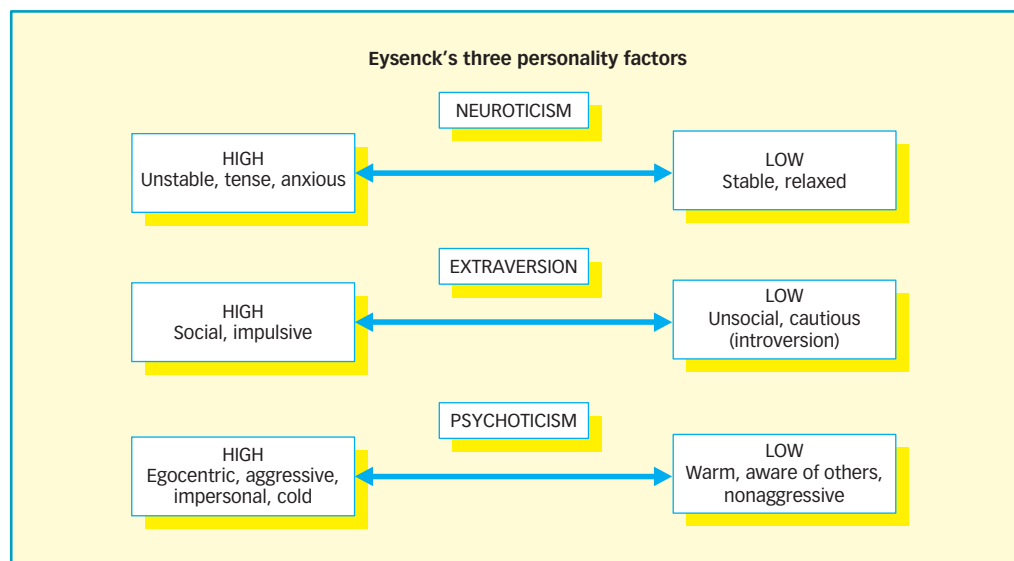
H.J. Eysenck's approach was very different to that of Cattell. He argued that the best strategy is to focus on a small number of *independent* personality traits or factors, all of which are entirely separate from each other. More specifically, he identified three major traits or "superfactors" (see the figure below), all of which are assessed by the Eysenck Personality Questionnaire (EPQ):

- **Extraversion:** Those scoring high on extraversion (extraverts) are more sociable and impulsive than those scoring low (introverts).
- **Neuroticism:** Those scoring high on neuroticism are more anxious and depressed than those scoring low.
- **Psychoticism:** Those scoring high on psychoticism are aggressive, hostile, and uncaring.

You may well feel that there must be more to personality than these three factors (and you're probably right!). However, many aspects of personality can be understood as consisting of combinations of two (or even all three) of these factors. For example, there is no trait of "optimism," but individuals high in extraversion and low in neuroticism are typically optimistic.

Where do individual differences in extraversion, neuroticism, and psychoticism come from? According to H.J. Eysenck (1982, p. 28), "genetic factors contribute something like two-thirds of the variance in major personality dimensions."

How do genetic factors produce individual differences in personality? According to H.J. Eysenck (e.g., 1967), heredity influences the responsiveness of parts of the



Key Terms

Extraversion: a personality factor based on sociability and impulsiveness.

Neuroticism: a personality factor based on negative emotional experiences (e.g., anxiety and depression).

Psychoticism: a personality factor based on aggression, hostility, and a lack of caring.



Jack Nicholson in *The Shining*—perhaps this character would score highly on Eysenck's psychoticism personality factor—he is aggressive, impulsive, impersonal, cold, and lacking in empathy and concern for others.

Personality dimensions

Personality dimensions such as extraversion are at one end of a continuum, with their opposite at the other end. Personality testing does not simply determine that a person is either extraverted or introverted, but places them at the relevant point on the continuum, showing their degree of extraversion/introversion, as in the diagram below. This person is slightly more extraverted than introverted.



In order to determine whether individual differences in personality are due to genetic factors or environment, a number of studies using twins have been carried out. Resulting evidence suggests that 40% of individual differences in personality are due to genetic factors.

Key Term

Visceral brain:

a brain system including the hippocampus, septum, and hypothalamus claimed by H.J. Eysenck to be most responsive in individuals high in neuroticism.

physiological system. Introverts were assumed to have a higher level of cortical arousal (activity in the brain) than extraverts. As a result, they are in danger of becoming over-aroused, and so tend to prefer reading books to going to exciting parties. In contrast, extraverts generally have a low level of brain activity, and so seek stimulating situations to prevent themselves being bored. Those high in neuroticism are supposed to have greater activity than those low in neuroticism in the **visceral brain**, which consists of several parts of the brain (hippocampus, amygdala, cingulum, septum, and hypothalamus). The physiological differences between individuals high and low in psychoticism remain shrouded in mystery.

Findings

We will consider each of H.J. Eysenck's main assumptions in turn. The notion that extraversion, neuroticism, and psychoticism are all major personality traits has been investigated numerous times (see H.J. Eysenck & M.W. Eysenck, 1985, for a review). The evidence indicates strongly that extraversion and neuroticism are very important but not psychoticism. For example, Saville and Blinkhorn (1981) studied Cattell's 16PF questionnaire to find out which independent factors it contained. They found that the 16PF largely measures extraversion and neuroticism (but *not* psychoticism). These are very important findings, because they suggest that the approach based on associated or correlated factors works less well than one based on independent or uncorrelated factors.

H.J. Eysenck's assumption that genetic factors account for 67% of individual differences in personality is an exaggeration. The best approach to this issue involves studying identical twins (known technically as monozygotic twins) and fraternal twins (dizygotic twins). Identical twins share 100% of their genes, whereas fraternal twins share only 50% of their genes. If genetic factors are important, then identical twins should be more similar in personality than fraternal twins. More precise conclusions can be drawn if we consider twin pairs brought up apart as well as twin pairs brought up together. As you can imagine, it is not exactly easy to find many pairs of identical twins brought up apart!

The most thorough attempt to study the influence of genetic factors on extraversion and neuroticism was reported by Pedersen, Plomin, McClearn, and Friberg (1988). Amazingly, they studied 95 identical twin pairs brought up apart, 220 pairs of fraternal twins brought up apart, as well as numerous identical and fraternal twin pairs brought up together. They found that 31% of individual differences in neuroticism and 41% of individual differences in extraversion were a result of genetic factors. So far as psychoticism is concerned, Zuckerman (1989) reviewed four twin studies, and concluded that 40% of individual differences in psychoticism come from heredity.

We turn now to evidence concerning the physiological bases of the personality factors. The only

one for which there is much support is extraversion, for which it will be remembered that H.J. Eysenck claimed that introverts are generally more cortically aroused than extraverts. One way of testing this hypothesis is by using electroencephalography (EEG), which provides a measure of brain-wave activity. Gale (1983) considered 33 EEG studies reporting a total of 38 experimental comparisons. Extraverts were significantly less cortically aroused than introverts in 22 comparisons, whereas introverts were significantly less aroused than extraverts in 5 comparisons. Introverts and extraverts didn't differ in the remaining 11 cases. Thus, introverts tend to be more cortically aroused than extraverts, but this is often not the case.

H.J. Eysenck (1967) argued that individual differences in neuroticism depend on the level of activity within the “visceral brain.” This hypothesis has been tested by taking various indirect measures (e.g., heart rate; skin conductance) from individuals high and low in neuroticism (or the closely related personality factor of trait anxiety) in stressful and nonstressful conditions. There is very little support for the hypothesis. Fahrenberg (1992, pp. 212–213) concluded his review of the literature as follows: “Over many decades research has failed to substantiate the physiological correlates that are assumed for emotionality [neuroticism] and trait anxiety. There is virtually no distinct finding that has been reliably replicated across studies and laboratories.”



A young woman (background) undergoing an EEG examination. The EEG test records the electrical activity of the brain via small electrodes attached to the scalp. Gale (1983) reviewed studies using this technique to test the hypothesis that introverts are more cortically aroused than extraverts.

Evaluation

- + It has proved more useful to identify a small number of unrelated or uncorrelated personality traits than a larger number of correlated ones (more evidence is discussed below).
- + Extraversion and neuroticism are major personality traits or factors.
- + H.J. Eysenck made a thorough attempt to explain the mechanisms underlying individual differences in his three personality factors.
- The role of genetic factors in determining individual differences in personality is much less than was claimed by H.J. Eysenck.
- There is little support for the physiological bases of personality proposed by H.J. Eysenck.
- Psychoticism is not a major personality trait or factor. It is also poorly named, being more closely related to psychopathy or antisocial personality than to psychosis (Eysenck, 1994).

BIG FIVE

For the past 20 years, the most popular view concerning the structure of human personality is that there are five major personality traits commonly referred to as the Big Five. It is assumed that these five factors are independent or uncorrelated. Theorists differ somewhat in terms of the exact traits or factors identified, but there is a fairly high level of consensus. The most influential version of the Big Five approach is that of McCrae and Costa (1985), who identified the following five factors (the first letters of which conveniently form the word OCEAN):

Openness (curious, imaginative, creative)
 Conscientiousness (hard-working, ambitious, persistent)
 Extraversion (sociable, optimistic, talkative)
 Agreeableness (good-natured, cooperative, helpful)
 Neuroticism (anxious, insecure, emotional).

Costa and McCrae (1992) produced the NEO-PI Five-Factor Inventory to measure the above five factors. What causes individual differences in these personality factors? It

is assumed that genetic factors play a significant role in determining individual differences in all five factors (e.g., McCrae & Costa, 1999), although environmental factors are also assumed to be important.

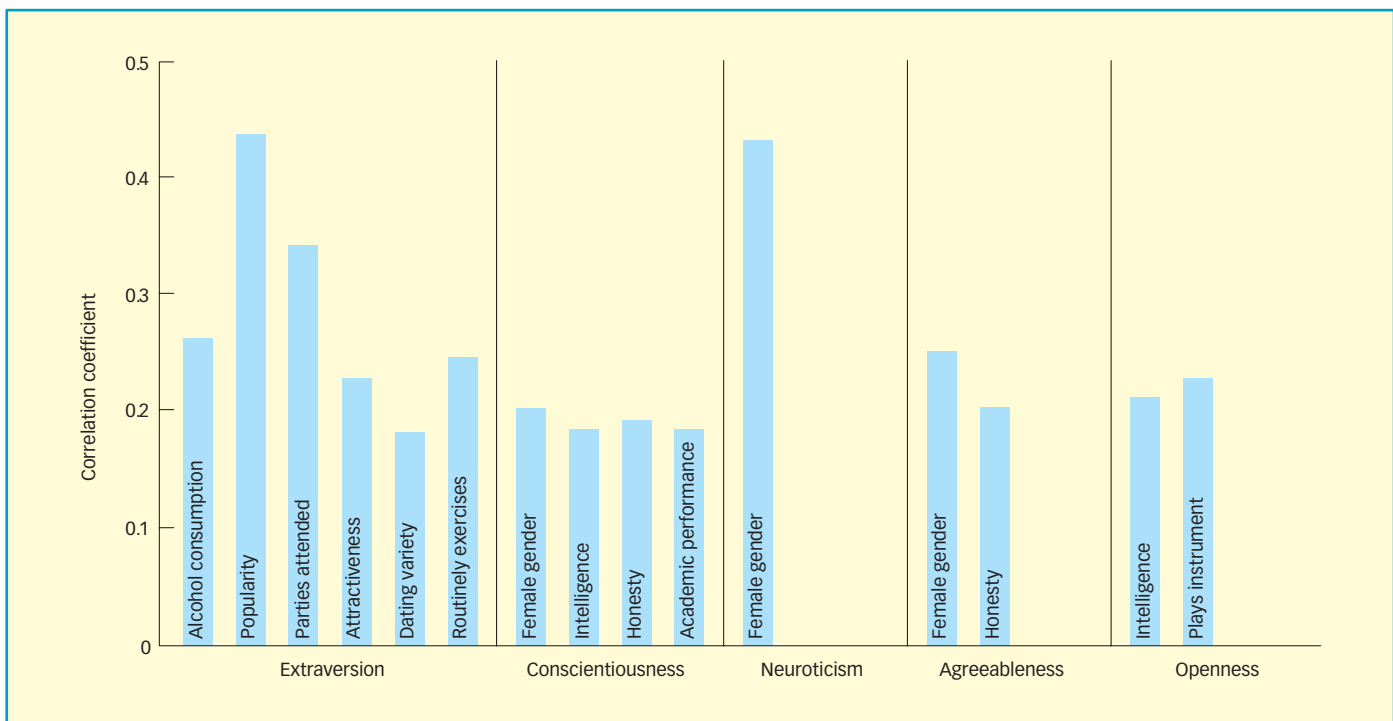
Findings

Five traits or factors closely resembling those put forward by McCrae and Costa (1985, 1990) have been found numerous times. For example, Goldberg (1990) collected more than 1000 words describing personality. Students produced self-descriptions based on these terms, and the data were factor analyzed in 10 different ways. The consistent finding was that five factors were extracted: emotional stability (opposite of neuroticism); agreeableness; conscientiousness; surgency (very similar to extraversion); and intellect (resembling openness).

Goldberg (1990) carried out a second study based on 479 common trait terms in which he obtained self-report and rating data. He found strong evidence for the same five factors in both kinds of data, and came to the following conclusion: “Analyses of any reasonably large sample of English trait adjectives in either self- or peer-descriptions [ratings] will elicit a variant of the Big-Five factor structure” (Goldberg, 1990, p. 1223).

There is reasonable evidence that questionnaires of the Big Five factors are valid. For example, McCrae and Costa (1990) compared self-report measures on each of the five factors with ratings by their marriage partners. All the correlations between self-report and rating data were moderately high, ranging from $+.53$ for neuroticism to $+.59$ for openness.

If the Big Five factors are important, then scores on these factors should correlate significantly with various external criteria. Paunonen (2003) found precisely that, and his findings based on averaging across four samples of students are shown in the figure below.



Correlations between the Big Five factors and several external criteria. Based on data in Paunonen (2003).

The assumption that the Big Five are all independent of each other and so are uncorrelated has been tested. The evidence indicates that this assumption is incorrect. For example, Costa and McCrae (1992) reported that the factors of neuroticism and conscientiousness correlated $-.53$ with each other, and that extraversion and openness correlated $+.40$. Thus, some factors are not nearly as separate from each other as they should be.

Key Study

Loehlin, McCrae, Costa, and John (1998): Big Five and genetic factors

The early research of Costa, McCrae, and other researchers showed that we can describe the structure of human personality reasonably well in terms of the five factors of openness, conscientiousness, extraversion, agreeableness, and neuroticism. That is a considerable achievement. However, it is important to move from *description* to *explanation*. For example, it is important to explain why there are individual differences with respect to all five personality factors. Paul Costa and Robert McCrae have consistently assumed that genetic factors help to account for individual differences in the Big Five personality factors, and Loehlin et al. (1998) carried out a thorough study to test this assumption.

As discussed earlier in the chapter, twin studies have been used to assess the role of genetic factors in determining individual differences in the Big Five factors. Since monozygotic or identical twins share 100% of their genes but dizygotic or fraternal twins share only 50% of their genes, then identical twins should resemble each other more in personality than fraternal twins if genetic factors are important. There are two kinds of environmental influences: (1) shared environment refers to the common influences within a family making children resemble each other; and (2) nonshared environment refers to all those influences unique to any given child. Loehlin et al. decided to estimate the importance of each kind of influence in producing individual differences in personality.

Loehlin et al. (1998) administered various self-report questionnaires and other measures to 807 pairs of twins, 490 of whom were identical and 317 were fraternal. What they found was as follows: “51% to 58% of individual difference variation along the Big Five dimensions is genetic in origin, 42% to 49% is due to experience unique to the individual, to temporary situational factors, and to gene–environment interaction, and none is due to effects of environment shared by the twins” (Loehlin et al., 1998, p. 447). The detailed findings are shown in the table below, with all figures indicating the percentage of individual differences in each personality factor attributable to a genetic or specific environmental source:

Factor	Genetic variance	Shared environmental variance	Nonshared environmental variance
Extraversion	57%	0%	44%
Agreeableness	51%	0%	49%
Conscientiousness	52%	0%	48%
Neuroticism	58%	0%	42%
Openness	56%	0%	44%

Discussion points

1. In what ways do you think the environments of identical twins would be more similar than the environments of fraternal twins?
2. How would you account for the differences in personality between identical twins?

KEY STUDY EVALUATION

The findings provide strong support for the assumption that genetic factors have a substantial influence on individual differences in personality. It is impressive that this is the case for all Big Five factors. The findings also indicate that similarities in personality between twins do not depend at all on shared environmental influences (e.g., within the family).

The study can be criticized in various ways. First, it is possible that the environments experienced by identical twins were on average more similar than those experienced by fraternal twins. If so, this may mean that the influence of

genetic factors is exaggerated in the figures. Second, the kinds of environmental factors influencing personality were not identified. However, Torgersen and Janson (2002) found that differences in personality in identical twins depended in part on differences in their reactions to childhood and adolescent stressors. Third, the fact that twins differ from nontwins in having a sibling of the same age may mean that we cannot safely generalize from twins to nontwins.

Evaluation

- + The Big Five personality traits have been obtained repeatedly in self-report and rating data, making the Big Five approach the dominant one in personality research.
- + Genetic factors are of importance in determining individual differences in all of the Big Five factors.
- + Each of the Big Five factors predicts important real-world behavior (Paunonen, 2003; see Ozer and Benet-Martinez, 2006, for a review).
- Some of the Big Five factors correlate with each other, and thus are not independent.
- There is a suspicion that some of the Big Five factors (e.g., openness; conscientiousness) may be less important than others.
- The Big Five approach is mostly descriptive, and fails to provide an adequate explanation of the processes underlying the various factors. As McCrae and Costa (1999, p. 147) admitted, “Shouldn’t a five-factor theory explain why there are five factors and not six? And why these factors and not others?”

CROSS-CULTURAL PERSPECTIVES

The “old” approach to studying personality across different cultures simply involved giving questionnaires developed in Western cultures to individuals in various non-Western cultures. For example, Steel and Ones (2002) discussed findings from the Eysenck Personality Questionnaire obtained from 40 different countries (see the box on the opposite page). The fact that we can use Western questionnaires all around the world doesn’t prove that the structure of personality is the same everywhere. However, detailed analyses of scores on the Eysenck Personality Questionnaire in 34 countries suggested that the same three personality factors were clearly present in all of them (Barrett, Petrides, Eysenck, & Eysenck, 1998).

McCrae, Terracciano, et al. (2005) collected *rating* data for the Big Five personality factors in 50 cultures. The same five factors within the same overall structure were found in nearly all of the cultures studied, with the major exceptions of Botswana and Nigeria. Thus, the structure of personality is similar in the great majority of cultures regardless of whether self-report (Barrett et al., 1998) or rating data (McCrae et al., 2005) are used.

The “new” cross-cultural approach involves using indigenous personality measures (those developed in the culture being studied) as well as standard Western measures. When that is done, it often turns out that some aspects of personality are culture specific. For example, Cheung and Leung (1998) administered the Chinese Personality Assessment Inventory and the Big Five items in China and Hong Kong. There was evidence for four of the Big Five factors but not for openness. In addition, they found a Chinese tradition factor having no relationship with the Big Five.

There are more profound differences among cultures than has been suggested so far. It has often been claimed that we should distinguish between individualistic and collectivistic cultures (see Chapter 1). The former emphasize personal responsibility and the latter focus on group obligations. The notion of semi-permanent personality traits determining behavior seems less applicable in collectivistic cultures, in which individuals are supposed to fit in flexibly with group expectations. Norenzayan, Choi, and Nisbett

Mean scores for psychoticism, extraversion, neuroticism, and lie scales for 40 countries or regions

Country or region	Psychoticism	Extraversion	Neuroticism	Lie	Sample size
Australia	6.96	19.31	15.48	7.58	654
Bangladesh	4.24	19.05	12.29	19.15	1075
Brazil	3.99	17.63	14.20	17.93	1396
Bulgaria	4.17	18.60	14.96	15.12	1038
Canada	4.28	18.05	12.77	13.92	1257
China	6.79	13.75	14.50	20.41	1000
Czechoslovakia	9.14	19.52	14.09	11.47	1912
Egypt	4.40	18.57	17.36	21.37	1330
Finland	4.90	16.26	14.60	11.57	949
France	5.49	17.75	15.09	14.59	811
Germany FR	6.23	18.40	13.68	10.96	1336
Greece	5.49	20.40	18.32	16.61	1301
Hong Kong	7.05	16.73	14.61	14.37	732
Hungary	3.86	16.57	14.58	15.63	962
Iceland	3.52	19.19	13.90	10.53	1144
India	8.17	22.80	16.26	18.38	981
Iran	4.52	17.69	16.05	17.13	624
Israel	3.60	22.62	8.51	17.39	1050
Italy	4.43	18.37	16.66	16.89	802
Japan	4.80	16.50	16.78	9.62	1318
Korea	4.97	16.49	18.71	15.74	1200
Lebanon	2.30	19.26	14.17	20.30	1239
Lithuania	5.01	16.45	15.10	17.50	1404
Mexico	4.49	20.63	14.15	15.41	988
Netherlands	2.88	17.36	11.52	16.09	876
Nigeria	3.58	24.50	9.43	17.76	430
Norway	2.22	18.65	10.33	11.68	802
Portugal	2.49	18.94	15.27	14.12	1163
Puerto Rico	4.43	21.01	14.15	17.01	1094
Romania	3.51	18.45	13.31	17.14	1014
Russia	3.41	16.55	18.04	14.18	1067
Sicily	5.89	17.36	17.16	16.42	785
Singapore	4.36	17.42	13.02	16.32	994
Spain	2.97	17.11	16.24	15.81	1030
Sri Lanka	4.27	18.67	12.09	20.86	1027
Uganda	6.06	19.44	15.78	13.56	1476
United Kingdom	3.84	18.03	14.97	12.11	1198
United States	3.32	21.53	15.20	9.46	879
Yugoslavia	7.46	17.32	14.37	17.94	971
Zimbabwe	6.50	19.75	14.65	15.20	838

Data supplied by Steel and Ones (2002).

