Comprehending text is part of the daily routine for most individuals 6 years of age and older. We read texts for a variety of reasons. We want to be informed about the state of the world, learn about new domains, escape into fictional worlds, and perform certain actions (e.g., fill out a tax form). The skill to accomplish all these things, based on perceptual processes by which our eyes fixate black marks on white and then jump to the next set of marks, is surely one of the most remarkable accomplishments of our species. Text comprehension researchers are slowly but steadily uncovering the highly complex set of cognitive mechanisms that underlie the skill to comprehend text. In doing so, they have developed an impressive range of theories and tools to test them. This chapter provides an overview of this work. The question of how people convert the proverbial black marks on white to “stories in their heads” is, in and of itself, a fascinating one. In our attempts to address this question, we are learning a great deal more about what makes us human. After all, the ability to understand language is shared by no other animal. However, text comprehension research also has a distinct practical relevance. For example, text comprehension research provides useful information on how documents (e.g., manuals, questionnaires, drug prescriptions, educational texts) can be written so they convey their information with maximum clarity.
A THEORETICAL FRAMEWORK FOR TEXT COMPREHENSION

A fundamental assumption of this chapter is that text comprehension is an instance of complex information processing, and that it therefore complies with general principles of cognition. In this section, we use the influential construction-integration theory (Kintsch, 1988, 1998; see also Kintsch, 1974; Kintsch & van Dijk, 1978) as an illustration of a theoretical framework for text comprehension. Although individual components of construction integration remain controversial, many of the principles that it encompasses are generally accepted in the field.

The theory assumes that comprehenders process text one chunk (a clause or sentence) at a time. The processing of each chunk includes a construction phase and an integration phase. During construction, lower level processes such as orthographic analysis, word retrieval, and grammatical parsing analyze the current text chunk into idea units called propositions (Carroll, 1978; Clark & Clark, 1977; Kintsch, 1974; Townsend & Bever, 1982). For example, the propositions derived from the input sentence *Alice unlocked the wooden door with the key* are represented as P1 (UNLOCK, AGENT:ALICE, PATIENT:DOOR, INSTRUMENT:KEY), P2 (WOODEN, DOOR). Each proposition includes one predicate that, using the present notation, is written first (e.g., UNLOCK). The proposition also includes one or more concepts called arguments, each of which fulfills a distinct semantic function. The psychological validity of the proposition has been supported by numerous demonstrations: Text reading time increases systematically with the number of (a) propositions, holding constant the number of words (Kintsch & Keenan, 1973); and (b) different arguments in the text, holding constant the number of propositions (Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975). This analysis also suggests that arguments appearing in the same proposition are more strongly connected in memory than are concepts from different propositions (Ratcliff & McKoon, 1978; Weisberg, 1969).

During construction, the propositions that are expressed directly in the text are organized into a coherence network. This network is proposed to include several other types of ideas. Contained in the network are: (a) close associates of text ideas (e.g., both SPY and INSECT are associated with BUG even in the spying context); (b) inferences that contribute to the coherence of the text; and (c) text generalizations (e.g., *the student sprinkled chalk dust on the chair* might be generalized to PRANK).

During the subsequent integration phase of construction integration, activation accumulates in the network propositions that are most
highly interconnected with one another. In this process, there is a de-
activation of contextually inappropriate concepts, such as the INSECT
association to BUG in the spying context. This distribution of activa-
tion has been analyzed in the form of mathematical algorithms (e.g.,
Kintsch, 1988; Rumelhart & McClelland, 1986).

These language and discourse processes continually interact with
different memory systems. The processes of the integration phase
modify the original coherence network to produce an enduring long-
term memory representation of the text (Kintsch & Welsch, 1991). At
the end of each construction-integration cycle, a small number of
highly active elements of the current clause are held over in a limited
capacity working memory (Baddeley, 1986) for further processing
(Fletcher, 1981; Kintsch & van Dijk, 1978). It is by virtue of this carry-
over that a coherent text representation may be constructed.

The long-term memory encoding of the text consists of multiple lev-
els of representation, the surface structure, a mental representation of
the actual wording of the text, the textbase, a mental representation of
the explicitly stated semantic information in the text, and the situation
model, a mental representation of the state of affairs denoted in a text.
These levels of representation are inspected later. However, we first ex-
amine the methods typically used in text comprehension research.

TEXT COMPREHENSION RESEARCH METHODS

The methods typically used in text comprehension can roughly be di-
vided into two categories. The first category taps the comprehension
process as it unfolds. These methods are called online methods. The
second category of methods, historically older than the first, focuses
on the results of the comprehension process, the mental representa-
tions stored in the comprehender’s long-term memory. These meth-
ods are characterized here as memory methods.

Online Methods

The online methods can be subdivided into four types: processing-
load measures, activation measures, information-content meas-
ures, and brain-activity measures.

Processing-Load Measures

Processing-load measures are used to make inferences about the
relative amounts of cognitive resources needed to process linguistic in-
formation. This idea relies on the old assumption (e.g., Donders,
1868/1969) that processes that require more attentional or memory resources take longer than processes that require fewer resources. There are several ways in which processing times can be measured. Perhaps the most widely used method in text comprehension research is the self-paced reading task, in which people see segments of a text, typically clauses or sentences, on a computer screen one at a time. They advance through the text by pressing a key on the keyboard or a mouse button and the computer measures their reading times. Although this method is not as precise as the methods to be discussed next, it is precise enough to tap many of the processes in which discourse psychologists are interested.

The moving-window task allows more precise measurements. In this task, texts are displayed in their entirety (or one page at a time) on the computer screen, but with all the letters changed into dashes or slashes. By pressing a button or key, the reader makes a word visible and, by pressing the button again, the word reverts back to dashes and slashes and the next word appears. It is as if the reader is passing a small window across the masked text. In this fashion, reading times can be obtained for each word in the text while the reader still has an idea of the overall structure of the text.

The two preceding tasks share the characteristic that they rely on keypresses for the measurement of reading times. Thus, the resulting reading times reflect not only the time to read the sentence, but also the time needed to switch attention from comprehension to the key press and the time needed to execute the motor response. Furthermore, there is the possibility that the response times are not dictated by the amount of cognitive effort involved, but by the tendency to tap the keys in a rhythmic fashion. The eyetracking method does not have these disadvantages (see Rayner & Sereno, 1994, for a useful discussion). With this method, texts are presented on a computer screen and a participant’s eye movements are registered by small cameras as he or she processes the text. Thus, there are no key presses involved. Contrary to most people’s intuitions, reading is not a series of smooth horizontal sweeps of the eyes across a page. Instead, the eye pauses to take in information and then jumps to the next word (some short and highly predictable words such as the are often skipped). The pauses are known as fixations and the jumps as saccades. A third type of eye movement is known as regressions. Regressions occur when the eye moves back to an earlier word in the sentence. Along with long fixations, regressions are viewed as indexes of processing difficulty. Thus, eyetracking allows more precision than do the other methods. However, often this precision is not necessary; given that the method is relatively expensive and labor intensive (the eyetracker has to be calibrated for each indi-
 Activation Measures

Various measures are used to measure the availability of information to readers as they comprehend a text. These measures allow the researcher an insight into the mental representation as it unfolds in the comprehender’s mind. The three most commonly used methods to assess activation are: lexical decision, naming, and probe recognition.

Lexical Decisions. In traditional lexical decision experiments (e.g., Meyer & Schvaneveldt, 1971), people are presented with a string of letters and have to determine whether it is a word. A central finding is that the lexical decision for a word like doctor is faster when it is immediately preceded by a close associate like nurse than by the unrelated word bread or by a row of Xs. This implies that nurse activates or primes its close associate, doctor. In applications of this procedure to text comprehension research, one or more words from the text are used as primes, and lexical-decision latencies are used to make inferences about the degree to which concepts are activated. Thus, one could present the opening sentence of Alice in Wonderland:

1. Alice was beginning to get very tired of sitting by her sister on the bank and having nothing to do.

followed by the lexical decision item bored. The prediction would then be that, if people made the inference that Alice was bored, lexical decision times to this word should be shorter than those to a control word. This is because bored was activated during the reading of the sentence, thus facilitating the lexical decision when the actual word appeared on the screen.