(1999) found that people in Western cultures regard personality traits as stable, whereas East Asians regard traits as much more flexible and changeable. In line with this analysis, personality traits don't predict behavior as well in collectivistic cultures as in individualistic ones (Church & Katigbak, 2000).

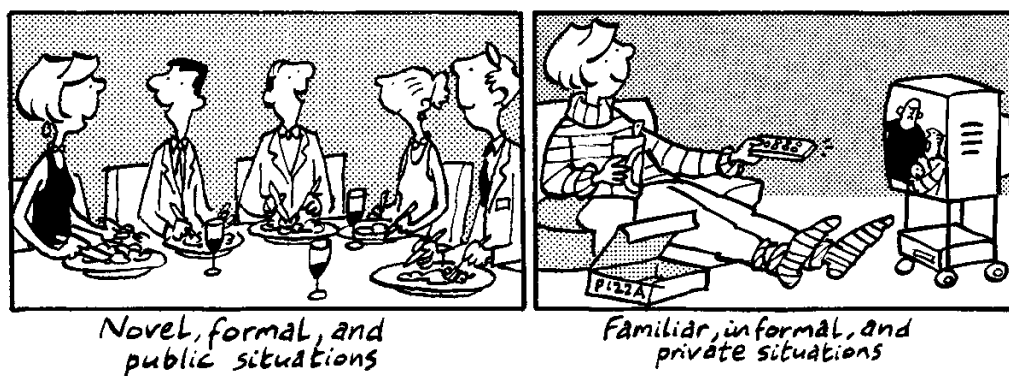
EVALUATION OF THE TRAIT APPROACH

Trait-based approaches to personality seem to assume that individuals will behave consistently across different situations. In other words, individuals' personality (which doesn't change from situation to situation) has a strong influence on their behavior in all situations. This is known as strong **cross-situational consistency**. According to Mischel (1968), there is actually much less cross-situational consistency than predicted by trait theories. His literature review indicated that personality measures rarely correlate more

Key Term

Cross-situational consistency:

the extent to which any given individual responds in the same way to different situations.



than about $+0.30$ with behavior in any given situation. This means that personality accounts for no more than 9% of individual differences in behavior. Mischel concluded that personality predicts behavior poorly, and that behavior is overwhelmingly determined by the situation in which the individual finds himself/herself.

There are two major problems with Mischel's analysis. First, he assumed that situational factors were much more important than personality in determining behavior. However, he didn't actually test to see whether this was true! Sarason, Smith, and Diener (1975) worked out the percentage of the variance (differences in behavior among individuals) accounted for by personality and by the situation across 138 experiments. On average, the situation accounted for 10.3% of the variance, whereas personality accounted for 8.7% of the variance. Thus, it is simply incorrect that behavior is determined substantially more by the situation than by personality.

Second, Mischel (1968) only considered the effects of personality on behavior in *specific* situations. Individual differences in personality might be much more important if we considered behavior over extended periods of time. Precisely this was done by Fleenor (2001). Participants carried personal data assistants with them for a few weeks. Several times a day, they indicated how much their behavior exhibited each of five major personality traits: extraversion; agreeableness; conscientiousness; emotional stability [opposite of neuroticism]; and intellect [resembles openness].

What did Fleenor (2001) find? He confirmed Mischel's findings indicating that any given individual's behavior fluctuates considerably from one situation or time to another. However, the findings were very different when he compared individuals' average levels of each personality trait in the first week against those in the second week. Here there was very high consistency with an average correlation of about $+0.9$. Fleenor (2004, p. 86) drew the following conclusions from these findings:

Everyone routinely acts in a wide range of ways on a given dimension of behavior, yet different people's ranges of behaviors are centered on different portions of the dimension, and each individual's center remains very stable across large periods of time.

SOCIAL COGNITIVE THEORY

Albert Bandura, an American psychologist born in 1925, has spent many years developing his social cognitive approach. Bandura (1999) argued that we need to consider personal factors, environmental factors, and the individual's own behavior to obtain a full picture of what is happening. All three factors are assumed to influence each other in complex ways. Thus, the environment influences our behavior, but our personality and behavior also help to determine our environment. For example, extraverts actively seek out social situations more often than introverts (Furnham, 1981). This approach is known as the **triadic reciprocal model** (see the figure on the following page).

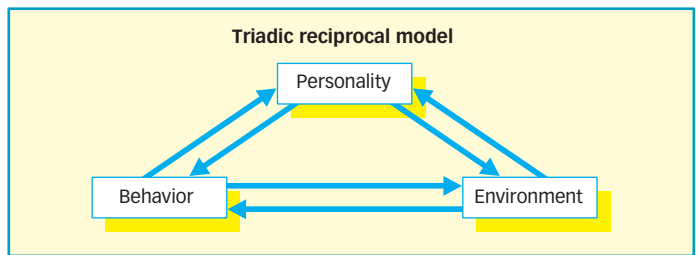
Bandura's theoretical approach is much more complex than that of traditional trait theories of personality. Trait theorists emphasize the notion that personality influences

Key Term

Triadic reciprocal model: Bandura's view that an individual's personality, the environment, and his/her own behavior all influence each other.

behavior, which corresponds to only one out of the six arrows in the triadic reciprocal model (the one running from personality to behavior). They also argue that the environment influences behavior (a second arrow in the triadic reciprocal model). However, few trait theorists focus on the other four arrows in that model. This is a real limitation. People's personalities influence the situations in which they find themselves as well as how they behave in those situations.

There are two other important differences between Bandura's social cognitive theory and trait theories. Bandura (e.g., 1999) argues strongly that we must take full account of the particular situation or context in which individuals find themselves if we are to predict their behavior. In contrast, most trait theorists claim that an individual's personality will influence their behavior in similar ways in most situations. Second, Bandura argued that individual differences in cognitive processes and strategies play an important role in determining differences in behavior. In contrast, trait theorists have generally had little or nothing to say about such processes.



SELF-EFFICACY

The notion of self-efficacy is of central importance within Bandura's social cognitive theory. **Self-efficacy** refers to the beliefs individuals have concerning their ability to cope with a particular task or situation and to achieve the desired outcome. In the words of Bandura (1977, p. 391), self-efficacy judgments are concerned, "not with the skills one has but with judgments of what one can do with the skills one possesses." It is assumed that high self-efficacy has beneficial effects on performance because it leads to increased task motivation. Self-efficacy is claimed to predict several aspects of behavior:

Given appropriate skills and adequate incentives . . . efficacy expectations are a major determinant of people's choice of activities, how much effort they will expend, and how long they will sustain effort in dealing with stressful situations (Bandura, 1977, p. 194).

Four factors determine an individual's sense of self-efficacy in any given situation:

1. *Previous experiences* of success and/or failure the individual has had in that situation.
2. *Relevant vicarious [second-hand] experiences*: For example, if you see someone else cope successfully with a situation, this may increase your self-efficacy beliefs.
3. *Verbal (or social) persuasion*: Your feelings of self-efficacy may increase if someone argues persuasively that you have the skills to succeed in that situation.
4. *Emotional arousal*: High levels of arousal are often associated with anxiety and failure, and can reduce feelings of self-efficacy.

Findings

The key assumption is that self-efficacy beliefs should predict behavior. There is considerable support for this prediction. For example, Dziewaltowski (1989) assessed the ability of various factors to predict the amount of exercise students would take over a 7-week period. Self-efficacy beliefs concerning their ability to take part in an exercise program when faced with competing demands emerged as the best single predictor. Convincing evidence that self-efficacy predicts work-related performance was reviewed by Stajkovic and Luthans (1998). The average correlation between self-efficacy and work performance across 114 studies was +.38. Thus, self-efficacy was a moderately strong



Previous success in a situation may make an individual more likely to believe they will succeed again, whereas previous failure may lead to reluctance to put much effort in or show much interest.

Key Term

Self-efficacy: an individual's beliefs concerning his or her ability to cope successfully with a particular task or situation.

predictor, producing a 28% increase in performance. Stajkovic and Luthans (p. 252) concluded as follows: “This 28% increase in performance due to self-efficacy represents a greater gain than, for example, those obtained in meta-analyses examining the effect on performance of goal setting (10.39% . . .), feedback interventions (13.6% . . .), or organizational behavior modification (17% . . .).

Stajkovic and Luthans (1998) discovered that self-efficacy was more strongly associated with high task performance on easy than on complex tasks. In addition, the strength of the association was consistently higher in laboratory settings than in more naturalistic field settings. Why did these differences occur? We would expect the relationship between self-efficacy and performance to be greater when participants possess detailed information about task demands, the best task strategy to adopt, and so on. Participants performing difficult tasks in field settings will often lack sufficient information to make accurate judgments of self-efficacy.

How does self-efficacy compare to standard personality factors in its ability to predict behavior? Caprara, Barbaranelli, and Pastorelli (1998) obtained measures of self-efficacy and of the Big Five factors in a study on academic achievement and peer preference. The Big Five factors nearly all failed to have any predictive power, except that openness predicted academic achievement. In contrast, self-efficacy predicted academic achievement and peer preference.

Evaluation

- + Strong associations have often been found between self-efficacy and performance (e.g., Stajkovic & Luthans, 1998). These associations are generally larger than those of other variables with performance.
- + The emphasis on self-efficacy shows the value of focusing on cognitive processes of relevance to motivation.
- Self-efficacy predicts performance less well when the task is difficult and/or is performed under naturalistic conditions.
- There are tricky issues about causality. The theoretical assumption is that self-efficacy influences performance. However, past performance influences self-efficacy judgments, so the association between self-efficacy and performance is difficult to interpret.

SELF-REGULATION

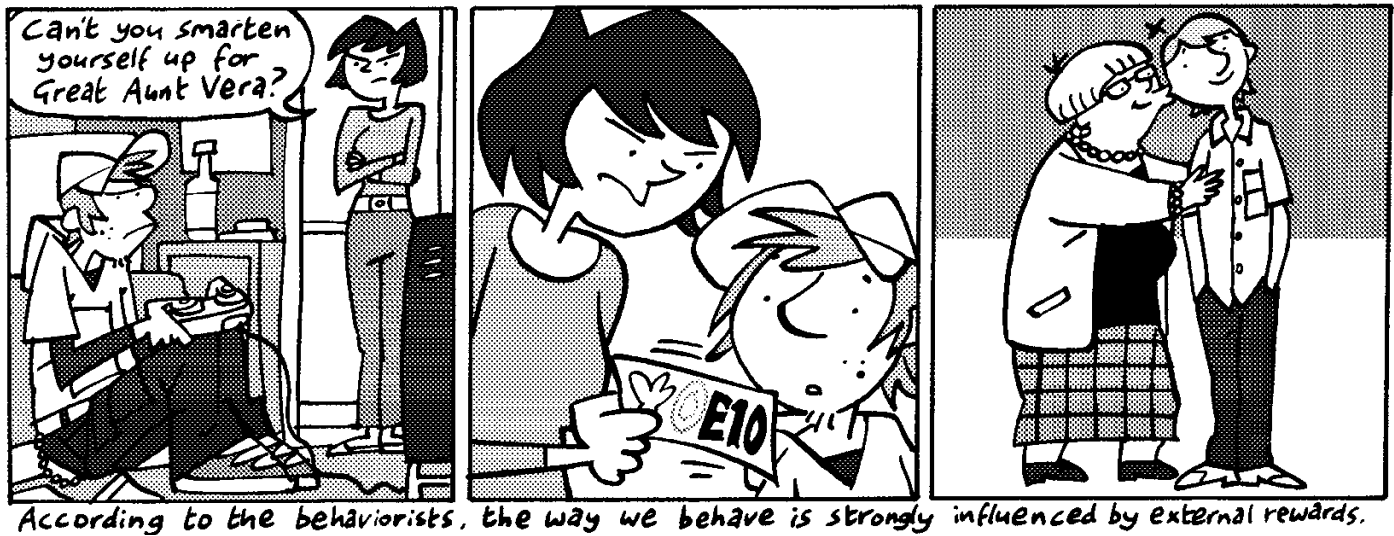
Bandura (1986) argued that our behavior is also influenced by self-regulation. **Self-regulation** involves using your cognitive processes to regulate and control your own behavior. For example, you may reward yourself if you achieve a given standard of performance. In broad terms, our behavior is often controlled by *internal* factors rather than the *external* ones (e.g., reward or reinforcement) emphasized by the behaviorists. Bandura (1977, pp. 128–129) provided vivid examples to support this position: “Anyone who attempted to change a pacifist into an aggressor or a devout religionist into an atheist would quickly come to appreciate the existence of personal sources of behavioral control.”

Bandura (1986) identified three processes of central importance to self-regulation:

1. *Self-observation*: Individuals observe their own behavior (e.g., the quality of their work; their productivity).
2. *Judgmental processes*: Individuals take account of their personal standards, of standard norms (i.e., other people’s performance), and of the role of personal and external factors in influencing their performance.
3. *Self-reaction*: Individuals experience positive self-reactions (e.g., pride; self-satisfaction) when their behavior reaches or exceeds their personal standards. They experience self-criticism or dissatisfaction when their behavior falls short of those standards.

Key Term

Self-regulation: using one’s own cognitive processes to control and regulate one’s own behavior and goals.

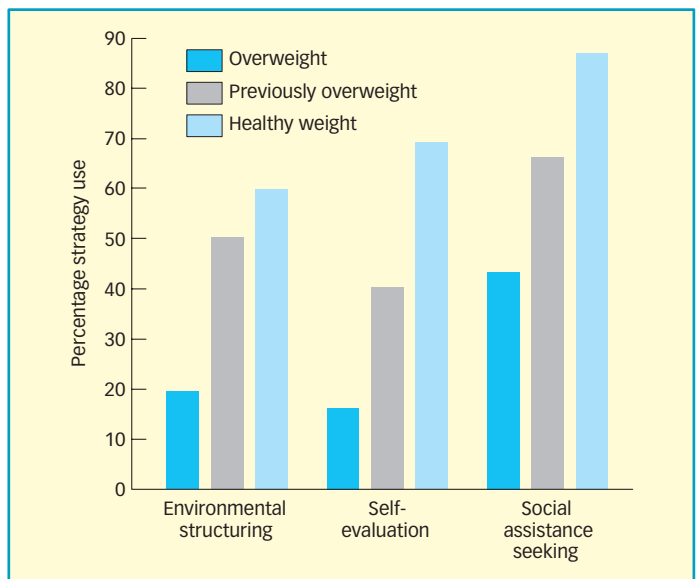


There are some links between self-regulation and self-efficacy. According to Bandura (1999, p. 176), “After people attain the standard they have been pursuing, those who have a strong sense of efficacy generally set a higher standard for themselves.”

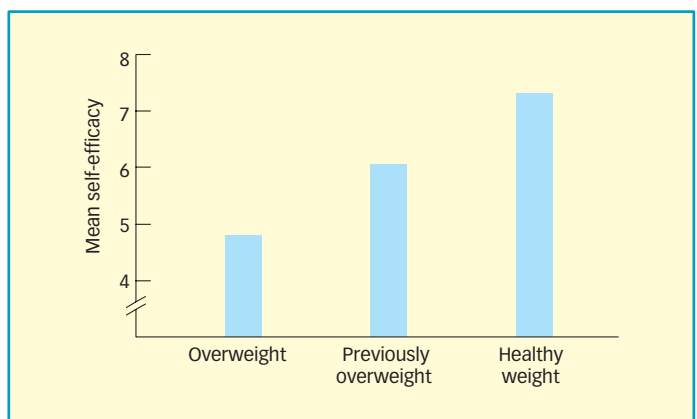
Findings

The prediction that people using self-regulation strategies should outperform those making little use of such strategies has been supported several times. Kitsantas (2000) reported positive evidence in a study on three groups of college students: overweight students who had failed to lose weight; previously overweight students who had successfully lost weight; and students with no weight problems. All participants completed a questionnaire indicating the self-regulation strategies they used, and their self-efficacy beliefs concerning their ability to use these strategies successfully. The self-regulation strategies considered included the following: goal setting and/or planning (e.g., desired weight); self-monitoring to keep track of progress in losing weight; self-evaluation of progress in weight control; and attempts to seek help in efforts to lose weight.

What did Kitsantas (2000) find? First, overweight students who didn't lose weight used fewer self-regulation strategies than did the other two groups (see the top figure on the right). This was especially the case for self-evaluation of progress, a strategy used far less by overweight students than by students in the other groups. Second, overweight students who didn't lose weight had lower levels of self-efficacy than students in the other groups (see the bottom figure on the right). Third, the use of several self-regulation strategies was only effective when combined with high self-efficacy. Thus, there is little point in using good strategies if you use them half-heartedly—you have to believe they are going to be effective.



Use of various self-regulation strategies by overweight, previously overweight, and healthy weight participants. Data from Kitsantas (2000).



Self-efficacy in overweight, previously overweight, and healthy weight groups. Data from Kitsantas (2000).

Evaluation

- + Much human behavior is motivated by self-reinforcement rather than directly by external rewards (e.g., money).
- + Most previous theories de-emphasized the role of internal factors (e.g., self-observation; self-reaction) in influencing motivation and behavior.
- Several factors influence self-regulation. As a result, an apparent failure to support the theory can be explained away with reference to factors not explicitly included in the study.
- “Sometimes a behavior is maintained with no obvious external reinforcer. In such a case, the [social cognitive] theorists assert that the behavior is being supported by self-reinforcement . . . if self-reinforcement accounts for behavior sometimes, why isn’t it adequate all the time? Why is external reinforcement ever necessary? How do you decide when it’s needed and when it isn’t?” (Carver & Scheier, 2000, p. 372).

Overall Evaluation

- + Self-efficacy and self-regulation both influence how any given individual will behave in a given situation.
- + Bandura’s approach has deservedly been very influential within health psychology. The extent to which people adopt healthy forms of behavior (e.g., giving up smoking; losing weight; taking exercise) depends to an important extent on self-efficacy and self-regulation.
- Social cognitive theory focuses on cognitive factors and de-emphasizes emotional ones. However, much human motivation and behavior is influenced by our emotions rather than by cool calculation.
- Bandura has paid little attention to the impact of genetic factors on individual differences in personality and behavior.
- Bandura has focused on predicting and understanding people’s behavior in *specific* situations. It is unclear whether his theory could account for individual differences in *broad* areas of life.

Chapter Summary

Theory of psychosexual development

- According to Freud, children pass through five stages of psychosexual development: oral; anal; phallic; latency; and genital. Various personality types are associated with problems at a given psychosexual stage.
- Excessive (or insufficient) gratification at any stage can lead to fixation. Adults experiencing severe problems often regress to the stage at which they fixated.
- Childhood experiences clearly influence the development of adult personality in some ways.
- There is evidence for at least some of the personality types identified by Freud.
- Freud’s general approach is supported by the evidence. However, his more specific assumptions (e.g., the Oedipus complex; regression to earlier stages of development) are mostly incorrect.

Personality assessment

- Personality can be assessed by self-report questionnaires, ratings, objective tests, and projective tests. Useful tests need to be standardized with good reliability and validity.

- Questionnaires are affected by social desirability bias, which can be detected by the use of lie scales. In spite of such bias, most well-known personality questionnaires have high reliability and low to moderate validity.
- Ratings suffer from the limitation that most raters have only partial information about the people they are rating. However, the reliability and validity of ratings (especially reliability) are generally reasonably good.
- Most objective and projective tests possess low reliability and validity and so are of rather limited usefulness.

Trait approach to personality

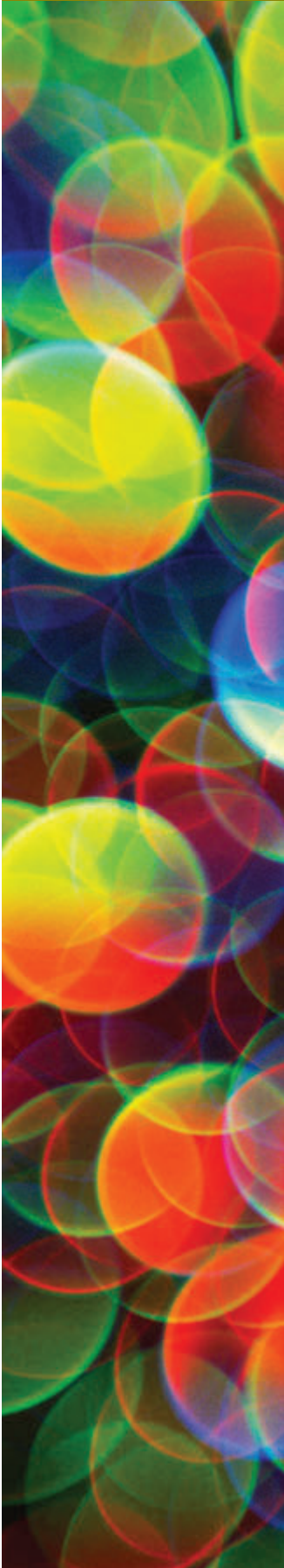
- Cattell made use of the fundamental lexical hypothesis, according to which each language contains words describing all of the main personality traits.
- Cattell assessed personality by using questionnaires, life or rating data, and objective tests. The personality traits identified by questionnaire and rating data were similar to each other.
- Cattell's 16PF questionnaire was designed to assess 16 personality factors, but actually measures only about half that number.
- H.J. Eysenck identified three superfactors: extraversion; neuroticism; and psychoticism. The evidence confirms the importance of the first two, but psychoticism is less important.
- Individual differences in all three superfactors depend in part on genetic factors.
- There is general agreement that there are five main personality traits (the Big Five), often identified as openness, conscientiousness, extraversion, agreeableness, and neuroticism. These factors are claimed to be independent of each other, but are not actually so.
- Individual differences in each of the Big Five factors depend in part on genetic factors.
- Most trait approaches have ignored the existence of culture-specific traits. In addition, the notion of stable and relatively unchanging personality traits is more applicable in individualistic than in collectivistic cultures.