Naming. Similarly, one can assess activation simply by having participants name words. The assumption here is that priming reduces naming latencies. It has been argued that naming has the advantage over lexical decision that it does not involve a yes–no decision component (Potts, Keenan, & Golding, 1988). Another advantage is that the procedure of the naming test requires fewer items than lexical decision, which requires that about half of the test items be nonwords.

Probe Recognition. A third way to assess activation is by using recognition probes. In this case, a word is presented and the participant’s task is to indicate whether it appeared in the text read up to that point.
Probe-recognition shares with lexical decision the fact that it involves a decision component and thus requires more items than does naming. Furthermore, it could be argued that probe recognition focuses the comprehender’s attention on the surface structure of the text. Nonetheless, many experiments have shown that the procedure is sensitive to situation-model levels of processing. Gordon, Hendrick, and Foster (2000) recently identified a new potential drawback of the probe-recognition task. In some experimental designs, participants seem to use a strategy in which they keep track of the words that might be probed. However, this problem is likely to be confined to situations in which the pool of to-be-recognized words is small.

**Information-Content Measures**

Activation measures allow the researcher to draw inferences about specific aspects of content and structure of the mental representation that is being constructed during comprehension. Information-content measures, in contrast, provide much more extensive information. However, the question is to what extent they reflect processes that are really going on during online comprehension or to what extent they reflect task demands.

One instance of information content measures is the think-aloud protocol. Think-aloud protocols have been used extensively in research on memory and expertise. In text-comprehension applications, participants are typically presented with one sentence or clause at a time and are instructed to comment on their understanding of the sentence or clause in the context of the text they are reading. The think-aloud method has been argued to provide an accurate window on cognitive processes provided the task is well defined (Ericsson & Simon, 1993). Although this is often the case in problem-solving experiments, it is often not the case in language-comprehension experiments. For that reason, think-aloud protocols are most often used as exploratory methods in conjunction with activation measures, with the latter being used as hypothesis-testing devices. A good example of this is the study of Trabasso and Suh (1993). Trabasso and Magliano (1996) developed a detailed procedure for analyzing think-aloud protocols (see also Magliano & Graesser, 1991).

Another way to elicit information is by using a question-answering procedure in which participants are asked specific questions about aspects of a text (e.g., why did X happen?). Graesser and Clark (1985) used this procedure to uncover the range of information that might be activated during online comprehension of a particular text segment. It is also possible to elicit information about the contents of compre-
henders’ mental representations by asking them to complete a text fragment. Presumably, some linguistic devices invite different continuations than others. For example, Gernsbacher and Schroyer (1989) showed how a cataphoric pronoun such as this (e.g., the egg in On the beach she found this egg) prompted participants to maintain its referent in their continuations to a greater extent than the indefinite pronoun an (as in On the beach she found an egg).

Content measures may provide useful insights into text comprehension. However, they are most profitably used as exploratory devices in conjunction with hypothesis-testing tools for tapping online processes, such as processing-load and activation measures. This is because of the potential vulnerability of content measures to strategic processes on the part of the participants.