Social cognitive theory

- According to Bandura's triadic reciprocal model, personal factors, environmental factors, and the individual's own behavior all influence each other in complex ways.
- Self-efficacy is an important determinant of behavior. The level of self-efficacy in any given situation depends on previous experiences of success in that situation, relevant vicarious experiences, verbal or social persuasion, and emotional arousal.
- Behavior is also influenced by self-regulation. Three processes of central importance to self-regulation are self-observation, judgmental processes, and self-reaction.
- There is empirical support for the predicted positive effects of self-efficacy and self-regulation on performance.
- Bandura's approach focuses on cognitive factors and behavior in fairly specific situations. It de-emphasizes emotional factors and general patterns of behavior.

Further Reading

- Carver, C.S., & Scheier, M.F. (2004). *Perspectives on personality: International edition* (5th ed.). New York: Allyn & Bacon. This textbook provides good basic coverage of approaches to personality.
- Mathews, G., Deary, I.K., & Whiteman, M.C. (2003). *Personality traits* (2nd ed.). Cambridge, UK: Cambridge University Press. Major trait approaches to personality are discussed in an authoritative way by the authors of this book.
- Ozer, D.J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, 57, 401–421. The authors show how each of the Big Five personality factors helps us to predict individuals' physical and psychological well-being, quality of relationships, and so on.
- Pervin, L.A., Cervone, D., & John, O.P. (2005). *Personality theory and research* (9th ed.). Chichester, UK: Wiley. Provides a thorough introduction to the field of personality.
- Triandis, H.C., & Suh, E.M. (2002). Cultural influences on personality. *Annual Review of Psychology*, 53, 133–160. The authors of this chapter provide a good overview of theory and research on cross-cultural research in personality.



- **Cognitive development: Specific abilities** *pp.313–335*
 - Perceptual development*
 - Memory development*
 - Language development*
 - Theory of mind*
- **Theories of cognitive development** *pp.337–353*
 - Piaget's theory*
 - Vygotsky's theory*
 - Contemporary approaches*
 - Educational implications*
- **Social development in everyday life** *pp.355–375*
 - Gender development*
 - Prosocial behavior*
 - Antisocial behavior*
 - Moral development*
- **Attachments and friendships** *pp.377–399*
 - Attachment*
 - Deprivation effects*
 - Divorce and day care*
 - Peer relationships*
 - Parents, peers, or genes?*

INTRODUCTION TO Developmental Psychology

The four chapters in this section of the book (Chapters 13–16) deal with developmental psychology. Developmental psychology is mainly concerned with the psychological changes occurring during the time between birth and adulthood. However, some developmental psychologists are interested in changes throughout the lifespan. Our primary focus is on infancy and childhood, because that is the period of life in which the most dramatic changes in development occur.

Developmental psychology (as Sigmund Freud was one of the first psychologists to acknowledge) is of crucial importance to an understanding of adult behavior. What we are now as adults depends to a large extent on the experiences we had during the years of childhood. In the words of the poet William Wordsworth in his poem, *The Rainbow*, “The Child is father of the Man.” The fact that the childhood years are vitally important means that society has a responsibility to ensure that all children are provided with the opportunities and support they need to develop into well-adjusted and successful adults.

It is also worth pointing out that the study of children is intrinsically fascinating, especially to parents. As a parent myself, I still remember very clearly being put in my place by my daughter Fleur, who was $2\frac{1}{4}$ at the time. We were on a cross-channel ferry, and I said to her, “Look, Fleur, there’s a boat.” I instantly felt deflated when she replied, “It’s not a boat, Daddy, it’s a yacht!” At the age of 4, Juliet (my other daughter) also managed to deflate me when she said earnestly, “A professor should know at least everything.”

Much research in developmental psychology involves laboratory studies. The potential danger with laboratory studies is that the findings obtained may be of limited relevance to everyday life. However, researchers are well aware of the dangers, and increasing numbers of studies are being carried out in naturalistic surroundings (e.g., school playgrounds). The main advantage of naturalistic studies is that they are more likely to provide findings of relevance to children’s everyday lives and behavior. However, such studies tend to suffer from the disadvantage that they are less well controlled than laboratory studies. The best way of dealing with these issues is to carry out both laboratory and naturalistic studies. If broadly similar findings are obtained from both kinds of studies, we can have some confidence that the findings are genuine and applicable to everyday life.

Most research on children used to involve obtaining relatively simple response measures (e.g., reaction times). However, there has been a substantial increase in studies in which much more complex types of behavior have been assessed. There are two key reasons for this change. First, the ready availability of video-recording equipment means that children’s behavior can be replayed over and over again to extract its full richness. Second, computer-based software packages for analyzing complex sets of data are now in wide use, and greatly simplify data analyses.



What we become as adults is largely dependent on the experiences we had in childhood.

ECOLOGY OF DEVELOPMENT

Bronfenbrenner (1979) proposed an approach to developmental psychology that has proved influential. According to his ecological model, children should be studied in terms of the ecology or the social and cultural environment in which they grow up. He argued that development occurs within various environmental structures arranged “like a set of Russian dolls” (p. 3). There are four basic structures or systems:

- *Microsystem*. A microsystem consists of a child’s direct experiences in a given setting. Children typically encounter various microsystems in their everyday lives (e.g., school microsystem; home microsystem). Most developmental research focuses on children’s behavior within a single microsystem.
- *Mesosystem*. This consists of the inter-relationships among the child’s various microsystems. For example, the child’s ability to form friendships at school may depend in part on how securely attached he/she is to his/her parents at home.
- *Exosystem*. This consists of factors (e.g., parents’ workplaces; mass media) that are not experienced directly by the child, but which nevertheless have an indirect impact on him/her.
- *Macrosystem*. This consists of the general beliefs and ideology of the culture, which can have various indirect influences on children. For example, if a child’s parents have occupations highly valued by society, this may influence his/her behavior at home.

Bronfenbrenner’s (1979) ecological approach is attractive for two reasons. First, it helps to integrate the closely related areas of developmental and social psychology. Second, it is much broader in scope than most developmental theories, in which the emphasis is often mainly at the level of specific microsystems. However, a limitation of Bronfenbrenner’s approach is that it does not generate many precise and testable predictions.

It follows from Bronfenbrenner’s approach that developmental psychologists should carry out cross-cultural studies. Such studies would help to clarify the role played by the macrosystem in influencing children’s development. There is an increasing amount of cross-cultural developmental research, some of which is discussed in the following chapters. For example, evidence that children in collectivistic cultures exhibit more altruistic behavior than those in individualistic cultures (e.g., Whiting & Whiting, 1975) is discussed in Chapter 15, and cross-cultural differences in attachment behavior (e.g., Sagi, van IJzendoorn, & Koren-Karie, 1991) are dealt with in Chapter 16. Evidence of cross-cultural differences even in preschool was reported by Tobin, Wu, and Davidson (1989). American children in preschool differed considerably from each other in their behavior, and sometimes fought each other for attractive toys. In contrast, Chinese children behaved in a highly regulated way—they all went to the toilet at the same time, and they were expected to play in a cooperative and helpful way with each other.

There are major cross-cultural differences in the duration of childhood and adolescence, both of which last for many more years in affluent cultures where life expectancy is high. Indeed, it is sometimes argued that adolescence only exists in Western cultures! Consider the !Kung San people living in the Kalahari desert. When children in this culture reach puberty, they have already acquired good hunting and gathering skills that allow them to be self-sufficient and economically independent (Cole & Cole, 1993).

ORGANIZATION OF CHAPTERS 13–16

Developmental psychology deals largely with cognitive and social development, and the four chapters in this section reflect that distinction. Chapters 13 and 14 are concerned with cognitive development, and Chapters 15 and 16 with social development.

Chapter 13 focuses on the development of various important abilities including perception, memory, language, and theory of mind. In Chapter 14, the emphasis shifts to general theories of cognitive development including the very influential approaches proposed by Jean Piaget and Lev Vygotsky. Chapter 15 deals with the development of several aspects of social development (e.g., moral development; development of prosocial and antisocial behavior; gender development). Finally, in Chapter 16, we consider the most general and important aspects of social development. These include the child's attachment to its parents, the consequences of parental deprivation and divorce, friendships, and relationships with children of the same age.

It is important not to infer from the organization of these chapters that cognitive and social development are completely separate. In fact, cognitive development is influenced by social development, and social development depends in part on cognitive development. The links between these two major forms of development are emphasized at several points in the four chapters.

chapter 13

Contents

Perceptual development	313
Memory development	317
Language development	322
Theory of mind	329

Cognitive development: Specific abilities

13

This chapter and the next are concerned with children's cognitive development. You only have to compare infants a few months old with children aged 12 or 13 to see that dramatic improvements in cognitive skills and abilities occur over a period of a few years. The task of developmental psychologists is twofold. First, they need to describe *what* changes during the course of cognitive development. That is the easy part. Second, and much more challenging, they have to work out *how* and *why* these changes occur.

In this chapter, we consider the development of major cognitive abilities. We start with perceptual development, which is of fundamental importance to the young child as he/she comes to grips (literally) with the environment. After that, we consider memory development, finding that children's memories improve considerably during childhood. We then turn to the crucial issues of language development. The central puzzle here is that language acquisition is a very complex accomplishment, but the overwhelming majority of young children seem to learn their native language with amazing speed. Psychologists are still struggling to understand precisely how this is done. Finally, we focus on another key aspect of cognitive development. Young children at about the age of 4 start to develop an understanding that other people often have different beliefs and intentions to their own. This is known as developing a theory of mind, and it has substantial effects on children's interactions with others.

PERCEPTUAL DEVELOPMENT

How much can the newborn baby (or neonate) see and hear? It used to be assumed that the answer was "very little." Towards the end of the nineteenth century, William James described the world of the newborn baby as a "buzzing, blooming confusion, where the infant is seized by eyes, ears, nose and entrails all at once." This suggests that the infant is bombarded by information in every sensory modality, and can't attach meaning to this information. In fact, that view underestimates the capabilities of infants. Many basic perceptual mechanisms are working at a very early age, and infants are *not* merely helpless observers of their world.

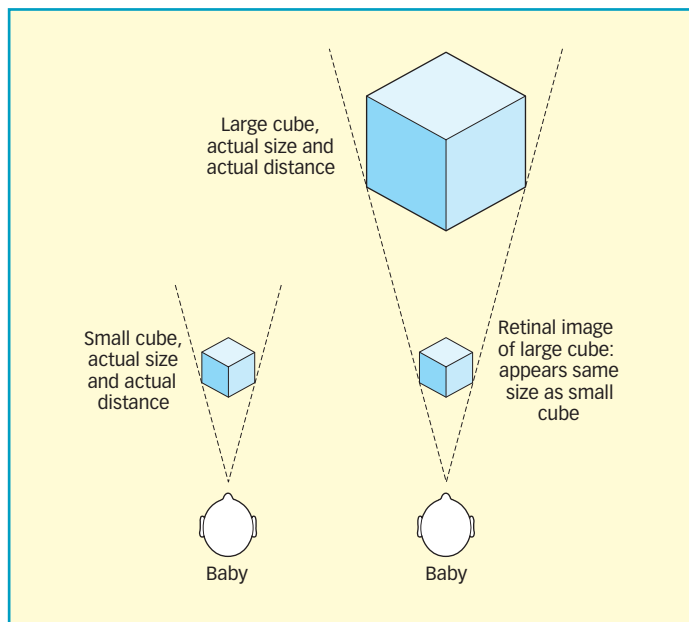
There is increasing evidence that even fetuses are capable of perceptual learning. For example, Kisilevsky et al. (2003) exposed 38-week-old fetuses to a tape recording of their mother or a female stranger reading a poem. Fetal heart rate increased when the poem was read by the mother but decreased when read by the stranger. Thus, fetuses can discriminate between different voices prior to birth.

BASIC ASPECTS OF VISION

Newborns are at a great disadvantage to adults with respect to several basic aspects of vision. First, “The visual field of the newborn human infant . . . consists of a narrow tunnel around the line of vision, and the ability to resolve visual details is roughly 40 times poorer than that of adult humans” (Sireteanu, 1999, p. 59). Second, color vision is either nonexistent or nearly so during the first weeks of life. According to Teller (1997, p. 2197), “By two months, rudimentary color vision has arrived. Most infants can probably discriminate red, blue, and green from each other, but not yet yellows and yellow-greens.” Third, the eyes of newborns have a fixed focal length for the first 3 months of life. In contrast, older children and adults show accommodation, in which the curvature of the eye’s lens alters to bring objects at different distances into focus. Newborns don’t show accommodation, and so only objects 8 inches in front of them can be seen clearly. Fourth, binocular disparity (in which there is a difference or disparity between the images projected on to the retinas of the two eyes) assists depth perception. Infants below the age of 3 months lack binocular disparity (Teller, 1997).



Newborns are able to relate visual and auditory information, for example, the sight and sound of a musical mobile.



Two cubes of different sizes may project retinal images of the same size, depending on their distance from the viewer.

Key Term

Size constancy: objects are perceived to have a given size regardless of the size of the retinal image.

On the positive side, newborns’ eye movements are fairly systematic rather than random. Of particular importance, newborns generally seem to search for the edges of objects. When an edge is detected, a newborn’s next few eye movements tend to be small so as to keep the visual focus close to it (Haith, 1980). In addition, newborns show some evidence of **size constancy**, which involves a given object being perceived as having the same size regardless of its distance from us. Slater, Mattock, and Brown (1990) familiarized newborns with either a small or a large cube over several trials. After that, the two cubes were presented successively. The larger cube was presented at a greater distance from the newborns than the smaller cube, with the size of the retinal image being the same in both cases (see the figure below). All the newborns could distinguish between the cubes, because they spent longer looking at the new cube than the old one. The fact the newborns could tell the difference between two cubes having the same-sized retinal image indicates they possessed at least some elements of size constancy. Sai (2005) found that newborns can combine visual and auditory information. Newborns only showed a preference for the mother’s face when they had been exposed previously to her face and voice at the same time.

The visual abilities that are deficient in newborns and infants develop rapidly during the early months of life. Improvements in most aspects of visual perception depend on naturally occurring developmental changes. For example, Maurer, Lewis, Brent, and Levin (1999) studied human infants who had been treated for cataracts in one or both eyes. Initially, their visual acuity was at the level of newborn infants. Strikingly, however, they showed substantial improvement in visual acuity after only a 1-hour exposure to the visual world. This shows the great plasticity of the visual system. It also shows that prolonged visual experience is not necessary for the development of visual acuity.

The position with respect to binocular disparity is more complicated. Banks, Aslin, and Letson (1975) studied adults who had had a problem with binocular vision because of having a squint in childhood that was subsequently corrected. Adults who had a squint

at birth and received surgery by 30 months of age had good binocular disparity. Those who had a squint diagnosed between 2 and 7 years of age and surgery within 3 years of diagnosis had reasonably good binocular disparity. However, adults who had a squint at birth and surgery between the ages of 4 and 20 had little or no binocular disparity.

The above complex findings suggest that there is a critical or sensitive period for the development of binocularity during the early years of life. If children don't develop binocularity during the early years of life because of an uncorrected squint, then it is very difficult to develop it later. It has also proved difficult to treat **amblyopia** (impaired vision resulting from disuse of an eye) in older children and adults, thus providing more evidence of a critical period. However, visual acuity in adults with amblyopia can be improved if they receive prolonged practice on a difficult visual task (Levi, 2005).

FACE PERCEPTION

Faces (especially its mother's face) form a very important part of the newborn infant's visual environment. This has led some theorists to argue that natural selection has equipped humans with an innate bias favoring facial stimuli. For example, Morton and Johnson (1991) argued that human infants are born with a mechanism containing information about the structure of human faces. This mechanism is known as **CONSPEC**, because the information about faces it contains relates to conspecifics (members of the same species).

The key prediction from this theoretical approach is that newborn infants should exhibit a clear preference for facial stimuli over other stimuli. There is much evidence supporting this prediction. For example, Johnson, Dziurawiec, Ellis, and Morton (1991) found that newborns in the first hour after birth showed more visual tracking of realistic faces than of scrambled but symmetrical faces. This suggests that some aspects of face perception don't depend on learning.

In spite of the plausibility of the notion that there is an innate bias for faces, there is increasing evidence that it is probably wrong. An alternative view has been developed by Turati and colleagues (e.g., Turati, 2004). According to this view, there is nothing special about faces. Instead, newborns simply have a preference for stimuli having more patterning in their upper than in their lower part. This is found with faces, but is also found with numerous other stimuli.

Support for the above view was reported by Simion et al. (2002) and by Turati, Simion, Milani, and Umiltà (2002). Simion et al. found that newborns preferred visual stimuli having more elements in the top half than in the bottom half. Importantly, this was the case even when the stimuli didn't look anything like faces. Turati et al. (2002) found (as predicted) that newborns aged between 1 and 3 days did *not* prefer face-like stimuli to nonface-like stimuli when the number of elements in the upper part was held constant. The key finding was that newborns preferred a stimulus of only a nonface-like arrangement in the upper part to a stimulus consisting only of a face-like arrangement in the lower part (see the figure below). These findings suggest that, "Newborns direct their gaze toward faces because they belong to a broader stimulus category that is characterized by a greater number of high-contrast areas in the upper portion of the pattern" (Turati et al., 2002).

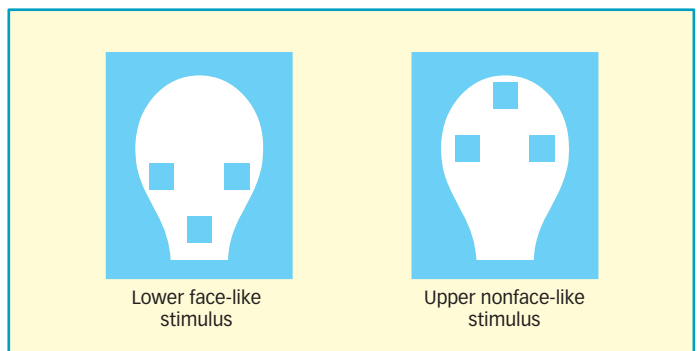
Simion, Turati, Valenza, and Leo (2006) replicated the finding that newborns prefer scrambled faces with more elements in the top area to natural faces. However, 3-month-olds preferred facial stimuli to scrambled faces. This suggests that cognitive mechanisms specialized for faces develop over the first few weeks or months of life.

DEPTH PERCEPTION

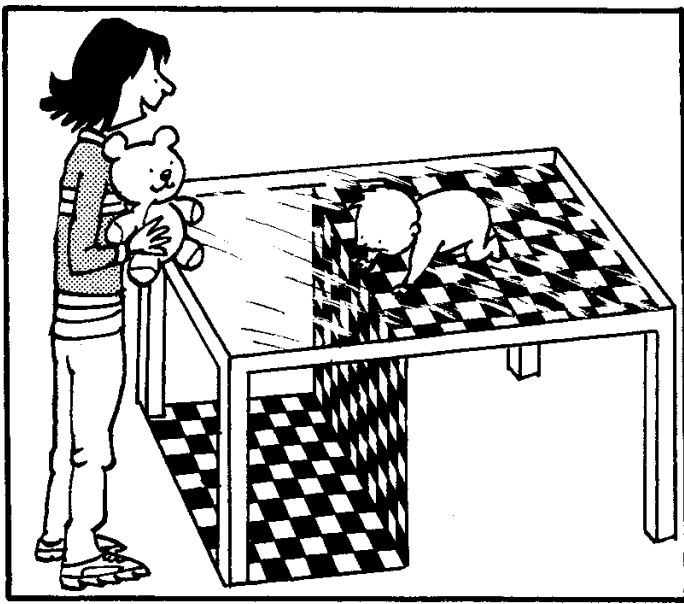
Gibson and Walk (1960) studied depth perception by designing a "visual cliff" involving a glass-top table

Key Term

Amblyopia: a condition involving impaired vision as a result of disuse of one eye in the absence of any obvious damage to it.



Newborns preferred an upper nonface-like stimulus (shown on the right) to the lower face-like stimulus (shown on the left). From Turati et al. (2002). Copyright © American Psychological Association. Reprinted with permission.



A drawing of Gibson and Walk's "visual cliff." Babies between 6.5 and 12 months of age were reluctant to crawl over the "cliff" edge, even when called by their mothers, suggesting that they perceived the drop created by the check pattern.

(see figure on the left). A check pattern was positioned close to the glass under one half of the table (the "shallow" side) and far below the glass under the other half (the "deep" side). Infants between the ages of 6.5 and 12 months were placed on the shallow side of the table, and encouraged to crawl over the edge of the visual cliff on to the deep side. Most failed to do so. This finding suggests they possessed some elements of depth perception. It may be relevant that binocular vision (which is helpful in depth perception) has typically developed by the age of 6 months (Teller, 1997).

Adolph (2000) argued that the development of depth perception is more complex than assumed by Gibson and Walk (1960). According to Gibson and Walk, infants acquire *general* knowledge (e.g., an association between depth information and falling) that stops them from crossing the visual cliff. In contrast, Adolph's sway model assumes that infants' knowledge is highly *specific*. According to this model, infants learn how to avoid risky gaps when sitting, but subsequently have to learn how to avoid such gaps when crawling.

Adolph (2000) obtained support for her sway model by studying 9-month-old infants more familiar with sitting than with crawling. The key findings were that, "The babies avoided reaching over risky gaps in the sitting posture but

fell into risky gaps while attempting to reach in the crawling posture" (Adolph, 2000, p. 290). The implication is that learning with the visual cliff is specific to a given posture (e.g., sitting) and new learning is needed when infants become more mobile. However, slightly older infants show general rather than specific learning. Witherington et al. (2005) found that newly walking infants consistently avoided the deep side of the visual cliff. These infants had learned to avoid risky gaps while crawling, and this learning simply transferred over to walking. The take-home message is that learning to avoid gaps is specific to a given posture in younger infants but becomes more general in older ones.