**Brain Activity Measures**

A recent development in the study of discourse comprehension is the use of brain-activation measures. The method of measuring event-related potentials (ERP), by attaching electrodes to a participant’s scalp, has been used for quite some time in the study of language processes. However, until recently, most of its applications were at the levels of the word or sentence. A pattern of electrical activity of particular relevance to language comprehension is known as the N400 (e.g., Kutas, Federmeier, & Sereno, 1999; Kutas & van Petten, 1994), which is a shift in negative activity occurring about 400 ms after the presentation of a stimulus. The N400 is thought to reflect difficulty in semantic integration. For example, in the sentence He dug the hole with a pizza, pizza will yield an N400 with a larger amplitude than if the last word were shovel. Recently, van Berkum, Hagoort, and Brown (1999) demonstrated that contradictions in discourse yield robust N400 effects, which suggests that ERPs might be a useful tool in the study of discourse comprehension. ERPs have several advantages over reaction-time methods, the main one being their temporal resolution. Whereas the shortest response latencies in reaction time experiments are between 350 and 400 ms (in naming tasks), ERPs show effects at shorter latencies (e.g., although N400 reaches its peak at about 400 ms, the onset is much earlier). ERPs provide some information about where in the brain processes are localized, although the spatial resolution of ERPs is rather low compared with brain-imaging methods. A drawback of ERPs is that, to obtain stable patterns within participants, large numbers of items (at least 30) are needed per condition.

Brain imaging methods such as positron-emission tomography (PET) and functional magnetic resonance imaging (fMRI) allow the re-
searcher to localize cognitive processes in the brain. As such these methods hold great promise with respect to uncovering the neural substrates of discourse comprehension. Neuroimaging methods have been used in several recent studies of text comprehension (Carpenter, Just, Keller, Eddy, & Thulborn, 1999; Fletcher et al., 1995; Maguire, Frith, & Morris, 1999; Robertson et al., 2000). Among other things, these studies converge on the rather surprising finding that areas in the right hemisphere (and not the traditional language areas in the left hemisphere) are involved in the integration of information across sentences and in the construction of spatial representations from text. Although neuroimaging research is clearly still in its infancy and various methodological hurdles have to be overcome, it is likely that neuroimaging studies of text comprehension will soon begin to constrain theorizing about language comprehension.

Memory Measures

Memory measures provide information on how the mental representation constructed during comprehension is stored in and retrieved from long-term memory. The method of free recall is one of the first methods used to study discourse comprehension (Bartlett, 1932). However, questions have arisen regarding the extent to which recall provides a window on the long-term memory representation laid down as a result of the comprehension process (e.g., Corbett & Dosher, 1978). As Bartlett noted, recall is a constructive process. Thus, what is or is not in the recall protocol does not necessarily reflect what was constructed during comprehension. For example, people may remember information and then decide to edit it out of their recall protocol because they think it does not fit very well. However, having to recall information from texts (oral or print) is a daily activity for most people. Therefore, studying recall from text is an important topic in and of itself. Memory for text is discussed in a later section.

In cued recall, the participant is presented with part of a discourse and is asked to provide a missing part, such as the second sentence of a pair. Such tasks have yielded important insights into how information is integrated in long-term memory (e.g., Myers, Shinjo, & Duffy, 1987). As discussed later, recognition (“Did you see this sentence before?”) and verification (“Is this statement true with respect to what you just read?”) are used to assess the relative strengths of different levels of representation.

When measures are not timed, it is always possible that they are contaminated by strategic processes. Ratcliff and McKoon (1978) developed the method of primed recognition, which is not sensitive or is
much less sensitive to strategic processes. In this procedure, sets of statements pertaining to texts are shown, one at a time, to participants. Their job is to indicate as quickly and accurately as possible whether each sentence occurred in the text. Each set of items contains a prime-target pair, sentences from the text that the researchers hypothesize are more related in one condition than in one or more others. The recognition latencies are thought to reflect the strength of the long-term memory link between the nodes coding for the statements or events denoted by the prime-target pair (e.g., Zwaan, 1996).

**MULTILEVEL REPRESENTATIONS**

Various theorists have argued that, during the comprehension of texts, readers construct a mental representation of the text as well as the situations described in the text. For example, van Dijk and Kintsch (1983) proposed that readers construct mental representations of (a) the text’s surface structure, (b) the semantic meaning explicitly conveyed by the text or textbase, and (c) the situation described in the text, the situation model. The first two levels—surface structure and textbase—are sometimes collapsed. For example, Johnson-Laird’s (1983, 1996) propositional representation appears to be an amalgam of the surface structure and textbase. Researchers have not agreed on a single representational format for situation models. For example, Kintsch (1998) viewed situation models largely as propositional representations, but allowed mental images as well. Johnson-Laird (1996) viewed mental models as nonpropositional, but also as different from mental images. Barsalou (1999) proposed an entirely different view, in which the building blocks of mental representation are not abstract and amodal, but analog representations called perceptual symbols that are the result of perceptual activity in the brain. Thus, at the moment, situation models are less characterized by their structure than by their content. Much of the effort in research on situation models has gone into delineating which aspects of described situations are encoded in situation models and which are not.

**SURFACE AND TEXTBASE REPRESENTATIONS**

Empirical evidence suggests that readers typically maintain a brief memory record of the wording of a text (J. Anderson, 1974; Bransford, Barclay, & Franks, 1972; Clark & Sengul, 1979; Gernsbacher, 1985; Graesser & Mandler, 1975; Jarvella, 1971a, 1971b; Kintsch & Bates, 1977; Sachs, 1967). Estimates are that, under most conditions, decay
of the surface representation can be measured in seconds. However, in some cases, surface representations are maintained over long periods of time. In a series of experiments that ruled out various alternative explanations, Murphy and Shapiro (1994) showed that a critical factor in verbatim memory is the pragmatic context of a sentence. If a sentence has high interactive value (e.g., if it is an insult or a joke), the comprehender’s attention is focused on its wording, leading to better encoding and thus to a better surface memory (see also Kintsch & Bates, 1977). Similarly, Zwaan (1993, 1994) showed that surface memory can be affected by comprehenders’ expectations about the genre of a text they are reading. When participants thought they were reading excerpts from novels, they exhibited better surface memory than when they thought they were reading newspaper articles. In both cases, the same texts were used, ruling out wording as a factor. However, this does not mean that wording plays no role. For example, poetry and songs provide constraints on fit, other than semantics—namely, meter and rhyme. Consequently, people exhibit excellent surface memory for songs and poems (Rubin, 1995).

The textbase is thought of as the explicitly stated meaning of the text. Kintsch and van Dijk (1978) developed an influential model of textbase construction, whose function it was to predict recall from text. In this model, texts were hand parsed into propositions, and the propositions were arranged into a network according to several principles. For example, a connection between two propositions was made only if they co-occurred in a working memory buffer. Mechanisms such as the leading-edge rule specified that it is predominantly the most recent and important propositions that are held over in the working memory. With appropriate assumptions about the size of the buffer, the Kintsch and van Dijk model proved accurate at predicting recall from text. However, van Dijk and Kintsch (1983) noted that their earlier model failed to capture the most important aspect of comprehension: the construction of a situation model.

**SITUATION MODEL REPRESENTATIONS**

Why is it necessary to posit situation models? Why is it not sufficient to assume that comprehenders construct a textbase? The following example from Sanford and Garrod (1998) suggests an answer.

2. Harry put the wallpaper on the table. Then he put his mug of coffee on the paper.

It is rather straightforward to integrate these sentences. They call for a spatial arrangement in which the paper is on top of the table, the mug
3. Harry put the wallpaper on the wall. Then he put his mug of coffee on the paper.

Most readers balk at the second sentence. Surely it is impossible to put a mug of coffee on a vertical surface. If one simply assumes that comprehenders construct propositional textbases, then no such effect would be expected. The sentence pairs are equivalent in their propositional structure and connections.

Various experiments have shown that equivalent propositional structures lead to different behavioral responses. Many of these studies were reviewed by Zwaan and Radvansky (1998). There is evidence that comprehenders keep track of at least five situational dimensions during comprehension: time, space, characters, causation, and motivation. Zwaan and Radvansky (see also Zwaan, Magliano, & Graesser, 1995) proposed that the building blocks of situation models are mental representations of single events. These event representations are integrated during comprehension based on their overlap on each of the situational dimensions. Thus, an event that occurred at the same time and place as the previous event will, all other things being equal, be more easily integratable into the evolving situation model than an event that takes place at a different time or place. By the same token, two events that overlap on multiple situational dimensions are more strongly connected in the comprehender’s long-term memory representation than two events connected on only a single dimension. It has been shown that readers simultaneously monitor multiple situational dimensions during comprehension and that this is reflected in their long-term memory representations (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan, Radvansky, Hilliard, & Curiel, 1998). However, in most research on situation models, the focus is on single dimensions.