THEORETICAL APPROACHES

Perceptual development is complex, with different perceptual abilities developing at different ages and in various ways. Some aspects of visual perception (e.g., relating visual and auditory information; aspects of face perception; aspects of size constancy) are present at birth or very shortly thereafter. These aspects reflect either innate visual capacities or very rapid learning. Other aspects of visual perception (e.g., visual acuity; color vision) develop several weeks after birth, and probably depend on maturational factors. There are still other aspects of perception (e.g., binocular disparity) for which there may be a critical or sensitive period of their development. Finally, some aspects (e.g., those relating to depth perception) develop only after several months of life and probably require certain kinds of learning.

Theorists differ in terms of how they interpret all this evidence. Some (e.g., Teller, 1997) emphasize the limitations of infants' visual perception whereas others (e.g., Slater, 1990, 1998) focus on the strengths of their perception. Here are Teller's (1997, p. 2196) views on the situation in which newborn infants find themselves:

Their acuity and contrast sensitivity are very poor but are measurable. Their . . . eye movements reveal the capacity to analyze the direction of motion of large, high-contrast objects . . . However, they should reveal no appreciation of stereo depth, no capacity to respond to low contrasts or to fine spatial details, and probably no color vision. Their visual worlds are probably marked less by blooming and buzzing than by the haziness of low-contrast-sensitivity, the blurriness of spatial filtering, and the blandness of monochrome [black-and-white].

In contrast, Slater (1990, p. 262) is more upbeat in his assessment of infants' perceptual skills:

No modality [none of the senses] operates at adult-like levels at birth, but such levels are achieved surprisingly early in infancy, leading to recent conceptualizations of the “competent infant” . . . early perceptual competence is matched by cognitive incompetence, and much of the re-organization of perceptual representation is dependent upon the development and construction of cognitive structures that give access to a world of objects, people, language, and events.

Evidence that visual perception in newborns is influenced by cognitive incompetence was provided by Slater et al. (1990). They presented newborns with a visual display in which a rod was seen moving from side to side behind a box. The newborns never saw the entire rod. Instead, they could see the top and bottom of the rod above and below the box. Adults and infants aged 3 or 4 months assume there is a single rod of which they can see only the ends (Kellman & Spelke, 1983). However, the newborn infants seemed to perceive the rod behind the box as a broken one.

In sum, newborns immediately possess some of the main aspects of visual perception. After that, there are rapid and substantial improvements in visual perceptual abilities during the early months of life. Teller (1997) emphasizes the role of maturational changes in producing these improvements. In contrast, Slater (1990, 1998) focuses on the development of the cognitive system and the infant's growing store of knowledge as the engines of improvements in visual perception.

MEMORY DEVELOPMENT

Children's ability to learn and remember information becomes better and better during development. Why does this happen? According to Siegler (1998), there are four possible explanations:

1. *Basic processes and capacity*: For example, the capacity of short-term or working memory may increase.
2. *Strategies*: Children possess more memory strategies as they develop, and may use these strategies more efficiently.
3. *Metamemory*: **Metamemory** is “knowledge about memory. The development of metamemory is the development of the ability to monitor and regulate one's own memory behavior” (Goswami, 1998, p. 206).
4. *Content knowledge*: Older children possess more knowledge than younger ones, and this may make it easier to learn and remember new information.

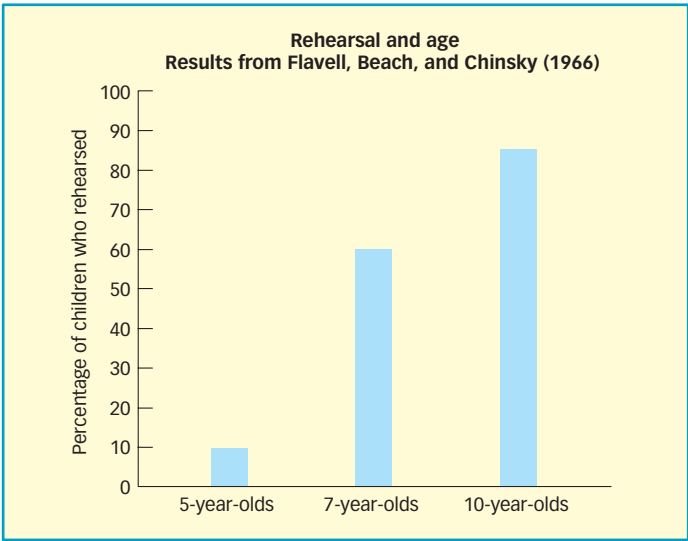
All four explanations are discussed below. As we will see, the evidence suggests that *all* of them possess some validity, and jointly they account for the dramatic improvement shown by children in the course of development. However, most research on memory development suffers from important limitations. We know older children possess memory skills and strategies not possessed by younger ones. However, we lack a clear sense of the processes involved in the development of these skills and strategies.

BASIC PROCESSES AND CAPACITY

One of the most important parts of the human memory system is working memory (see Chapter 8). In essence, working memory allows us to process and store information at the same time. An example of a task requiring the use of working memory is as follows (Swanson, 1999). The participants are given a sentence such as, “Now suppose somebody wanted to have you take them to the supermarket at 8651 Elm Street.” The task was to recall the name of the street and then to recall the number. Swanson found with children aged from 6 upwards that working memory improved continuously throughout childhood. Swanson also found that age differences in working memory predicted reasonably well children's performance in reading and arithmetic. This suggests that working memory capacity is an important factor in cognitive development.

Key Term

Metamemory: knowledge about one's memory and about how it works.



Baddeley (1986) argued that working memory consists of a central executive (resembling an attentional system), a phonological loop (used for verbal rehearsal), and a visuospatial sketchpad that stores visual and spatial information (see Chapter 8). Gathercole et al. (2004) found that the capacity of all three components of working memory increased progressively from the age of 4 through to early adolescence.

MEMORY STRATEGIES

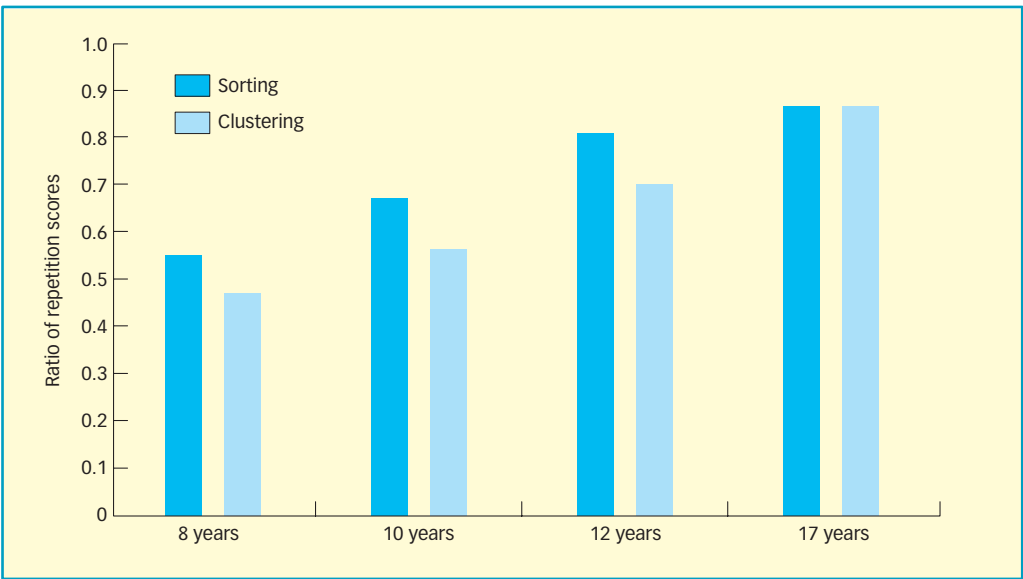
Adults use several memory strategies (e.g., verbal rehearsal; mnemonics) to assist them in learning and remembering. As we would expect, there are developmental changes in the use of strategies. For example, Flavell, Beach, and Chinsky (1966) found on a picture-learning task that far more 10-year-olds than 7- or 5-year-olds rehearsed the learning material (see figure on the left).

Consider a task such as remembering a list of four words belonging to each of six categories, with the words presented in a random order. Adults would typically use an organizational strategy in which they rehearsed (and recalled) the words category by category. This organizational strategy leads to improved recall of the list compared to strategies in which the list isn’t organized category by category.

What do children do when learning similar lists? Schneider, Knopf, and Stefanek (2002) presented children of various ages with pictures belonging to familiar categories (e.g., animals; food). The children sorted the pictures and then they recalled them in any order (free recall). Of central interest was the extent to which the children organized the stimuli into categories at learning (sorting task) and at recall. As expected, free recall increased steadily between the ages of 8 and 17. This increase was accompanied by progressive increases in use of an organizational strategy at learning and recall (see the figure below).

Schlagmüller and Schneider (2002) studied the rate of acquisition of an organizational strategy during sorting in children aged between 8 and 12. The findings were very clear-cut: “The transition from non-strategic to strategic occurred rapidly, indicating that children ‘jumped’ from random behavior to nearly perfect sorting scores” (Schlagmüller & Schneider, 2002, p. 313).

Sorting by categories at learning and clustering (recall by categories) as a function of age. From Schneider et al. (2002). Copyright © American Psychological Association. Reprinted with permission.



METAMEMORY

As children grow older, they show increased metamemory or knowledge about their own memory and how it works (see Schneider, 1999, for a review). For example, Yussen and Levy (1975) found that preschoolers' memory span was five items less than they had predicted, whereas 9-year-olds only overestimated their memory span by one item.

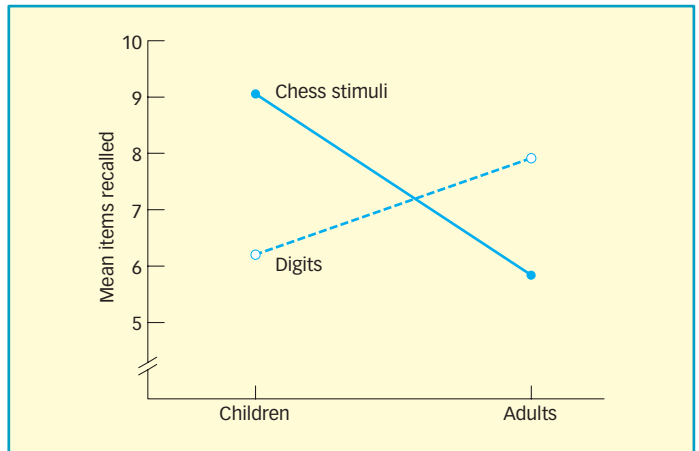
Does metamemory knowledge generally predict memory performance? Schneider and Pressley (1989) found in a meta-analysis of 60 studies that the correlation between metamemory and memory performance was $+0.41$, indicating a moderate relationship. Why isn't the relationship stronger? Children may not be motivated to use effective memory strategies they possess, they may feel a good memory strategy isn't needed when a word list is short, and so on.

CONTENT KNOWLEDGE

If the amount of relevant knowledge is of major significance in determining memory performance, then a well-informed child might remember some things better than an ill-informed adult. This prediction was tested by Chi (1978), who studied digit recall and reproduction of chess positions in 10-year-olds skilled at chess and adults knowing little about chess. The adults performed better than the children on digit recall. However, the children's recall of chess positions was more than 50% better than that of the adults (see figure on the right).

Schneider, Gruber, Gold, and Opwis (1993) compared children and adults having similar chess expertise. Both groups remembered chess positions comparably well and much better than nonexpert children and adults. Thus, memory for chess positions depends largely on expertise and hardly at all on age.

Content knowledge also influences memory in other ways. For example, there is evidence that the memory performance of even very young children is influenced by scripts, which are knowledge structures indicating the typical sequence for common events. Bauer and Thal (1990) studied script knowledge in children aged 21 or 22 months. The children were presented with a sequence of actions (e.g., giving teddy a bath involved putting the bear in the tub, washing it with the sponge, and then drying it with the towel). In one condition, sequences of actions were presented in the wrong or an unnatural order. The children tended to recall the sequences in the correct order even when they had been presented in the wrong order. Presumably this occurred because the children possessed (and made use of) the relevant scripts. Bauer, Hertsgard, Dropik, and Daly (1998) found that even 16-month-olds could recall events presented in a logical sequence (e.g., starting the engine before driving a car) after a 2-week retention interval. This presumably occurred because they could create organized representations of such events.



Immediate recall of chess positions and digits in children (mean age = 10 years 6 months) with expert knowledge of chess and in adults with limited knowledge of chess. Adapted from Chi (1978).

IMPLICIT MEMORY

Virtually all the research discussed so far has involved explicit memory (based on conscious recollection of information). This can be contrasted with implicit memory, a form of memory *not* dependent on conscious recollection. There are important differences between explicit and implicit memory (see Chapter 8). As we have seen, there are generally large improvements in explicit memory performance during development. This is in marked contrast to the findings from studies of implicit

Earliest memories

Lamont (2001) has demonstrated musical memories from before birth in research that is part of the UK's BBC *Child of Our Time* project. She compared babies from 11 families where in the last 3 months of pregnancy daily half-hour sessions of particular music had been played with babies from another 11 families who had had no such experience. The research is ongoing, but results indicate strongly that the experimental babies do have a musical long-term memory.

memory. According to Murphy, McKone, and Slee (2003, p. 125), “The usual finding of implicit memory studies to date is of no age-related increase . . . Of 18 published studies, 15 have concluded that implicit memory is stable across a wide age range (3 years to adulthood).”

A typical study of implicit memory was carried out by Russo, Nichelli, Gibertoni, and Cornia (1995). Children identified degraded pictures of objects, some of which had been seen before in intact form. Implicit memory was measured by the extent to which performance was better with objects previously seen. Very similar levels of implicit memory were observed in 4-year-olds, 6-year-olds, and young adults. In contrast, explicit memory for the pictures was highest in the adult group and lowest in the 4-year-olds.

Key Study

Murphy et al. (2003): Age and implicit memory

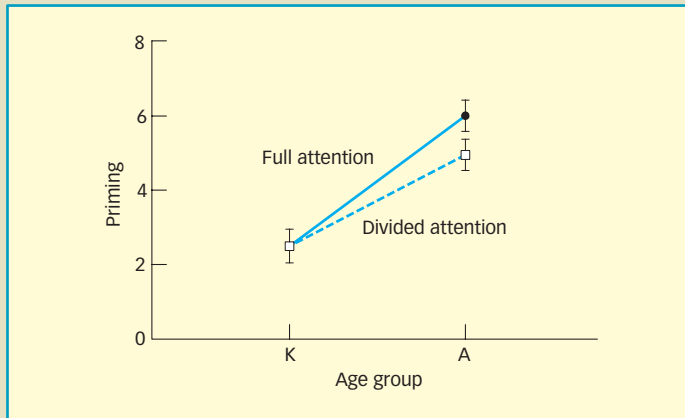
As we have seen, there is much evidence suggesting that age doesn't affect implicit memory even though it has large effects on explicit memory. Why are there such small (or nonexistent) effects of age on implicit memory? One possibility is that implicit memory doesn't involve metamemory or strategic processes, because participants don't know their memory is being tested. This is probably part of the answer. However, Murphy et al. explored another possibility. They pointed out that explicit memory is strongly influenced by the amount of relevant knowledge that a child possesses (e.g., Chi, 1978), and wondered whether the same was true of implicit memory. Previous research didn't provide an answer because most of the implicit memory tests used involved identification of common objects presented in degraded form. As a result, even young children possessed good knowledge about the stimuli used.

Accordingly, Murphy et al. used an implicit memory condition in which older children would possess much more knowledge than younger ones. They presented a word list with one word belonging to each of 20 categories. The category members were selected to be atypical but known to even the younger children (e.g., kite as a toy; mushroom as a vegetable). At test, the participants were simply asked to say the first five members of each category that came to mind. In this condition, adults showed 2½ times as much implicit memory as 5- to 7-year-old children and 1½ times as much implicit memory as 10- to 13-year-old children. Murphy et al. also used a very similar condition involving explicit memory and found the usual superiority in recall of adults over young children.

It could be argued that Murphy et al. hadn't really demonstrated changes in implicit memory with age because there was a possibility that explicit processes were involved despite the instructions they used. However, no participants in the implicit memory condition reported explicitly trying to recall information from the study list. Murphy et al. (2003) pursued this issue. Children aged 4 or 5 and adults performed the same implicit memory task as before under conditions involving full attention (only study list) or divided attention (detect sequences of odd or even numbers as well as study list). The assumption was that it would be very difficult to use explicit memory under divided attention conditions. Implicit memory increased with age even when the opportunity to use explicit processes was much reduced (see figure on the following page).

Discussion points

1. What are some of the practical implications of this research?
2. What are the main differences between implicit and explicit memory?



Implicit memory (priming) in kindergarten (K) and adult (A) groups in full-attention and divided-attention conditions. From Murphy et al. (2003). Copyright © 2003 Elsevier. Reproduced with permission.

KEY STUDY EVALUATION

The study by Murphy et al. (2003) is important because it shows that there can be large age effects in implicit memory. It also identified increases in children's knowledge base over time as the probable factor determining whether increasing age is accompanied by improved implicit memory. However, the study is limited in that it doesn't conclusively eliminate the possibility that explicit memory was involved in the allegedly implicit memory condition. In addition, it remains unclear precisely how increased knowledge helps to produce enhanced implicit memory performance.

In sum, implicit memory is more basic than explicit memory, and makes much less use of metamemory and strategic processes. This helps to explain why there are generally few or no developmental changes in implicit memory. However, implicit memory may differ less dramatically from explicit memory than is sometimes supposed, in that both forms of memory are influenced by the amount of relevant knowledge possessed by participants.

EYEWITNESS TESTIMONY

There has been a substantial increase in the number of children testifying in court in sexual abuse cases. This raises important issues about the accuracy of the information they provide, and has led to much research (see Bruck & Ceci, 1999, for a review). It is worrying that many children seem to be suggestible, which can lead to systematic errors in their recall of events. In general, suggestibility is greater during the early years of childhood than the later ones (Bruck & Ceci, 1999).

In a representative study, Thompson, Clarke-Stewart, and Lepore (1997) studied the effects of interviewer bias on children's memories. Five- and 6-year-olds witnessed one of two events. In the innocent event, a janitor called Chester cleaned some dolls and other toys in a playroom. In the abusive event, Chester handled the dolls roughly and in a mildly abusive way. Some children were then questioned by an accusatory interviewer who suggested the janitor had been abusive. Other children were questioned by an exculpatory [free from blame] interviewer who suggested the janitor was innocent. The remaining children were questioned by a neutral interviewer who avoided making any suggestions. The children were asked by their parents to describe what the janitor had done immediately after the interview and 2 weeks later.

Thompson et al. (1997) found that children's eyewitness memories were generally accurate when questioned by the neutral interviewer. However, when questioned by the accusatory or exculpatory interviewer, the children's accounts typically conformed to the interviewer's suggestions. Thus, the janitor was reported by the children as having behaved abusively when the interviewer was accusatory but as having behaved innocently

when the interviewer was exculpatory. When the children were then asked neutral questions by their parents, their descriptions of the event were consistent with what they had said to the interviewer.

It is perhaps unsurprising that young children were influenced by blatant interviewer bias in the Thompson et al. (1997) study. However, children are fairly suggestible in many situations. For example, Bruck, Ceci, and Hembrooke (1997) asked preschool children on five separate occasions to describe two true events (e.g., a recent punishment) and two false events (e.g., witnessing a thief stealing food). By the third interview, nearly all the children accepted that both the false events had actually happened. The children continued to argue that the false events were true when questioned later by a new interviewer adopting a nonsuggestive approach. The descriptions of the true and false events were similar. For example, they contained similar numbers of spontaneous statements and details (e.g., about conversations).

Why do children produce systematically distorted reports of events when exposed to suggestive influences? Roebbers and Schneider (2005) argued that two main factors are involved. First, there is social compliance. Children may yield to social pressure and a lack of social support even when their own recollection is accurate. Second, there is cognitive incompetence. Children may come to believe their own distorted memory reports because of limitations in processing, attention, or language.

The evidence suggests that both factors play a part. The notion that social compliance is important is supported by the work of Poole and Lindsay (1996). Children produced false memories after being repeatedly questioned by a suggestive interviewer. However, many of these false memories faded when the children weren't re-interviewed for a reasonably long period of time. The notion that children's memories may sometimes actually be altered is supported by several studies in which children continued to produce false memories after being warned that the interviewer may have been mistaken in his/her suggestions (see Bruck & Melnyk, 2004, for a review).

What can be done to increase the accuracy of children's eyewitness reports? Several answers are discussed by Baddeley, Eysenck, and Anderson (2009), but we will focus on one of the simplest and effective strategies. Gross and Hayne (1999) argued that much of the information about an event stored away by young children is likely to be in non-verbal form. If so, recall might be improved by asking children to draw what they could remember about an event before asking them to describe what they had seen. In their study, Gross and Hayne asked 5- and 6-year-old children to remember a visit to a Cadbury's chocolate factory 1 day, 6 months, and 1 year later. At the two shorter intervals, children in the drawing condition recalled 30% more information than those who only provided verbal reports. After 1 year, children in the drawing condition recalled almost twice as much information as those in the verbal-report-only condition.

LANGUAGE DEVELOPMENT

Young children seem to acquire language with breathtaking speed. From the age of about 16 months onwards, children often acquire upwards of 10 new words every day. By the age of 5, children have mastered most of the grammatical rules of their native language. However, very few parents are consciously aware of the rules of grammar. Thus, young children simply "pick up" the complex rules of grammar without much formal teaching.

STAGES OF LANGUAGE DEVELOPMENT

We need to distinguish between *receptive language* (language comprehension) and *productive language* (language expression or speaking). Children (and adults) have better receptive than productive language. For example, some young children produce only a few words but nevertheless have a comprehension vocabulary in excess of 150 words (Bates, Bretherton, & Snyder, 1988).

Children need to learn four kinds of knowledge about language:

1. *Phonology*: The sound system of a language.
2. *Semantics*: The meaning conveyed by words and sentences.

3. **Syntax:** The set of grammatical rules indicating how words may be combined to make sentences.
4. **Pragmatics:** The principles determining how language is modified to fit the context (e.g., we speak more simply to a child than to an adult).

Children generally acquire the above kinds of knowledge in the order listed. They initially learn to make sounds, followed by developing an understanding of what those sounds mean. After that, they learn grammatical rules, and how to alter what they say to fit the situation. Evidence that some aspects of phonological development occur remarkably early in life was reported by Mehler et al. (1994). Four-day-old French infants could discriminate between the French and Russian languages, showing a clear preference for French.

Early vocalizations

The babbling of infants up to 6 months of age is similar in all parts of the world. By about 8 months of age, however, infants show signs of the language they have heard in their vocalizations. Indeed, adults can sometimes guess accurately from their babbling whether infants have been exposed to French, Chinese, Arabic, or English (De Boysson-Bardies, Sagart, & Durand, 1984).

Up until the age of 18 months, young children are limited to single-word utterances (although they may be trying to convey much more meaning than their short utterances would suggest). Almost two-thirds of the words used by young children in Europe and the United States are nouns referring to objects or to people. Children naturally refer to things of interest to them (e.g., the people and objects surrounding them). Gentner (1982) found the same emphasis on nouns in children learning Japanese and Kaluli. Words are sometimes used to cover more objects than they should (over-extension), such as when my younger daughter referred to every man as “Daddy.” Words are also sometimes used too narrowly (under-extension). For example, a child may think the word “cereal” refers only to the brand of cereal he/she eats for breakfast.



Overextension occurs when a child hears a word in the presence of an object, and then goes on to label all similar things with that same word. Hence, all furry creatures with four legs may be called a “cat.”

Telegraphic period

The second stage of language development starts at about 18 months. This is the **telegraphic period**, during which children’s speech is abbreviated like a telegram costing so much per word. Content words such as nouns and verbs are included, but function words (e.g., “a”; “the”; “and”), pronouns, and prepositions are omitted.

Even though young children are mostly limited to two-word utterances, they can still communicate numerous meanings. For example, “Daddy chair” may mean “I want to sit in Daddy’s chair,” “Daddy is sitting in his chair,” or “Daddy, sit in your chair!”

Brown (1973) argued that young children possess a basic order rule: a sentence consists of agent + object + location (e.g., “Daddy eats lunch at home”). Their two-word utterances follow the basic order rule. For example, an utterance containing an agent and an action will be in the order agent–action (e.g., “Daddy walk”) rather than the reverse “walk Daddy”). Similarly, action and object will be spoken in the order action–object (e.g., “drink Coke”).

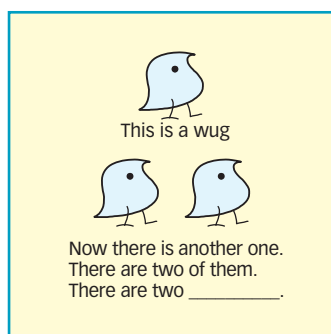
Subsequent developments

Children’s language develops considerably between 2 and 5 years of age. For example, the maximum sentence measured in terms of morphemes (meaningful units) increases from four morphemes at 24 months to eight morphemes at 30 months (Fenson et al., 1994). Children gradually acquire various **grammatical morphemes**, which are modifiers that alter meaning. Examples include prepositions, prefixes, and suffixes (e.g., “in”; “on”; plural –s; “a”; and “the”). Nearly all children learn the various grammatical morphemes

Key Terms

Telegraphic period: a stage of language development in which children use short but informative utterances, as in a telegram.

Grammatical morphemes: modifiers (e.g., prefixes; suffixes) that alter the meaning of words and phrases.



in the same order, starting with simple ones (e.g., including “in” and “on” in sentences) and moving on to more complex ones (e.g., reducing “they are” to “they’re”).

Do children simply imitate adults’ speech rather than learning rules? Evidence that they don’t comes from children’s grammatical errors. For example, a child will say, “The dog runned away,” which adults are unlikely to do. Presumably the child makes that mistake because he/she is applying the rule that the past tense of a verb is usually formed by adding –ed to the present tense. Using a grammatical rule in situations in which it doesn’t apply is known as **over-regularization**. Evidence that over-regularization doesn’t occur simply because children are imitating other children was reported by Berko (1958). Children were shown two pictures of an imaginary animal

or bird. They were told, “This is a wug. This is another wug. Now there are two . . .” (see figure on the left). Even young children produced the regular plural form “wugs” despite not having heard the word before.

At one time, it was thought that the development of vocabulary and of grammar occurred independently. In fact, the two forms of development are actually closely related. For example, Dionne, Dale, Boivin, and Plomin (2003) assessed vocabulary and grammar in children at the ages of 2 and 3. Vocabulary at age 2 predicted grammar at age 3, and grammar at age 2 predicted vocabulary at age 3. Such findings suggest (but don’t prove) that the processes underlying vocabulary and grammar development are similar. Conboy and Thal (2006) studied children between the ages of 20 and 30 months who were bilingual in English and Spanish. Their key finding was that grammatical development (e.g., complexity of utterances) in each language was associated mainly with vocabulary development in the same language. Thus, grammatical development within a given language depends more on the child’s vocabulary in that language than on its more general maturational or conceptual development.

THEORIES OF LANGUAGE DEVELOPMENT

Numerous theories have been put forward to explain children’s language acquisition. Most can be categorized as inside-out theories or outside-in theories (Hirsch-Pasek & Golinkoff, 1996). Inside-out theorists (e.g., Chomsky; Pinker) argue that language acquisition depends heavily on innate factors and only modestly on the child’s own experiences. In contrast, outside-in theorists (e.g., Bruner; Tomasello) argue that experience is of central importance in children’s language acquisition.

There are other differences between inside-out and outside-in theorists (Harris & Butterworth, 2002). Most inside-out theorists claim that language development occurs in relative isolation from other forms of cognitive and social development. However, outside-in theorists argue that general cognitive and social mechanisms (e.g., those involved in perception or thinking) are involved in language development. These two theoretical positions are discussed below. Note that it is entirely possible that there is an element of truth in both positions, with innate factors *and* experience both being of vital importance in language development.

Early language acquisition

Age	0–6 months	6 months–1 year	1–2½ years	2–5 years
Babbling	✓			
Some phonemes learned	✓	✓		
First spoken word		✓		
Beginning of grammatical rules			✓	
Basic rules of grammar acquired				✓

Key Term

Over-regularization: extending a grammatical rule to situations in which it does not apply.

Inside-out theories

Inside-out or nativist theorists are very impressed by the speed with which most young children acquire language. They claim that language acquisition is simply too fast for it to depend entirely on learning by experience. Accordingly, they argue that infants are born with knowledge of the structure of human languages. For example, Chomsky (1965) argued that humans possess a **language acquisition device** consisting of innate knowledge of grammatical structure. Children require some exposure to (and experience with) the language environment provided by their parents and other people to develop language. Such experience determines which specific language any given child will learn.

Chomsky (1986) later replaced the notion of a language acquisition device with the idea of a universal grammar, which forms part of our innate knowledge of language.

According to his Principles and Parameters Theory, there are **linguistic universals**, which are features common to nearly every language. There are substantive and formal universals. Substantive universals concern categories common to all languages (e.g., noun and verb categories). Formal universals are concerned with the general form of syntactic or grammatical rules.

Pinker (1984, 1989) is broadly sympathetic to Chomsky's approach. However, he argued that exposure to language is more important than admitted by Chomsky. According to Pinker, children use "semantic bootstrapping" to allocate words to their appropriate word class. Suppose a young child hears the sentence, "William is throwing a stone," while watching a boy carrying out that action. The child will realize from his/her observations that "William" refers to the actor, "stone" to the object acted on, and "is throwing" to an action. The child then uses his/her innate knowledge of word categories to work out that "William" is the subject of the sentence, "stone" is the object, and "throwing" is the verb.

Lenneberg (1967) put forward the critical period hypothesis. According to this hypothesis, language learning depends on biological maturation, and so is easier prior to puberty. This might help to explain why most people report that it was easier to learn their own language than to learn a second language when they were teenagers.

Findings

Chomsky argued that word order was a linguistic universal. We can test this by considering the preferred word order for the subject, verb, and object in sentences in numerous languages. In principle, there are six possible orderings, two of which (object-verb-subject; object-subject-verb) are not found among the world's languages (Greenberg, 1963). The most popular word order is subject-object-verb (44% of languages), followed by the subject-verb-object word order found in English (35% of languages). The subject precedes the object in 98% of languages, presumably because it makes sense to consider the subject (typically the actor) of a sentence before the object.

Bickerton (1984) put forward the language bioprogram hypothesis, according to which children will create a grammar even if not exposed to a proper language during their early years. Support was obtained by considering laborers from China, Japan, Korea, Puerto Rico, Portugal, and the Philippines taken to the sugar plantations of Hawaii 100 years ago. These laborers developed a pidgin language which was very simple and lacked most grammatical structures. Here is an example: "Me cape buy, me check make." The meaning is, "He bought my coffee; he made me out a check" (Pinker, 1984). The offspring of these laborers developed a language known as Hawaiian Creole. This is a proper language and is fully grammatical. Here is an example of this language: "Da firs japani came ran away from japan come." This means, "The first Japanese who arrived ran away from Japan to here."

We don't know the extent to which the development of Hawaiian Creole depended on the laborers' prior exposure to language. Clearer evidence that a language can develop in groups almost completely lacking in exposure to a developed language was reported by Senghas, Kita, and Özyürek (2003). They studied deaf Nicaraguan children at special schools. Attempts (mostly unsuccessful) were made to teach them Spanish. However, these deaf children began to develop a new system of gestures that expanded into a basic sign language passed on to successive groups of children who joined the school. Since

Key Terms

Language acquisition device:

an innate knowledge of the grammatical structure of language.

Linguistic universals:

features that are found in all (or virtually all) languages.

Case Study: *Genie*

Genie spent most of her time up to the age of 13 in an isolated room (Curtiss, 1977). She had practically no contact with other people, and was punished if she made any sounds. After Genie was rescued in 1970, she learned some aspects of language, especially vocabulary. However, she showed very poor learning of grammatical rules. There are problems in interpreting the evidence from Genie. She was exposed to great social as well as linguistic deprivation, and her father's "justification" for keeping her in isolation was that he thought she was very retarded. Thus, there are various possible reasons for Genie's limited ability to learn language.

Ethical issues: Deprivation studies are useful examples from which we can draw some inferences, but they rarely provide data that can be regarded as scientific. What are some of the ethical issues that arise from looking at the effects of deprivation? Should the psychologist be concerned with compensation for the deprivation experienced, e.g., linguistic support for individuals like Genie?

Do the ethical problems concerning work like this outweigh any practical advancement of our understanding as psychologists?

Nicaraguan Sign Language bore very little relation to Spanish or to the gestures made by hearing children, it appears that it is a genuinely new language owing remarkably little to other languages.

There is mixed evidence relating to the critical period hypothesis. Support for the hypothesis was reported by Johnson and Newport (1989), who studied Chinese and Korean immigrants to the United States. Those who had arrived in the United States at an early age performed much better than those who had arrived later when asked to decide whether sentences were grammatically correct. That suggests that there may be a critical period for the learning of syntax.

In principle, the best way to test the critical period hypothesis would be to consider children having little chance to learn language during their early years. For example, Genie spent most of her time up to the age of 13 in an isolated room (Curtiss, 1977). She had practically no contact with other people, and was punished if she made any sounds. After being rescued, she managed to learn the meanings of numerous words, but her learning of grammatical rules was very poor. These findings seem consistent with the critical period hypothesis. However, there are various possible reasons for Genie's poor learning of language. She was exposed to great social as well as linguistic deprivation, and her father claimed that he thought she was very retarded.

Evaluation

- + Nativist theories potentially explain why nearly all children master their native language very rapidly.
- + Chomsky's theory is supported by the way in which pidgin languages develop into creole languages and by the development of Nicaraguan Sign Language.
- + Chomsky's theory makes sense of the fact that language is rule-based even though few speakers of a language can express these rules clearly.
- + There is support for a weakened version of the critical period hypothesis, according to which some aspects of language are more difficult to acquire outside the critical period (Harley, 2001).
- Chomsky's theory is difficult to test. For example, he assumed that very young children have access to a considerable amount of grammatical knowledge. However, if their language performance doesn't reflect this assumed knowledge, this can be explained away by arguing that the children were not motivated, had problems with attention or memory, and so on.
- Chomsky argued that the language children hear doesn't contain enough information to allow them to work out grammatical rules from scratch. This argument is not persuasive. As discussed later, mothers and other adults typically talk to young children in simple, short sentences so as to facilitate their language acquisition.
- The entire idea of an innate grammar seems implausible. According to Bishop (1997, p. 123), "What makes an innate grammar a particularly peculiar idea is the fact that innate knowledge must be general enough to account for acquisition of Italian, Japanese, Turkish, Malay, as well as sign language acquisition by congenitally deaf children." Tomasello (2005) described the problem of relating a universal innate grammar to any given language as the "linking problem."

Outside-in theories

Outside-in theories (e.g., Tomasello's constructivist theory) emphasize the central role of experience in allowing young children to acquire language. In essence, outside-in theorists argue that the language input to which young children are exposed is adequate for language acquisition, whereas inside-out theorists are skeptical. We start with a discussion of research on child-directed speech, which is very relevant to this disagreement.

When mothers (and fathers) talk to their young children, they use very short, simple sentences; this is known as **child-directed speech**. Other features of child-directed speech are that it involves a slow rate of speaking, use of a restricted vocabulary, and extra stress on key words (Dockrell & Messer, 1999). Mothers, fathers, and other adults also try to help children's language development by expansions. **Expansions** consist of fuller and more grammatical versions of what the child has just said. For example, a child might say, "Dog out," with its mother responding, "The dog wants to go out."

Saxton (1997) argued that many expansions provide children with an immediate contrast between their own incorrect speech and the correct version. For example, a child may say, "He shooted fish," to which the adult might reply, "He shot the fish!" Children seem to process expansions fairly thoroughly, because they are more likely to repeat adult expansions than other adult utterances (Farrar, 1992).

Child-directed speech helps children to learn vocabulary. For example, Weizman and Snow (2001) considered factors influencing children's vocabulary at the age of 7. Children with the largest vocabularies had mothers who included many sophisticated words when speaking to them. In addition, their mothers tended to introduce these sophisticated words in helpful ways (e.g., providing explicit information about the meaning of each word).

From what has been discussed so far, it may seem as if we know for sure that language acquisition relies heavily on child-directed speech. In fact, however, finding that children having the most exposure to child-directed speech acquire language the fastest doesn't prove that the child-directed speech was responsible. Another possibility is that adults use more child-directed speech with children who find it easiest to acquire language.

Cross-cultural research suggests that child-directed speech may *not* be essential for language acquisition. Schieffelin (1990) studied the Kaluli of New Guinea. Kaluli adults talk to children as if they were adults, but Kaluli children develop language at about the normal rate. However, cultures in which child-directed speech is little used may provide different kinds of assistance to children learning language. For example, Ochs and Schieffelin (1995) argued that children in such cultures become involved in social and communal activities, which assists shared understanding and language development.

In an important study, Haggan (2002) discovered many Kuwaiti mothers who claimed they did not use child-directed speech with their children. However, observations of their actual behavior indicated some evidence that all of them did actually make use of child-directed speech.

Cognitive and constructivist approaches

Piaget (see Chapter 14) argued that children acquire language only after developing relevant cognitive abilities. More specifically, he claimed that children form schemas consisting of organized knowledge about the world and events they have experienced. Children need to have



Child-directed speech helps young children to acquire language by using very short, simple sentences, often with target words receiving extra stress "What a pretty DAISY!"

Key Terms

Child-directed speech:

the simplified sentences spoken by mothers and other adults when talking to young children.

Expansions:

fuller and more detailed versions of what a child has just said provided by an adult or an older child.

developed various cognitive processes and structures (e.g., schemas) before they can develop language. This is the cognition hypothesis.

It follows from Piaget's cognition hypothesis that children having slow cognitive development should also have slow linguistic development. This is often (but by no means always) the case. For example, consider children with **Williams syndrome**. This is a rare genetic disorder involving unusual facial characteristics ("elfin face") and typically an IQ in the 50–65 range. They have slow cognitive development. However, their language abilities are relatively good and they often have surprisingly large vocabularies (Tager-Flusberg, 1999). Such findings seem at odds with the cognition hypothesis. However, the onset of language in children with Williams syndrome is later than in healthy children, and their language comprehension is poor (Karmiloff-Smith, Grant, Berthoud, Davies, Howlin, & Udwin, 1997).

Tomasello has developed a constructivist theory based in part on Piaget's views. The essence of his theory is as follows:

Children acquire linguistic competence in the particular language they are learning only gradually, beginning with more concrete linguistic structures based on particular words and morphemes [small units of meaning], and then building up to more abstract and productive structures based on various types of linguistic categories, schemas, and constructions (Tomasello & Brooks, 1999, p. 161).

According to Tomasello's theory, children's language development is based on their cognitive understanding of the scenes or events they experience in their everyday lives. They generally find it easiest to learn nouns, because nouns refer to concrete objects or people within scenes. It is more difficult to learn verbs, because their meaning is typically more abstract. Tomasello (1992) put forward the Verb Island hypothesis, according to which children initially treat each verb independently as if it had its own island of organization.

We can relate the central assumptions of constructivist and nativist theories to the issue of language productivity or creativity. According to Tomasello's constructivist theory, young children should initially show very limited productivity. According to nativist theories, young children should rapidly show high levels of productivity or creativity in their language utterances.

Findings

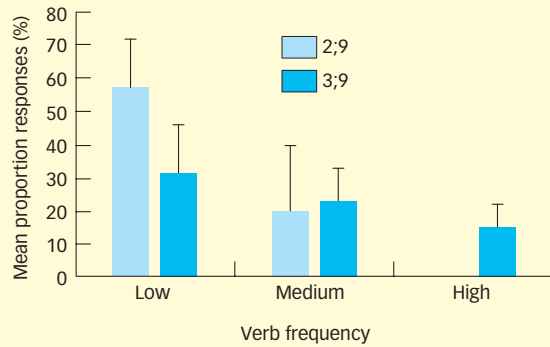
A central assumption of the constructivist approach is that initial language learning (especially of verbs) should be slow. Supporting evidence was reported by Tomasello (1992) based on observations of his 1-year-old daughter. She used some verbs in only one kind of sentence frame (e.g., "Cut ____"), which indicates limited learning. Other verbs were used in several kinds of sentence frames (e.g., "Draw ____"; "Draw ____ on ____"; "Draw ____ for ____"). Such differences among verbs are consistent with the Verb Island hypothesis discussed above.

Tomasello, Akhtar, Dodson, and Rekau (1997) taught children aged between 18 and 23 months two new nouns and two new verbs. The children produced 10 times more combinations of words using the novel nouns than the novel verbs. This relative lack of productivity with novel verbs is more consistent with the constructivist approach than with the nativist one.

According to Tomasello's theory, most language processes develop over time as a result of increased exposure to language. Thus, the learning of specific language skills should be gradual. In contrast, nativist theories tend to assume that specific language skills should emerge suddenly. Matthews, Lieven, Theakston, and Tomasello (2005) tested these predictions in a study using two groups of children, one with a mean age of 2 years 9 months and the other with a mean age of 3 years 9 months. The children were presented with short sentences containing low-, medium-, and high-frequency verbs in an ungrammatical subject-object-verb order (e.g., "Bear elephant tugged").

Key Term

Williams syndrome: a genetic disorder involving low IQ but reasonable language development.



Mean proportion of responses copying the ungrammatical word order heard in younger children (mean age 2 years 9 months) and older children (mean age 3 years 9 months). From Matthews et al. (2005). Copyright © 2005 Elsevier. Reproduced with permission.

Did the children in their own speech copy the ungrammatical order they had heard or did they use the grammatically correct order? Both groups of children (especially the younger group) were more likely to copy the ungrammatical order with low-frequency verbs than with high-frequency ones (see the figure above). Thus, using the correct grammatical order in spite of having heard an incorrect grammatical order is a gradual process that is found initially mainly with very familiar verbs.

Evaluation

- + The notion that language acquisition depends on various cognitive processes and is facilitated by child-directed speech is plausible.
- + The finding that children gradually learn to use the appropriate word order (e.g., Matthews et al., 2005) is more in line with the constructivist approach than with the nativist one.
- + Children show relatively little productivity early in language acquisition (especially with verbs), as predicted by the constructivist approach.
- + Children acquiring English find verb learning more difficult than noun learning, as predicted by Tomasello's theory.
- More thorough longitudinal research is needed to show how cognitive processes influence language acquisition.
- We need more detail on the processes children use in proceeding from initial concrete linguistic structures to more general abstract schemas.

THEORY OF MIND

One of the crucial differences between most 5-year-olds and most 2- or 3-year-olds is that the former understand that other people's beliefs about the world may differ from their own. This is really important, because social communication is limited if a child assumes that everyone else has the same beliefs that he/she has. Research in this area revolves around the notion of **theory of mind**, which "conveys the idea of understanding social interaction by attributing beliefs, desires, intentions, and emotions to people" (Astington & Jenkins, 1999, p. 1311).

Theory of mind has been assessed by using various false-belief tasks. For example, Wimmer and Perner (1983) used models to present children with the following story. A boy called Maxi puts some chocolates in a blue cupboard. While he is out of the

Key Term

Theory of mind: the understanding that other people may have different beliefs, emotions, and intentions than one's own.