Spatial Situations

The dimension that received the most attention in early research on situation models was space. A decade prior to the coining of the terms mental model and situation model, Bransford, Barclay, and Franks (1972) demonstrated empirically that the spatial structure of the described situation can have a powerful effect on the comprehender’s memory. In this study, the participants listened to sentences such as 4
4. Three turtles rested on a floating log, and a fish swam beneath them.
5. Three turtles rested on a floating log, and a fish swam beneath it.
6. Three turtles rested beside a floating log, and a fish swam beneath them.
7. Three turtles rested beside a floating log, and a fish swam beneath it.

People who had heard 4 frequently incorrectly recognized 5, whereas people who had heard 6 rarely incorrectly recognized 7. This discrepancy cannot be explained by differential changes at the surface structure level of the test items. The only surface structure difference between the members of the pairs 4–5 and 6–7 is that the pronoun them has been replaced with it. However, the pairs differ with respect to the spatial layout they describe. Sentences 4 and 5 describe essentially the same situation: The turtles are on top of the log and the log is above the fish. Sentences 6 and 7, in contrast, describe different spatial situations. Sentence 6 has the fish beneath the turtles but not the log, whereas 7 has the fish beneath the log but not beneath the turtles. Thus, 4 and 5 are being confused because they describe the same situation. In contrast, 6 and 7 are less likely to be confused because they describe different situations.

Many more recent experiments have examined the role of spatial representations in text comprehension (see Zwaan & Radvansky, 1998, for a review). The most accurate characterization at this moment is that readers spontaneously construct spatial representations of some sort, but these spatial representations are often not detailed unless at least one of the following holds: (a) The reader has detailed prior knowledge about the spatial layout of the environment in which a story takes place, or (b) the reader is instructed, constrained, or intrinsically motivated to construct a detailed spatial representation.

Temporal Situations

More recently, researchers have started investigating the time dimension (A. Anderson, Garrod, & Sanford, 1983; Bestgen & Vonk, 2000; Carreiras, Carriedo, Alonso, & Fernandez, 1997; Mandler, 1986; Münte, Schiltz, & Kutas, 1998; Zwaan, 1996; Zwaan, Madden, & Whitten, 2000). Time is the situational dimension that is the most explicitly coded in language. For example, in languages such as English and French, a tense morpheme is attached to the main verb of each sentence, specifying the temporal position of the described event rela-
tive to the moment of utterance. Temporal markers such as time adverbials (e.g., “a few minutes later”) locate events even more precisely in time. The cited studies show that temporal markers have rather subtle effects on comprehension. Together the results of these studies suggest that comprehenders use temporal markers to organize narrated events into meaningful structures and that these markers have rather immediate effects on processing (Münte et al., 1998).

COHERENCE AND INFERENCE

To be comprehensible, texts must be coherent: Readers must be able to identify relations among the text ideas. The issue of coherence pervades the study of many central phenomena of text comprehension as becomes clear in this section.

The given-new analysis is a central principle of text coherence (Clark & Haviland, 1977; Haviland & Clark, 1974). Most sentences convey both given and new information. For example, the grammatical construction of *What Alice painted were the roses* conveys that the given information is that Alice painted something. The new information is that roses were painted. To fully grasp a sentence, a given-new strategy must be executed: The reader must distinguish the given and new information underlying a sentence, identify—in memory—a referent for the given idea, and link the new information to that referent.