Theory of mind

According to Premack (2007), *"When an infant sees an individual in distress, for example, a child who cries when her teddy bear breaks, the infant consoles her, pats her, speaks softly to her, and may even try to fix the teddy bear."* This illustrates theory of mind, i.e., being able to understand other people's feelings and beliefs. It is thought that this attribute is crucial for good social cognition, and research suggests that some other animals may have a degree of theory of mind. Obviously this would be advantageous to social animals, such as chimpanzees. Premack describes a laboratory experiment where young chimpanzees chose between two containers. Only one of the containers held a food reward. The apes could not see the reward being placed but they could see that a human was sometimes able to observe this. When allowed to seek help from the human three out of four chimps pointed to the person who had been able to observe the food being placed. This suggests that they did have the understanding that the human's perception was different from their own, that they do have a theory of mind.

Goodall's view is even stronger. In a 2004 UK BBC *Horizon* television program she said *"There's absolutely no question that chimpanzees understand the needs and the emotions of other chimpanzees and respond correctly. They can even understand the needs of another human being, so clearly they do have theory of mind."* Her view is supported by experiments at the Max Planck Institute involving subordinate and dominant chimpanzees. Normally a subordinate will not take food if a dominant is present. But if a subordinate chimp saw a banana being hidden out of the dominant chimp's view then the former would snatch it. This clearly implies that the subordinate chimp understood that the other chimp's perception and knowledge was different from its own.

room, his mother moves the chocolate to a green cupboard. The children indicated where Maxi would look for the chocolate when he returned to the room. Most 4-year-olds argued mistakenly that Maxi would look in the green cupboard. This indicates an absence of theory of mind, because these children simply assumed that Maxi's beliefs were the same as their own. In contrast, most 5-year-olds produced the right answer.

Similar findings have been reported in numerous other studies. Wellman, Cross, and Watson (2001) reported a meta-analysis based on 178 false-belief studies. In general, most 3-year-olds performed poorly on false-belief tasks, whereas a substantial majority of 5-year-olds were correct. The findings were similar in seven different countries (United States, United Kingdom, Korea, Australia, Canada, Austria, and Japan), except that a theory of mind developed slowest in Austria and Japan. Thus, evidence from several countries suggests that a theory of mind develops at about the age of 4.

Flynn (2006) studied the development of theory of mind in detail over a period of several months. She found that young children, most of whom were 3 years old at the start of the experiment, showed a gradual development in the understanding of false beliefs on theory-of-mind tests. For some time before the children were able to explain other people's behavior with reference to their false beliefs they were in a state of confusion in which they failed to provide any explanation.

How can we explain children's development of theory of mind? Various answers have been suggested, of which we will focus on two. Astington and Jenkins (1999) argued that language development is central to children's development of theory of mind. In contrast, Riggs, Peterson, Robinson, and Mitchell (1998) argued that what is crucial is the development of reasoning ability. For example, children given the Maxi problem have to imagine a state of affairs (i.e., chocolate in the blue cupboard) that would exist if his mother hadn't moved the chocolate to the green cupboard). This is known as counterfactual reasoning, and was assumed by Riggs et al. to be the key to theory of mind.

FINDINGS

While it is generally assumed that theory of mind typically develops around the age of 4, some aspects of theory of mind may develop earlier. For example, O'Neill (1996) carried out a study in which 2-year-old children watched an attractive toy being placed on a high shelf in the presence or absence of their parent. The children subsequently asked their parent to let them have the toy. Those children whose parent had been absent previously were much more likely to name the toy and to gesture towards it than were those whose parent had been present. Thus, even 2-year-olds may have some awareness of the knowledge possessed by others.

The role of language ability in theory of mind was examined by Lohmann, Carpenter, and Call (2005) on two false-belief tasks with 3- and 4-year-old children. Their key finding was that performance on both tasks was better among children with higher vocabulary and grammar scores. Astington and Jenkins (1999) carried out a longitudinal study on 3-year-olds to test the assumption that language development underlies improvement on false-belief tasks with age. The children were given various theory-of-mind tasks (e.g., false-belief tasks) and measures of language competence three times over a 7-month period. Language development at one point in time predicted later theory-of-mind performance. Similar findings were reported by Slade and Ruffman (2005) over a 6-month period. These findings suggest that language ability may play a causal role in the development of theory of mind.

According to Riggs et al. (1998), the ability to engage in counterfactual reasoning underlies successful performance on false-belief tasks. They tested their theory by presenting a story to 3- and 4-year-old children. In this story, Maxi and his mother put chocolate in a cupboard. While Maxi is out at school, his mother uses some of the chocolate to make a cake, putting the rest of it in the fridge. After that, Maxi comes home. The children were asked a standard false-belief question: "Where does Maxi think the chocolate is?" They were also asked a second question: "If mum had not made a cake, where would the chocolate be?" Answering this question correctly required counterfactual reasoning but not an understanding of false beliefs.

What did Riggs et al. (1998) discover? Children who were correct on the false-belief question were nearly always correct on the reasoning problem, and those incorrect on the false-belief question nearly always failed the reasoning problem. Thus, they concluded that deficient counterfactual reasoning is the main reason for failure on false-belief tasks. However, matters may be more complicated than that. Perner, Sprung, and Steinkogler (2004) found that 3½-year-olds achieved 75% correct performance in simple counterfactual reasoning but only achieved that level on false-belief questions 1½ years later. Thus, the ability to engage in counterfactual reasoning is clearly not sufficient to guarantee success on false-belief tasks.

Not everyone thinks that theory of mind depends on *general* abilities such as language or counterfactual reasoning. For example, Saxe, Carey, and Kanwisher (2004) carried out a meta-analysis of brain-imaging studies on false-belief tasks performed by adults. Theory-of-mind processing was associated with activation in parts of the medial prefrontal cortex and the temporo-parietal junction toward the back of the brain. These areas are *not* associated with general processes relating to language or attention. Theory-of-mind tasks may involve relatively *specific* processes (whose nature is poorly understood).

AUTISM

Autism is a serious condition characterized by very poor social interaction, impaired communication (e.g., reluctance to maintain a conversation), and repetitive patterns of behavior. According to various theorists (e.g., Baron-Cohen, Leslie, & Frith, 1985; Leslie, 1987), the central problem of autistic children is that they lack a theory of mind. More specifically, they don't understand that other people have different ideas and knowledge to their own, and they don't appreciate that behavior is influenced by beliefs and thoughts. As a result, autistic children can't make sense of the social world, and this prevents them from communicating effectively with other people.



O'Neill's (1996) study, in which 2-year-olds gestured towards and named the toy that was placed in the room while their parents were absent, demonstrates a possession of at least some aspects of theory of mind in young children.

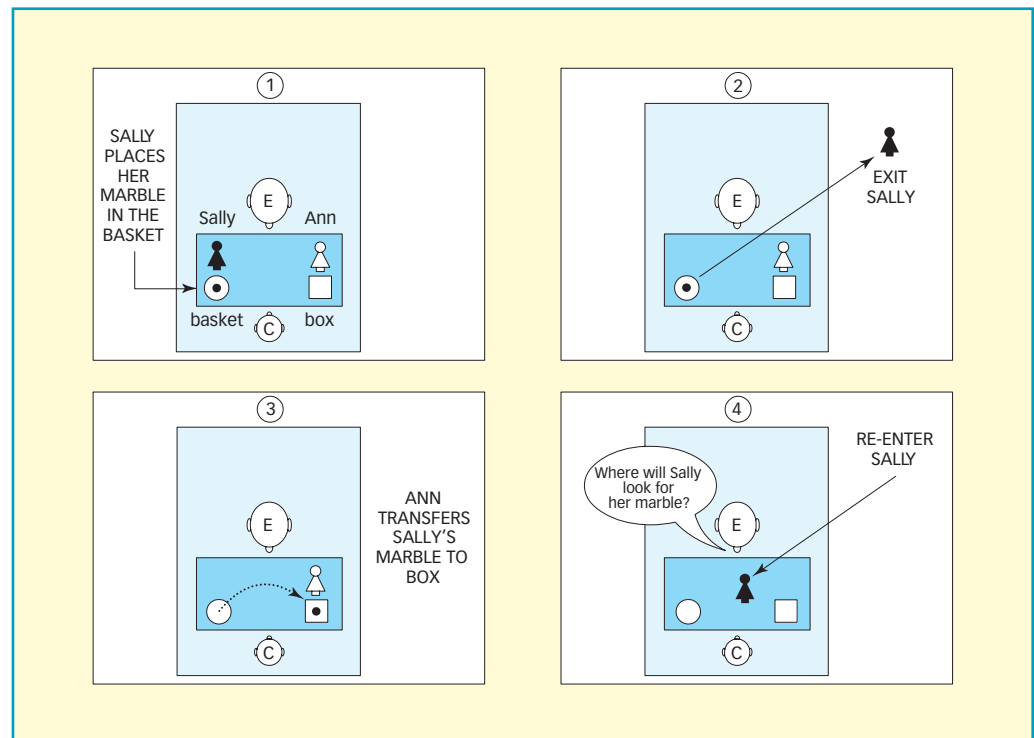
Key Term

Autism: a severe disorder involving very poor communication skills, deficient social and language development, and repetitive behavior.

Baron-Cohen et al. (1985) studied healthy 4-year-old children as well as autistic and Down's syndrome children with a mental age of at least 4. All these children were given the following story (see the figure below):

Sally puts her marble in the basket. Then she goes out. Anne takes Sally's marble, and puts it into her box. Then Sally comes back from her walk. Where will she look for her marble?

The Sally–Anne test. C denotes the child observer, and E is the experimenter.



More than 80% of the healthy and Down's syndrome children correctly pointed to the basket, compared to only 20% of the autistic children. Thus, the autistic children didn't understand that other children may have a different perspective to their own.

Autistic children's poor performance on false-belief tasks may well reflect fairly *general* cognitive deficits. Performance on false-belief tasks depends on executive functions (e.g., Müller, Zelazo, & Imrisek, 2005). Executive functions include planning, shifting set, and inhibition of dominant responses, and there is evidence that autistic children are impaired with respect to these executive functions (see Hill and Frith, 2003, for a review). For example, Hughes and Russell (1993) used a task in which participants had to turn a knob or flick a switch to obtain a marble in a box, but could not obtain it by reaching into the box. Autistic children found it much harder than children with learning difficulties to inhibit the response of reaching immediately for the marble. Planning in autistic children has been studied by observing their performance on the Tower of Hanoi, in which disks have to be moved according to various rules to attain a specified goal state. It has been found consistently that autistic children show very poor planning on this task (Hill & Frith, 2003).

Case Study: *Autistic Talents*

Not all aspects of autism are as negative as might be imagined. Some autistic children have startling artistic abilities and can produce drawings in full detail and perspective much earlier than other children.

An autistic girl named Nadia was studied by Selfe (1983). When she was only 5 years old, Nadia could draw realistic pictures of horses, cockerels, and cavalymen from memory, although she did not speak and had various severe motor problems.

Other talents shown by autistic children and adults include feats of mental arithmetic, for example being able to calculate the day of the week for any given date in the previous 500 years. There have also been gifted autistic musicians who learn to play musical instruments by ear, with no formal training.

These talents may all be linked in some way to the autistic child's narrow focus on the world, through which they can become preoccupied with certain objects or processes in great detail.

Kanner (1943) called gifts like these "islets of ability," which suggests that other aspects of autistic children's intelligence are hidden beneath the surface of a sea of difficulties.



Drawings by 5-year-old Nadia, who is autistic (left), and an average 6½-year-old child (right).

Overall Evaluation

- + The development of a theory of mind is of real importance in allowing children to communicate well with other people.
- + Some aspects of cognitive development (e.g., counterfactual reasoning; language) associated with the development of a theory of mind have been identified.
- + Autistic children's difficulties with false-belief tasks reflect at least in part problems with general executive functions.
- Deficiencies in executive functions are found in several clinical conditions (e.g., attention-deficit/hyperactivity disorder), and so autism must involve deficiencies over and above those in executive functions.
- False-belief tasks are complex and there is no consensus concerning the balance of general and specific processes underlying performing on such tasks when carried out by healthy individuals (Apperly, Samson, & Humphreys, 2005).
- There is probably no *single* theory of mind. What is more likely is that children acquire a progressively deeper understanding of others' beliefs over time.
- The theory-of-mind approach fails to account for all of the problems faced by autistic children. As Smith, Cowie, and Blades (2003, p. 481) pointed out, "It is not obvious how specific language problems . . . obsessive behaviors, or 'islets of ability' [in autistic children] could be linked to a lack of understanding minds."

Chapter Summary

Perceptual development

- Newborns have poor visual acuity, no color vision, and lack binocular disparity.
- Newborns make systematic eye movements, show elements of size constancy, and can relate visual and auditory information.
- It has been claimed that newborns have an innate bias for faces. However, it is more likely that they have a preference for stimuli with more patterning in their upper part.
- Infants of 6 months show some elements of depth perception by avoiding going over the edge of a “visual cliff.” There is some evidence that learning with the visual cliff is specific to a given posture (e.g., sitting), and so new learning is needed when the infant starts crawling.
- Some theorists talk of the “competent infant” who possesses early perceptual competence accompanied by cognitive incompetence.

Memory development

- The capacity of the working memory system increases during childhood.
- Children make increased use of memory strategies (e.g., rehearsal; organization) as they grow older.
- Children gradually acquire metamemory or knowledge about their own memory.
- Children’s memory performance improves as they acquire relevant content knowledge.
- Children’s explicit memory (involving conscious recollection) improves much more than their implicit memory (not involving conscious recollection), because implicit memory depends much less on strategic processes and metamemory.
- Children’s eyewitness testimony can be distorted by suggestibility. The main factors responsible are social compliance and cognitive incompetence.

Language development

- Children typically acquire knowledge of language in the following order: phonology; semantics; syntax; and pragmatics.
- Children start with single-word utterances and then move into the telegraphic period at 18 months.
- Inside-out theorists (e.g., Chomsky) argue that language acquisition depends heavily on innate factors (e.g., a language acquisition device).
- Inside-out theorists argue that linguistic universals are found in nearly every language, and that there is a critical period during which language learning is easiest.
- There is some evidence that children can create a grammar even if not exposed to a proper language during their early years as predicted by inside-out theories.
- The entire idea of an innate grammar is implausible and difficult to test.
- Outside-in theorists (e.g., Tomasello) emphasize the central role of exposure to language (especially child-directed speech) in allowing young children to acquire language.
- As predicted by outside-in theories, most language learning is gradual.
- As yet, little is known of the precise cognitive processes involved in language acquisition.

Theory of mind

- Most 5-year-olds have developed a theory of mind, meaning that they understand that other people may have beliefs, knowledge, and emotions different from their own.
- Factors underlying the development of theory of mind include language development and the development of reasoning ability (especially counterfactual reasoning).
- Autistic children lack a theory of mind and so can’t make sense of the social world. Impaired executive functioning is one reason for their deficiencies in developing a theory of mind.
- Some of the problems of autistic children (e.g., specific language difficulties; obsessive behavior) aren’t caused by a lack of theory of mind.

Further Reading

- Baddeley, A.D., Eysenck, M.W., & Anderson, M.C. (2009). *Memory*. Hove, UK: Psychology Press. The development of memory during childhood is discussed in this introductory textbook.
- Goswami, U. (Ed.). (2006). *Cognitive development: Critical concepts in psychology*. Hove, UK: Psychology Press. This edited book contains chapters by leading experts on the major areas within cognitive development.
- Harris, M., & Butterworth, G. (2002). *Developmental psychology: A student's handbook*. Hove, UK: Psychology Press. The development of language in the early years of life is covered in an accessible way in Chapters 7 and 8 of this textbook.
- Slater, A.M., & Bremner, J.G. (Eds.). (2003). *An introduction to developmental psychology*. Oxford, UK: Blackwell. Issues relating to cognitive development are discussed in separate sections on infancy, childhood, and adolescence in this edited book.
- Smith, P.K., Cowie, H., & Blades, M. (2003). *Understanding children's development* (4th ed.). Oxford: Blackwell. Several chapters in this textbook (e.g., Chapters 10 and 11) provide detailed coverage of important aspects of cognitive development.

chapter 14

Contents

Piaget's theory	337
Vygotsky's theory	344
Contemporary approaches	347
Educational implications	350

Theories of cognitive development

14

In the last chapter, we considered the dramatic changes in perception, memory, and language that occur during childhood. In this chapter, our focus shifts to more general theoretical approaches to cognitive development. The question of central importance in this area of research is the following: How can we describe (and explain) the enormous advances in thinking, reasoning, and problem solving shown by nearly all children during the years of childhood?

The first general systematic theory of cognitive development was proposed by Jean Piaget (1896–1980), and has been the most influential developmental theory of all time. Another important theoretical approach to cognitive development was put forward by the Russian psychologist, Lev Vygotsky (1896–1934). He was one of untold millions who suffered under Stalin, with his various writings being suppressed in Russia. As a consequence, few people knew of his work during his life, which ended prematurely because of tuberculosis. Vygotsky's work began to be translated into other languages in the 1960s and 1970s, and was then recognized as of major importance.

The pioneering approaches of Piaget and Vygotsky are discussed in this chapter. After that, we switch our attention to more contemporary approaches. It is generally accepted that the best way of understanding cognitive development is to identify the *learning* processes used by children. This approach has been followed in several theories in which the emphasis is on the processes and structures associated with information processing and learning. In this chapter, we focus on Robert Siegler's (1998, 2005) influential theory.

A key reason why it is important to study cognitive development is because of its potential relevance to education. If we can understand the processes involved in learning and cognitive development, we will be well placed to improve the educational system. The educational implications of the theories discussed in this chapter are considered in the final section.

PIAGET'S THEORY

Piaget's central focus was on cognitive development. More specifically, he was interested in how children learn and adapt to the world. In order for adaptation to occur, there must be constant interactions between the child and the world. According to Piaget, two processes are of key importance:

- **Accommodation:** The individual's cognitive organization is altered by the need to deal with the environment. Thus, the individual adjusts to the outside world.
- **Assimilation:** The individual deals with new environmental situations on the basis of his/her existing cognitive organization. Thus, the individual child's interpretation of the outside world is adjusted to fit him/her.

Key Terms

Accommodation:

in Piaget's theory, changes in an individual's cognitive organization to deal with the environment.

Assimilation:

in Piaget's theory, dealing with new environmental situations by using existing cognitive organizations.



The Swiss psychologist Jean Piaget, 1896–1980.

The clearest example of the dominance of assimilation over accommodation is play. In play, reality is interpreted according to the individual's whim (e.g., a stick becomes a gun). In contrast, dominance of accommodation is seen in imitation. In imitation, the child simply copies the actions of someone else.

There are two other key Piagetian concepts: schema and equilibration. **Schema** refers to organized knowledge used to guide action. The first schema infants develop is the body schema, when they realize there is an important distinction between “me” and “not me.” This body schema helps the infant in its attempts to explore and make sense of the world.

Equilibration is based on the notion that the individual needs to keep a stable internal state (equilibrium) in a changing environment. When a child tries unsuccessfully to understand its experiences in terms of existing schemas, there is an unpleasant state of *disequilibrium* or lack of balance. The child then uses assimilation and accommodation to restore a state of equilibrium or balance. Thus, disequilibrium motivates the child to learn new skills and knowledge to return to the desired state of equilibrium.

We can identify two extreme positions with respect to the changes occurring during cognitive development. At one extreme, the amount of knowledge available to children increases considerably, but there are no dramatic changes in *how* they think. At the other extreme, the ways of thinking found in adolescence differ profoundly from those of early childhood. Piaget identified himself with the latter position, believing that there are fundamental differences in cognition among children of different ages.

Of key importance, Piaget argued that all children pass through several stages. We will briefly mention three of the main assumptions of his stage theory: First, there are large enough changes in the course of cognitive development to permit the identification of separate processing stages. Second, all children pass through the same *sequence* of stages in the same order (although at different rates). Third, the cognitive operations and structures defining each stage form an integrated whole.

What stages of cognitive development did Piaget identify? He argued that there are four major stages (see figure on the following page):

1. *The sensori-motor stage (birth to 2 years)*: This stage is basically intelligence in action, with the child's main focus on the impact of his/her behavior on the environment.
2. *The pre-operational stage (2–7 years)*: Thinking in this stage is dominated by perception.
3. *The concrete operations stage (7–11 years)*: Logical reasoning is only applied to objects that are real or can be seen.
4. *The formal operations stage (11 upwards)*: Logical reasoning is applied to potential events and abstract ideas as well as concrete objects.



An example of the dominance in play of assimilation over accommodation—pretending hairbrushes are microphones.

Key Terms

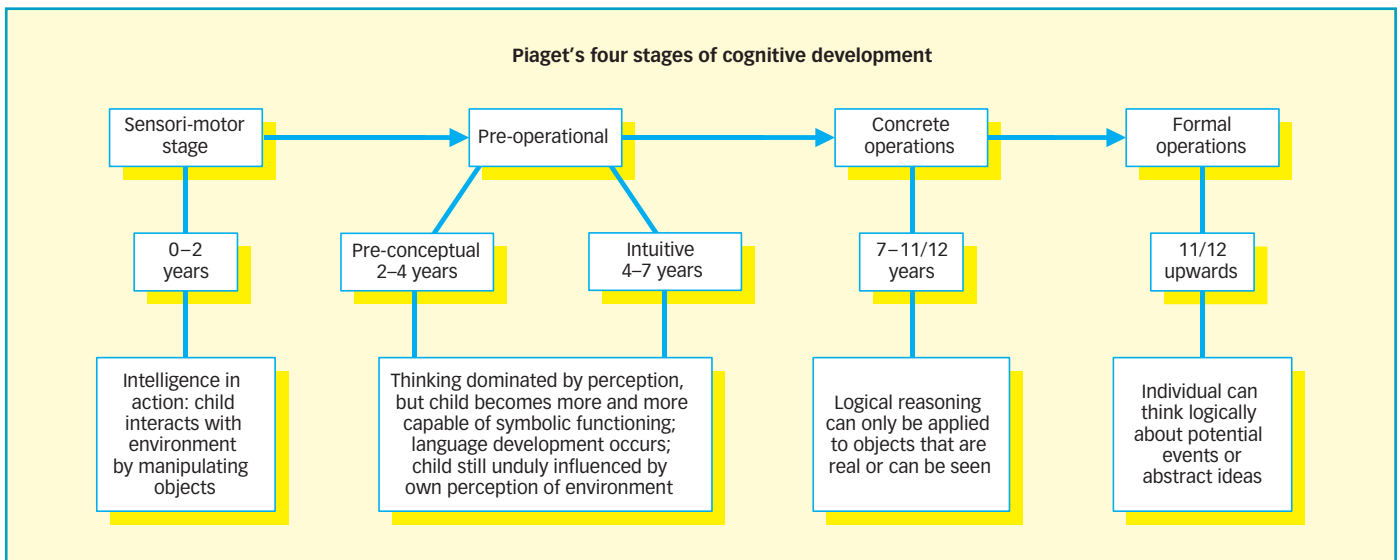
Schema:
in Piaget's theory, organized knowledge used to guide action.

Equilibration:
using the processes of accommodation and assimilation to produce a state of equilibrium or balance.

Clinical method:
an informal question-based approach used by Piaget to assess children's understanding of problems.

How did Piaget test his stage theory? He was rather skeptical about the value of the typical experimental approach based on strict use of the experimental method. Instead, Piaget preferred a less structured and formal approach. He used the **clinical method**, in which children are questioned informally to reveal the nature of their understanding of problems.

There is one other crucial point. Piaget focused very much on describing the strengths and limitations of



children's thinking at different developmental stages. However, he failed to explain precisely *why* and *how* children's thinking develops, and he attached remarkably little importance to the role of learning. Piaget's approach can be defended by arguing that we need to know *what* needs to be explained before trying to provide an adequate explanation.

FINDINGS

According to Piaget, the key achievement of the sensori-motor stage is **object permanence**. This involves the child being aware that objects continue to exist when no longer in view. Initially, the infant has no awareness at all of object permanence; it is literally a case of "out of sight, out of mind." Towards the end of the first year, however, the infant seems to show partial object permanence. What happens is that the experimenter hides an object at location A, which the infant then finds. After that, the same object is hidden at location B, but the infant mistakenly searches at location A. This finding (**perseverative search**) allegedly occurs because the infant doesn't regard objects as existing independently of his/her own behavior.

Piaget's assumption that infants show perseverative search because they don't remember where the object has been hidden is rejected by some researchers. In one study (Baillargeon & Graber, 1988), 8-month-old infants saw a toy being hidden behind one of two screens. Fifteen seconds later they saw a hand lift the toy out either from the place in which it had been hidden or from behind the other screen. The infants were only surprised when the toy was lifted from behind the "wrong" screen. This indicates that they *did* remember where it had been put (see figure on the following page).

Another major achievement of the sensori-motor stage is imitation, which allows the infant to add considerably to its range of actions. Towards the end of the sensori-motor stage, the infant shows evidence of **deferred imitation**, which is the ability to imitate behavior seen before. In a study by Hayne, Boniface, and Barr (2000), the experimenter demonstrated a



Object permanence is the term used to describe the awareness that objects continue to exist even when they are no longer visible.

Key Terms

Object permanence:

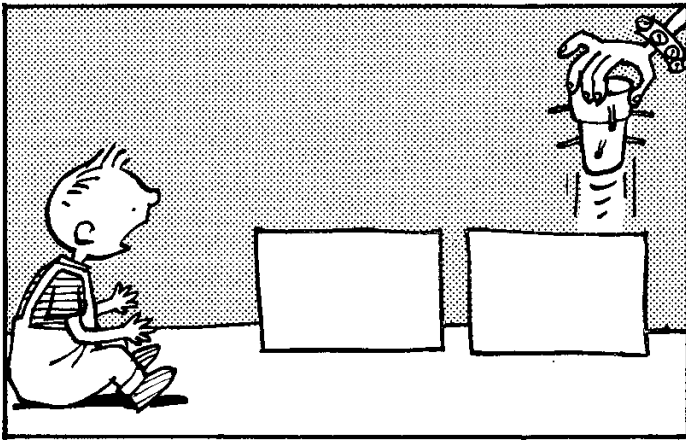
an awareness that objects continue to exist when they can no longer be seen.

Perseverative search:

mistakenly searching for an object in the place in which it was previously found rather than the place in which it is currently hidden.

Deferred imitation:

in Piaget's theory, the ability to imitate behavior observed at an earlier time.



Baillargeon and Graber found that 8-month-old infants were surprised when a cup they had seen being put behind the left-hand screen was then retrieved from behind the right-hand screen.

novel action using an unfamiliar object, and infants then tried to reproduce that action after an interval of time. Infants aged 6, 12, and 18 months all showed deferred imitation when the same stimulus object was used for the demonstration and the subsequent test.

We turn now to the pre-operational stage. According to Piaget, pre-operational children are unduly influenced by their immediate perception of the environment. They pay attention to only one aspect of the total situation (this is called **centration** by Piaget), which can make them prone to error. For example, they don't show **conservation**, which is the understanding that certain aspects of a visual display don't vary in spite of changes in perceptual aspects.

The thinking of pre-operational children is also characterized by egocentrism. **Egocentrism** is the tendency

Key Study

Piaget: Conservation

In his classic studies on conservation of quantity, Piaget presented children with two glasses of the same size and shape containing the same quantity of liquid. When the child agrees there is the same quantity of liquid in both glasses, the liquid from one of them is poured into a glass that is taller and thinner. The child is then asked if the two glasses (the original one and the new one) contain the same amount to drink, or if one contains more. Pre-operational children fail to show conservation. They argue mistakenly either that there is more liquid in the new container ("because it's higher") or that there is more liquid in the original container ("because it's wider"). In either case, the child centers or focuses on only *one* dimension (height or width).



Piaget claimed that pre-operational children fail on conservation tasks partly because of centration. However, they also fail because they lack the cognitive operation of reversibility. **Reversibility** involves the ability to undo (or reverse mentally) some operation. Thus, reversibility allows the realization that the effect of pouring liquid from one container into another could be negated by simply pouring it back into its original container.

Discussion points

1. In what real-life situations might the ability to conserve be important?
2. What are the limitations of Piaget's research?

Key Terms

Centration: attending to only one aspect of a situation.

Conservation: the principle that quantities remain constant under various changes to their appearance.

Egocentrism: the assumption that one's way of thinking is the only possibility.

Reversibility: the ability to undo, or reverse mentally, an action or operation.

KEY STUDY EVALUATION

Piaget used the conservation of liquid task to show that pre-operational children lack the internalized cognitive operations of reversibility and decentration. However, it might be interesting to try the same experiment with children from a non-Western environment, such as the bush people of the African Kalahari desert, who are not likely to be familiar with glass beakers filled with water. Would they show conservation or not? Would a lack of conservation necessarily mean that these children could not decenter?

to assume that one's way of thinking about things is the only possible way. It often involves a lack of differentiation between the self and the world, so the child can't distinguish clearly between psychological and physical events. This produces:

- *Realism*: The tendency to regard psychological events as having a physical existence.
- *Animism*: The tendency to endow physical objects and events with psychological qualities.
- *Artificialism*: The tendency to consider that physical objects and events were created by people.

An example of artificialism concerns my elder daughter Fleur at the age of 3. We were on Wimbledon Common in London, and I told her the sun would come out when I had counted to 10. When it did so, she was very confident that Daddy could control the sun, and often begged me to make the sun appear on overcast days!

Children in the pre-operational stage often fail to show conservation. However, they possess more knowledge of conservation than seemed to be the case in Piaget's studies. For example, Wheldall and Poborca (1980) argued that children often fail on conservation tasks because they don't understand the question. Accordingly, they devised a nonverbal version of the liquid conservation task. Only 28% of their 6- and 7-year-old children showed conservation with the standard verbal version compared to 50% with the nonverbal version.

McGarrigle and Donaldson (1974) argued that children presented with a conservation task typically assume that the experimenter deliberately *intends* to change the amount of liquid or other substance. This assumption biases them against showing conservation. It follows that children would show more evidence of conservation if the change made on a conservation task appeared to be accidental. This prediction was supported by McGarrigle and Donaldson—see the Key Study below.

Key Study**McGarrigle and Donaldson (1974): The "naughty teddy"**

McGarrigle and Donaldson showed that changing the way in which a conservation task is presented can make a large difference. They presented 6-year-old children with two rows of counters. All the children agreed there were equal numbers of counters in each row. In one condition, the experimenter deliberately messed up one of the rows. Number conservation was shown on 40% of the trials. This finding suggests that most of the children lacked the underlying competence necessary to show number conservation. However, the findings were rather different in a second condition, in which a "naughty teddy bear" messed up one of the rows in what looked like an accidental way. In this condition, number conservation was obtained on 70% of the trials.

Why did McGarrigle and Donaldson (1974) find such a large difference between the two conditions? The high level of performance in the “naughty teddy” conditions presumably occurred because most of the children had some understanding of number conservation. In the other condition, the fact that the experimenter deliberately altered the situation may have led the children to assume that the experimenter *intended* to change the number of counters in one of the rows. Whether or not that is correct, the fact remains that performance in that condition failed to reflect the underlying level of competence. However, it should be noted that Eames, Shorrocks, and Tomlinson (1990) failed to replicate McGarrigle and Donaldson’s findings.

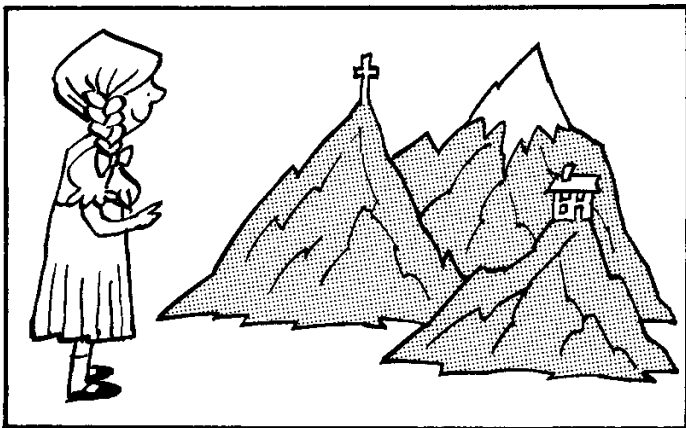
Discussion points

1. Why do you think that McGarrigle and Donaldson found such a large difference between their two conditions?
2. What problems for Piaget’s theory arose from his failure to distinguish carefully between performance and competence (i.e., the child’s potential)?

KEY STUDY EVALUATION

Recent research suggests that McGarrigle and Donaldson may have been mistaken. It is possible that the children were so absorbed in the “naughty teddy” routine that they didn’t actually notice the transformation and that is why, with naughty teddy, they said the display hadn’t changed. To test this possibility, Moore and Frye (1986) arranged for naughty teddy to actually add a counter (or take one away). Children said no change had taken place, which suggests that they were simply not attending to the display.

However, other evidence provides support for McGarrigle and Donaldson’s (1974) findings. Light, Buckingham, and Robbins (1979) also found that the existence of number conservation in children depended very much on whether or not the changes introduced by the experimenter seemed deliberate. It is likely that the percentage of children showing conservation is influenced by a number of factors (e.g., whether the experimenter changes the situation deliberately; whether the children attend to the changes).



A drawing of the model used in Piaget’s three mountains task. Children were shown the model from one angle, then shown photographs of the model from other viewpoints, and asked to choose which view someone standing at one of the other labeled points would see. Pre-operational children usually selected the view from the point at which they themselves had seen the model.

Piaget studied egocentrism in pre-operational children by using the three mountains task (see figure on the left). Children aged between $3\frac{1}{2}$ and 5 looked at a model of mountains, and decided which out of various pictures showed the view that would be seen by someone looking at the display from a different angle. Children younger than 8 nearly always selected the photograph of the scene as they themselves saw it. According to Piaget, this error occurred because young children can’t escape from an egocentric perspective.

Hughes (1975) argued that children performed poorly on the three mountains task because it doesn’t relate to their experience. He tested this argument by using a piece of apparatus in which two walls intersected at right angles to form what looked like a plus sign. A boy doll and a policeman doll were put into the apparatus, and the child was asked whether the policeman doll could see the boy doll. After that, the child was told to hide the boy so the policeman couldn’t see him. Nearly all the children could do this, and

any errors were corrected. Finally, a second policeman was used, and the children were told to hide the boy doll so that neither of the policemen could see him. According to Piaget, the children should have hidden the boy doll so they themselves couldn't see it, and so should have failed the task.

What happened in Hughes' (1975) experiment? Ninety percent of the children performed the task successfully, which is a much higher figure than reported by Piaget with the three mountains task. The main reason for the difference is probably that the task used by Hughes was much more meaningful and interesting. However, the task was also simpler than the one used by Piaget.

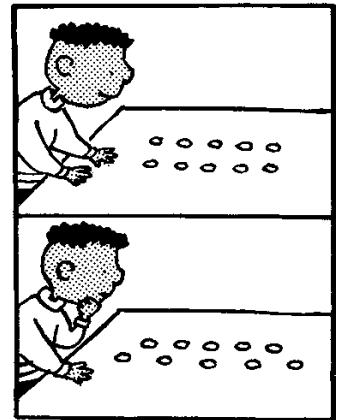
When children move from the pre-operational to the concrete operational stage they show an increasing independence of thought from perception. Part of this change involves an understanding of conservation and of reversibility. It also involves the development of various cognitive operations of a logical or mathematical nature, including the actions implied by mathematical symbols (e.g., $+$, $-$, \div , \times , $>$, $<$, and $=$). These cognitive operations are usually combined or organized into a system or structure. For example, the operation "greater than" can't really be considered independently of the operation "less than." You have only grasped the full meaning of "A is greater than B" if you realize that statement means that "B is less than A." Piaget used the term **grouping** to refer to a set of logically related operations.

We will consider two of the achievements that Piaget claimed were associated with the concrete operational stage. First, there is **transitivity**, which allows three or more elements to be placed in the correct order. For example, if Mark is taller than Peter, and Peter is taller than Robert, then it follows from the notion of transitivity that Mark must be taller than Robert. Piaget found that children in the concrete operational stage could solve such problems. However, Pears and Bryant (1990) found evidence of transitivity in 4-year-olds. The children were shown several small towers, each consisting of two colored bricks (e.g., a red brick on a yellow brick; a yellow brick on a green brick; and a green brick on a blue brick). They were told to construct a large tower in which the order of the bricks corresponded to that in the small towers. Thus, the large tower would have red at the top, followed in sequence by yellow, green, and blue. Before building the large tower, the children showed a reasonable ability to work out the correct order of the bricks, thus making transitive inferences several years earlier than assumed by Piaget.

Second, children in the concrete operational stage achieve various forms of conservation. As predicted by Piaget, more complex forms of conservation (e.g., conservation of volume) are achieved some years after easier forms of conservation (e.g., conservation of number; see the figure above right) (Tomlinson-Keasey et al., 1979). Why is conservation of volume only achieved at about the age of 11 or 12? According to Piaget, it involves taking account of the operations involved in the conservation of liquids *and* of mass. We can see this by considering how it is assessed. For example, two identical balls of clay are placed in two identical transparent containers filled to the same level (see figure on the right). One ball of clay is then molded into a new shape, with conservation being shown if the child realizes that this will not change the amount of water it displaces.

Piaget de-emphasized the role of specific learning experiences and cultural factors in determining performance on conservation tasks. For example, Price-Williams, Gordon, and Ramirez (1969) found that the children of Mexican potters had slow development of conservation of volume using beakers, but fast development when a ball of clay was stretched into an oblong shape.

According to Piaget, children and adults in the formal operations stage have the ability to think in terms of many possible states of the world. This allows us to



One of the tasks used to test conservation of number. Children are asked if there are the same number of beads in the two rows before and after they are rearranged.

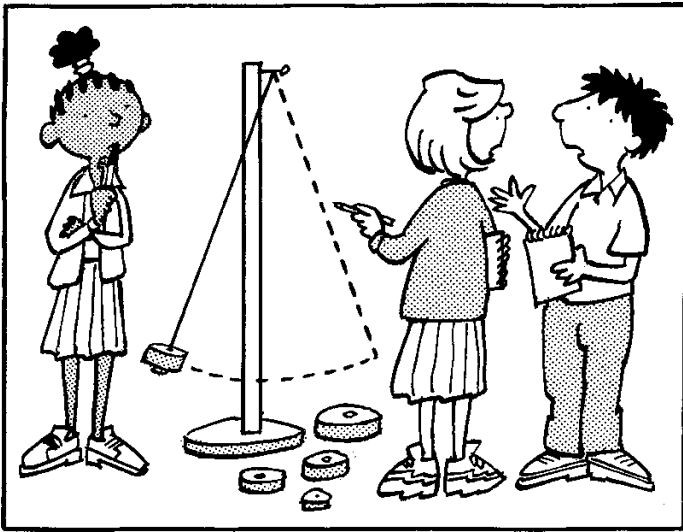


This apparatus tests conservation of volume. Children are asked if the liquids will be at the same level again when the new shape of clay is put back into the glass. Conservation of volume is not usually attained until about the age of 11 or 12.

Key Terms

Grouping:
in Piaget's theory, a set of logically related operations.

Transitivity:
this involves the ability to place at least three entities in the correct order.



Children were asked to work out what would affect the frequency of the swings of the pendulum. They were asked to consider changing the weights on the pendulum, the length of the string, how hard they pushed it, and which direction it was pushed in.

think in abstract terms and so escape from concrete reality. Piaget studied formal operational thought by presenting children with a set of weights and a string that could be lengthened or shortened. The task was to work out what determines the frequency of the swings of a pendulum formed by suspending a weight on a string from a pole (see figure on the left). *Only* the length of the string is important, but younger children often claim mistakenly that the strength of the push given to the pendulum is the main factor.

Piaget greatly overestimated the intellectual prowess of most adolescents. Bradmetz (1999) assessed formal operational thinking in 62 15-year-olds using various Piagetian tasks. Only *one* participant showed substantial evidence of formal operational thought. Bradmetz also found that overall performance on the tests of formal thought correlated +.61 with general intelligence. Thus, the cognitive abilities associated with formal thought resemble those assessed by traditional intelligence tests.

Evaluation

- + Piaget's theory was a very ambitious attempt to describe children's cognitive development from being irrational and illogical to being rational and logical.
- + The notions that children learn certain basic operations (e.g., reversibility), and that these operations then allow them to solve numerous problems, are valuable ones.
- Stage theories such as Piaget's *exaggerate* the differences between stages and *minimize* the differences within stages. For example, children in the concrete operations stage show conservation of quantity for familiar materials before they show it for unfamiliar materials (Durkin, 1995). Thus, successful performance depends on *specific* learning experiences as well as on *general* cognitive operations emphasized by Piaget.
- Piaget *underestimated* the cognitive abilities of young children, but *overestimated* those of adolescents and adults. However, as Lourenço and Machado (1996) pointed out, Piaget was interested in children's deep understanding of problems, and simply producing correct answers to problems at a young age doesn't necessarily require deep understanding.
- Piaget provided a detailed description of the major changes in cognitive development but didn't adequately explain the processes involved. In the words of Siegler and Munakata (1993), there appears to be a "miraculous transition" from one developmental stage to the next.
- Piaget de-emphasized the role of social factors in cognitive development. For example, children's cognitive development benefits from social interactions with adults and other children.
- Piaget virtually ignored individual differences. He admitted, "I'm not really interested in the individual. I'm interested in what is general in the development of intelligence and knowledge" (Bringuier, 1980, p. 86).

The French connection

Larivee, Normandeau, and Parent's (2000) overview of French-speaking researchers' work in Switzerland and France shows how Piaget's ideas have developed. The normal variations between children—what psychologists call individual differences—are seen as important areas for further research. Differing sensitivities to situations and differing preferences for strategies or ways of problem solving could lead to different developmental pathways. Nowadays one pathway of developmental stages is not thought to be applicable to the child population of even one or two cultures.

VYGOTSKY'S THEORY

Textbook writers typically argue that Vygotsky's approach to cognitive development was radically different from that of Piaget, and there are some major differences. However, we should not exaggerate the scale of these differences. Smith (1996) assembled 10 quotations

from Vygotsky and 10 from Piaget, and asked various experts to decide who had written each one. On average, these experts performed at little better than chance level!

In some ways, the approaches of Piaget and Vygotsky complement each other. As Shayer (2003) pointed out, Piaget was mainly concerned with determining children's thinking abilities at different stages of cognitive development. In contrast, Vygotsky was more interested in the *dynamics* of change, namely, the factors responsible for cognitive development.

One of the difficulties in discussing Vygotsky's ideas is that he kept changing them. However, a constant feature of his approach (and a real difference from Piaget) was his emphasis on the importance of social factors in influencing cognitive development. According to Vygotsky (1930/1981, p. 163), "Social relations or relations among people genetically [developmentally] underlie all higher functions and their relationships." More specifically, Vygotsky (p. 163) argued as follows: "Any function in the child's cultural development appears twice, or on two planes. First, it appears on the social plane, and then on the psychological plane." Within this approach, the child is an apprentice who learns directly from social interaction and from communication with older children having knowledge he/she lacks (Durkin, 1995).

We can contrast Vygotsky's approach with that of Piaget, who argued that children acquire knowledge through self-discovery. However, there are some important similarities. Vygotsky and Piaget both agreed that activity forms the basis for learning and for the development of thinking. In addition, they both argued that learning is most effective when the information presented to children is closely related to their current knowledge and abilities.

Vygotsky argued that there are four stages in children's formation of concepts. He identified these four stages on the basis of a study (one of the very few he carried out) in which children were presented with wood blocks having labels consisting of nonsense syllables. Each nonsense syllable consistently referred to blocks having certain characteristics (e.g., circular and thin blocks). The children had the concept-formation task of working out the meaning of each nonsense syllable. Vygotsky's four stages were as follows:

1. *Vague syncretic stage*: Children fail to use systematic strategies and show little understanding of the concepts.
2. *Complex stage*: Nonrandom strategies are used, but are not successful in finding the main features of each concept.
3. *Potential concept stage*: Systematic strategies are used, but are limited to focusing on one feature at a time (e.g., shape).
4. *Mature concept stage*: Systematic strategies relating to more than one feature at a time are used, and lead to successful concept formation.

It is interesting that Vygotsky's findings resembled those of Piaget with very different tasks. Vygotsky found that children had problems with concept formation because they focused on only one salient or obvious feature of stimuli. This is very similar to Piaget's discovery that pre-operational children fail on conservation tasks because they attend to only one aspect of the situation.

ZONE OF PROXIMAL DEVELOPMENT

Vygotsky emphasized the notion of the **zone of proximal development**. This was defined by Vygotsky (1978, p. 86) as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers." In other words, the zone of proximal development involves "problem solving that is beyond one's unassisted efforts but which can be achieved with assistance" (Granott, 2005, p. 141).

Two aspects of the zone of proximal development are of particular importance. First, children apparently lacking certain skills when tested on their own may perform more effectively in the social context provided by someone having the necessary knowledge. Second, when a given child's level of understanding is moderately challenged, he/she is most likely to acquire new knowledge rapidly and without a sense of failure. Vygotsky assumed that children differ in the size of the zone of proximal development. Those with

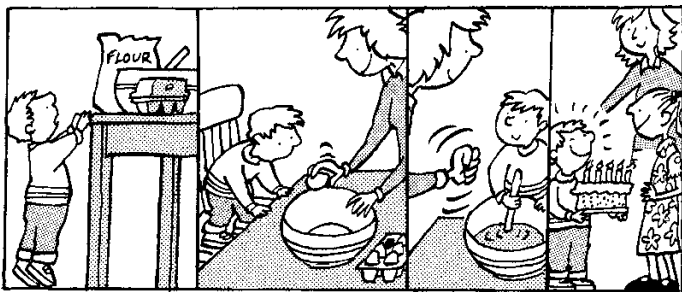


The Russian psychologist Lev Semeonovich Vygotsky, 1896–1934.

Key Term

Zone of proximal development:

in Vygotsky's theory, the gap between the child's current problem-solving ability and his/her potential ability.



Left to his own devices, could this boy make his sister a birthday cake? His mother uses scaffolding to create a situation within which he can begin to move into a zone of proximal development.

Case Study: Weaving

Evidence that the zone of proximal development and scaffolding are used effectively in cultures other than typical Western ones was reported by Greenfield and Lave (1982) in a study on the Zinacanteco Mexicans. Young girls who wanted to learn weaving skills started by spending almost half their time simply watching expert women weavers. After that, the girls were closely supervised by the skilled weavers as they acquired the necessary skills. The skilled weavers were generally successful at structuring the assistance they provided so that the girls remained within the zone of proximal development. Finally, the girls developed sufficient skills so that they could take responsibility for their own weaving.

larger zones of proximal development derive more benefit from instructions than those with smaller zones.

Wood, Bruner, and Ross (1976) extended the notion of a zone of proximal development. They introduced the concept of **scaffolding**. This refers to the context provided by knowledgeable people such as adults to help children to develop their cognitive skills. Effective scaffolding means the child doesn't need to climb too far at any point. Another important aspect of scaffolding is that there is a gradual withdrawal of support as the child's knowledge and confidence increase. Scaffolding and the zone of proximal development are closely related, but scaffolding focuses more on the strategies used by the adult and less on changes in the child.

Granott (2005) identified four major components of efficient scaffolding:

1. It is temporary and is dismantled when the child makes sufficient cognitive progress.
2. The person providing scaffolding increases his/her input if the child reduces its input, and decreases his/her input if the child increases its input.
3. The input of the person providing scaffolding is at a higher level than that of the child; this supports and stimulates the child.
4. Both partners should find their interactions on the task pleasant and rewarding.

LANGUAGE

Vygotsky attached great importance to the development of language through three stages. Language and thought are

unrelated during the first stage of development. During the second stage, language and thought develop in parallel and continue to have very little impact on each other. During the third stage, children use the speech of others and talking to themselves to assist in their thinking and problem solving. By the age of 7, egocentric speech (i.e., speaking without heeding anyone else present) gives way to inner speech.

According to Vygotsky, language becomes increasingly central to cognitive development over the years. Berk (1994, p. 62) described some of the processes that Vygotsky had in mind: "When a child discusses a challenging task with a mentor [person providing guidance], that individual offers spoken directions and strategies. The child incorporates the language of those dialogs into his or her private speech and then uses it to guide independent efforts."

Vygotsky's views on the role of language in cognitive development were very different from those of Piaget. At the risk of oversimplification, Vygotsky argued that cognitive development depends crucially on language development and use. In contrast, Piaget argued that cognitive development typically precedes (and is little affected by) language development.

Findings

Moss (1992) reviewed studies on the scaffolding provided by mothers during the preschool period. There were three main aspects to the mothers' scaffolding strategies. First, the mother instructed her child in new skills the child couldn't use on its own. Second, the mother encouraged her child to maintain useful problem-solving tactics it had shown spontaneously. Third, the mother persuaded the child to discard immature and inappropriate forms of behavior. In general, scaffolding emerged as an effective technique for promoting learning in preschool children.

Conner, Knight, and Cross (1997) studied the effects of scaffolding on 2-year-olds performing various problem-solving and language tasks. Most previous studies had focused only on mothers' scaffolding, but Conner et al. also considered fathers'

Key Term

Scaffolding: the context provided by an adult or other knowledgeable person helping children to develop their cognitive skills.

scaffolding. Mothers and fathers were equally good at scaffolding, and the quality of scaffolding predicted the children's performance on various tasks during the teaching sessions. The beneficial effects of good scaffolding were still evident at a follow-up session. Children who had originally received better scaffolding continued to perform better than those who had received poor scaffolding.

Vygotsky's notion that inner speech is of value in thinking has received support. Behrend, Harris, and Cartwright (1992) used whispering and observable lip movements as measures of inner speech. Children using the most inner speech performed difficult tasks better than children making little use of inner speech. Berk (1994) found that 6-year-olds spent 60% of the time talking to themselves while solving problems in mathematics. Those whose speech contained numerous comments about what needed to be done on the current problem did better at mathematics over the following year. This confirmed Vygotsky's view that self-guiding speech makes it easier for children to direct their actions.

Vygotsky argued that private speech becomes more internal as children's level of performance improves. Berk (1994) discussed a study in which 4- and 5-year-old children made Lego models in each of three sessions. As predicted by Vygotsky, the children's speech became increasingly internalized from session to session as their model-making performance improved. Thus, as Vygotsky assumed, private speech is of most value to children confronted by novel tasks they don't fully understand.

Vygotsky assumed that children do not produce egocentric or private audible speech after the age of 7. However, Girbau (2002) found evidence against that assumption in a study in which children of 8 and 10 played in pairs with a Lego construction set. Egocentric speech was found in both age groups, and was somewhat more frequent in the older group.

Evaluation

- + Piaget underestimated the importance of the social environment in cognitive development, and Vygotsky deserves credit for recognizing its key role.
- + Vygotsky's ideas have led to the introduction of several useful teaching techniques (e.g., scaffolding; see later in chapter).
- + As Vygotsky predicted, inner speech helps the problem-solving activities of young children.
- Many of Vygotsky's ideas were rather speculative, and he carried out very little research.
- Vygotsky exaggerated the importance of the social environment. Children's rate of cognitive development is determined more by internal factors (e.g., level of motivation; interest in learning) than he believed.
- Vygotsky didn't specify clearly what kinds of social interaction are most beneficial for learning (e.g., general encouragement vs. specific instructions).
- Social interactions are not always beneficial to learning. As parents discover, interactions with their children can lead to confrontations and stubbornness rather than to enlightenment.

CONTEMPORARY APPROACHES

Piaget and Vygotsky discussed in detail the changes in thinking and behavior shown by children during development. However, there were serious limitations in their approaches. Piaget was surprisingly uninterested in the role played by learning in cognitive development, and devoted little effort to understanding how children move from stage to stage. Vygotsky was interested in the processes responsible for children's cognitive development, but carried out practically no research. These gaps in our knowledge have recently been addressed by several theorists, and we will shortly consider the contributions of one of them (Robert Siegler).

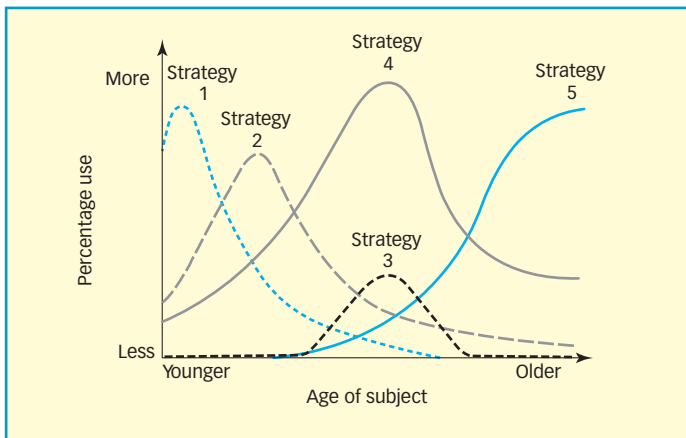
Before discussing Siegler's theory, we will consider the complex issue of how to study children's learning processes. Piaget and his followers tested children of different ages at a given point in time, an approach that didn't allow them to observe cognitive changes as they occurred. The obvious alternative is to use a longitudinal approach in which a given group of children is studied over a relatively long period of time. This is superior to Piaget's cross-sectional approach. However, it is hampered by the fact that major changes in children's development often take place over years rather than over weeks or months.

The **microgenetic method** has proved an especially valuable longitudinal approach. Use of this method involves carrying out short-term longitudinal studies in which intensive training is provided to accelerate changes in thinking (Miller & Coyle, 1999). The other key feature of the microgenetic method is that a substantial amount of data is obtained from each child in order to maximize the chances of understanding the learning process. The microgenetic method has a potential limitation: "Although [it] reveals how behavior *can* change, it is less clear whether behavior typically *does* change in this way in the natural environment" (Miller & Coyle, 1999, p. 212). This issue will be discussed in the context of Siegler's research, much of which involves use of the microgenetic method.

OVERLAPPING WAVES THEORY

Siegler (e.g., 1998, 2004, 2005) proposed an influential approach to cognitive development known as overlapping waves theory. According to this theory, children tackle most problems by using strategies that are goal-directed approaches varying from trial to trial. The key assumption is that children at any given time typically have various strategies or ways of thinking about a problem, and they exhibit considerable *variability* in strategy use (see the figure below). Common sense might suggest that any given child will adopt a particular strategy to solve problems of a given type. In contrast, it is assumed within overlapping waves theory that children often change their strategy from problem to problem.

How do children decide which strategy to use on a given problem? According to the theory, children choose adaptively among strategies, taking account of the speed and accuracy of problem solving with each strategy. Over time, there are several reasons why strategies change. First, children make increasingly adaptive choices among strategies based in part on the feedback they receive. Second, they execute each strategy more efficiently. Third, they make increasing use of the more advanced strategies among those at their disposal. Fourth, they acquire new strategies, sometimes by combining elements of pre-existing ones.



Siegler's overlapping waves model.

FINDINGS

Children are remarkably variable in their strategy use across a wide range of problems (see Siegler, 2004, for a review). This is so even when children are presented with the same problem on two occasions close in time. For example, Siegler and Shrager (1984) with preschoolers and Siegler and McGilly (1989) with older children found that about one-third changed strategies in such circumstances. Most of these changes in strategy were *not* a result of learning, because almost half of them involved a shift from a more advanced to a less advanced strategy.

One reason for children's variability of strategy use is that they often fail to use the optimal strategy after having discovered it. For example, Siegler and Stern (1998) gave 7-year-olds problems of the form $A + B - B = ?$ (e.g., $18 + 24 - 24 = ?$). The best strategy with such problems is simply to say the first number. The children took an average of only seven trials to discover this strategy, but they didn't use it consistently even 100 trials later. Part of the reason for this was that almost 90% of the children used the best strategy on one or more trials before being able to report verbally that they were using it.

Key Term

Microgenetic method: an approach to studying children's changes in cognitive strategies by means of short-term longitudinal studies.

It could be argued that variable strategy use is of relatively little interest because it merely reflects a lack of intellectual ability in children who often change strategies. That is *not* the case. Children who initially use several strategies with a given type of problem learn *faster* than those who initially use only one or two (see Siegler, 2005, for a review). Why is this? According to Siegler and Araya (2005), new strategies are often constructed from components of two existing strategies. This construction process is most likely to occur when both existing strategies have been used recently.

Detailed information about the processes involved in the adoption of a new strategy was reported by Siegler and Jenkins (1989). Children aged 4 and 5 were given addition problems (e.g., $3 + 8$). One of the most effective strategies used by young children with such problems is the count-on strategy. This involves starting with the larger number and counting on from that point. Thus, in our example, children would start with 8, and count 9, 10, 11.

Siegler and Jenkins (1989) gave children 30 sessions devoted to solving addition problems. None of the children used the count-on strategy initially, but nearly all of them discovered it during the course of the study. Most of the children took much longer than usual on the problem immediately before their first use of the strategy. This suggests they were thinking carefully about the best strategy to use.

As predicted by overlapping waves theory, the children often used other strategies after discovering the count-on strategy. However, they showed a much more consistent use of the count-on strategy after being presented with some challenging problems in the eighth week of the study. These problems (e.g., $26 + 2 = ?$) were easy to solve using the count-on strategy but very difficult to solve with other strategies (e.g., counting all the way from 1).

There was a final interesting finding from the study by Siegler and Jenkins (1989). Several children who used the count-on strategy had little conscious understanding of it. Indeed, some of them denied using the strategy at all in spite of very clear videotape evidence! Thus, implicit knowledge (knowledge without conscious awareness) can guide children's choice of strategy.

We return now to the issue of whether the rapid strategy changes found with use of the microgenetic method are basically similar to those shown by children in the natural environment. Relevant evidence was reported by Siegler and Svetina (2002). Children were given matrix problems in which they selected an object to complete a matrix or visual display. Changes in performance occurring naturally were assessed by giving matrix problems to children aged 6 and 7 and looking for differences between them. Other 6-year-olds were given systematic training in solving matrix problems using the microgenetic method. There were striking similarities between the patterns of change produced by the microgenetic method and by the natural environment. More specifically, there was matching on 10 out of 11 measures of change when the two conditions were compared.

Evaluation

- + It is very useful to study changes in children's strategies using short-term longitudinal studies involving intensive training.
- + The microgenetic method produces changes in cognitive strategies faster than happens in more naturalistic conditions, but the changes in both cases are similar (Siegler & Svetina, 2002).
- + Children often have several strategies available to them, and cognitive development depends in part on competition among these strategies.
- It remains relatively unclear how children discover new strategies.
- “The theory seems most applicable to domains in which children use clearly-defined strategies; its applicability to areas in which strategies are less well defined remains to be demonstrated” (Siegler, 1998, p. 97).

EDUCATIONAL IMPLICATIONS

The theoretical views of Piaget and Vygotsky have been very influential in education. We start by considering how their influence has been felt. After that, we discuss the Cognitive Acceleration through Science Education (CASE) program. This program represents an ambitious attempt to combine the insights of Piaget and Vygotsky to promote more effective teaching of science.

PIAGET'S APPROACH

Piaget himself didn't focus very much on the usefulness of his theory for educational practice. However, many of his ideas are of clear relevance to education, and many educationalists have applied them to teaching. According to Piaget, children learn best when engaged in a process of **self-discovery** involving initiative and an active approach. Teachers can foster self-discovery in their students by creating a state of disequilibrium in which the child's existing schemas or cognitive structures are shown to be inadequate. Disequilibrium can be produced by asking children difficult questions and encouraging them to ask questions.

Piaget's preferred approach can be contrasted with the traditional approach in which the teacher provides relatively passive children with knowledge. Piaget argued that this approach (**tutorial training**) is much less effective than self-discovery. In his own words, "Every time we teach a child something, we prevent him from discovering it on his own."

Brainerd (2003) reviewed the literature on self-discovery, and found it was generally less effective than tutorial training. He discussed five studies in which self-discovery of conservation concepts was compared to guided discovery, in which teachers directed students' attention to relevant features of the conservation task. Guided discovery was clearly more effective than self-discovery in all five studies.

Piaget argued that cognitive development depends very much on children acquiring a range of schemas or cognitive structures mostly based on mathematical or logical principles. It follows that it should be useful for children to study mathematics and logic as well as science subjects that provide illustrations of those principles at work. Of crucial importance is the notion that the learning material shouldn't be too complex or too far removed from the child's existing schemas.

The major weakness of Piaget's position is that the cognitive structures he emphasized are of limited relevance for many kinds of learning. For example, concrete and formal operations are of little assistance in the learning of foreign languages or history. Thus, Piaget's approach applies only to a small number of school subjects.

VYGOTSKY'S APPROACH

According to Vygotsky (1986, p. 188), "The only good kind of instruction is that which marches ahead of development and leads it; it must be aimed not so much at the ripe as the ripening functions." This can best be achieved when children's learning efforts are guided and encouraged by someone more knowledgeable. Thus, children can be regarded as apprentices who are taught the necessary skills by those already possessing them. This is known as scaffolding. As we saw earlier in the chapter, it can enhance children's learning. Effective teachers reduce their control over the learning process when children are performing successfully, and increase their control when children start making errors.

According to Vygotsky, those involved in teaching children should focus on the child's zone of proximal development. They should concentrate on knowledge only slightly beyond the child's current competence. It can be argued that the ideal tutors are children slightly older and more advanced than those being taught. Such tutors have useful knowledge to communicate but they still remember the limitations in their own knowledge and understanding when they were 1 or 2 years younger.

The approach described above is known as **peer tutoring**, and it is generally effective. For example, Ellis and Gauvain (1992) compared 7-year-old Navaho and Euro-American children performing a maze game. They were tutored by either one or two 9-year-old tutors working together. The children from both cultures benefited more from the paired tutors than from the individual tutors, and the benefit was the same in both cultures. Van Keer (2004) studied the effects of peer tutoring on the reading comprehension ability of

Key Terms

Self-discovery:

an active approach to learning in which the child is encouraged to use his/her initiative.

Tutorial training:

a traditional approach in which the teacher imparts knowledge to fairly passive students.

Peer tutoring:

teaching of one child by another, with the child doing the teaching generally being slightly older than the one being taught.

11-year-olds. Peer tutoring was effective when the tutor was a child who was older than the learner but not when the tutor was the same age as the learner. In the latter case, the knowledge of the peer tutor was insufficient to promote the learner's reading performance.

In spite of many examples of enhanced learning produced by scaffolding (including peer tutoring), these techniques have four main limitations. First, the learner may become uninvolved in the learning process if the tutor has too much status. Second, scaffolding seems better suited to some kinds of tasks (e.g., construction tasks) than to others. For example, Howe, Tolmie, and Rodgers (1992) found that peer tutoring was of very little benefit on a task concerned with understanding motion down an incline. Third, the main focus of scaffolding is on the contribution made by the tutor to the child's understanding. In fact, however, the success (or otherwise) of scaffolding often depends crucially on the responsiveness of the tutor to the child's thoughts and actions. Fourth, scaffolding can be hard to control and to analyze because it involves complex interactions between tutor and child (Granott, 2005).



Peer tutoring: two brothers teach their younger sister to read.

CASE PROGRAM

At a superficial level, it may appear that the approach to education based on Vygotsky's ideas is radically different from that of Piaget. As DeVries (2000) pointed out, Vygotsky seemed to emphasize factors *external* to the child (e.g., tutors; teachers) in promoting learning. In contrast, Piaget emphasized *internal* factors (e.g., adaptation), with the child in charge of the learning process. In fact, however, Vygotsky and Piaget were both fully aware of the importance of both external and internal factors. For example, Vygotsky's notion that tutors should teach within the zone of proximal development is a clear recognition of the importance of internal factors within the child.

Some of Piaget's and Vygotsky's ideas have been incorporated into the Cognitive Acceleration through Science Education (CASE) program (see Shayer, 1999, for a review). This program was originally used about 25 years ago in secondary schools in the UK and has since been developed and extended. The CASE program has five main features, and a typical lesson involves working through them in the order listed:

1. *Concrete preparation*: This involves the teacher setting the scene for what is to come. He/she ensures that the students understand the scientific terms that will be used and know how to use any equipment that will be needed.
2. *Cognitive conflict*: This is created by exposing students to unexpected ideas or findings that don't fit their preconceptions and that can't easily be understood on the basis of their current ways of thinking.
3. *Construction*: This involves resolving the cognitive conflicts that have been created by means of discussions involving either small groups or the entire class.
4. *Metacognition*: This involves asking students open-ended questions requiring them to explain their thinking and focus on tricky issues. One student argued that this motivated students to have something worth saying, "Otherwise you'd feel a right prat" (Shayer, 2003, p. 482).
5. *Bridging*: This involves relating students' new understanding to other aspects of science and their everyday experience. For this to be successful, teachers must have identified good examples to produce the required bridging.

How does the CASE approach use Piaget's and Vygotsky's ideas? First, the teacher's use of his/her greater knowledge than the students to provide information during the concrete preparation and bridging phases is in line with Vygotsky's thinking. Second, the use of cognitive conflict resembles Piaget's emphasis on the importance of disequilibrium. Third, the phase of metacognition requires students to engage in a process resembling