The given and new ideas that underlie sentences are differentiated by many text characteristics (Clark & Clark, 1977). The grammatical structure of *What Alice painted were the roses* performs this function. Likewise, in *A dormouse drank the tea*, the definite article *the* designates *tea* as given information, whereas the indefinite article *a* signifies *dormouse* as new information. To be pragmatically cooperative, the writer must, by means of various devices, distinguish given and new sentence information in a manner that best coincides with his or her beliefs about the reader’s knowledge and the previous discourse history.

Coherence and Coreference

In identifying the referent of the given information, the reader establishes that the given idea and its referent corefer to the same entity in the world. Coreference can be signaled by many linguistic devices and semantic relations. Consider the sentence, *The dormouse drank the coffee*. This sentence has many possible continuations, such as *The coffee . . .*, *The java . . .*, and *The beverage . . .*. These continuations are, respectively, recognized as coreferential with *the coffee* by virtue
of lexical identity (coffee-coffee), synonymy, and category relations. Each continuation (e.g., The beverage . . .) is a definite noun phrase—that is, the nouns are modified by the definite article the. These noun phrases function as anaphors—expressions that refer back to an entity or constituent previously denoted in the text.

Anaphoric Resolution

The prototypical anaphor is the pronoun. In The dormouse drank the tea. It was tasty, it constitutes a definite pronoun that is coreferential with tea. Extensive research concerning the role of definite pronouns has imparted many general principles of anaphoric resolution. A central finding is that the surface, textbase, and situational representations of texts individually and interactively influence pronoun resolution.

In their surface expression, pronouns vary in gender (she–he), number (she–they), and person (she–I). As a result, pronouns may (sentence 8) or may not (sentence 9) unambiguously signify their referents.

8. Sally rewarded Ron because he was on time.

9. Tom rewarded Ron because he was on time.

Sentences like 8 with pronouns that, by virtue of their syntactic characteristics, have a single referent and take less time to read than those with multiple candidate referents (Ehrlich, 1980; Frederiksen, 1981; Springston, 1975). Likewise, the pronoun is understood more quickly when the text provides only one referent (Caramazza, Grober, Garvey, & Yates, 1977; Vonk, 1985).

If anaphoric resolution were a strictly accurate process, then the anaphor would access only its correct referent. In that event, the pronoun he in 10 would access only KEVIN:

10. Gary gave Kevin a lot of money and he spent it foolishly.

In a test of this hypothesis (Corbett, 1984), people read sentences such as 10 and 11, and then had to indicate whether a test name, such as Gary, appeared in the sentence.

11. Gary gave Kevin a lot of money and Kevin spent it foolishly.

People needed less time to recognize Gary after sentence 10 than 11, which suggested that the pronoun he accessed GARY as well as KEVIN.
A general finding in this realm is that if a text provides several candidate referents, the pronouns temporarily access all of them (Corbett & Chang, 1983; McKoon & Ratcliff, 1980).

Another surface characteristic that influences pronoun resolution is the physical distance in the text between the pronoun and its referent. In sentence 12c, the only possible referent for the pronoun it is BOOK, but the pronoun and its antecedent are separated by intervening sentence b.

12. a. Yesterday I met a woman who had written a book on viruses.
   b. She had studied them for years and years.
   c. It was selling very well.

People take longer to read sentence 12c in a sequence such as 12a–c than when book appears in the sentence immediately preceding 12c (Lesgold, Roth, & Curtis, 1979; see also McKoon & Ratcliff, 1980; O’Brien, Duffy, & Myers, 1986). This outcome is likely a joint function of (a) the deletion of BOOK from working memory on reading 12b, and (b) the change of focus away from BOOK in sentence b.

Semantic factors interact with surface variables in pronoun resolution as suggested by 13.

13. Clinton confessed to Archie because he wanted forgiveness.

Although Archie is the most recent noun that agrees syntactically with he, the implicit causality of the verb confessed guides the resolution of he to Clinton—that is, it is usually the characteristics of the confessor that prompt confessing. In a study designed to probe these issues, participants read sentences with because clauses that were consistent (sentence 13) or inconsistent (sentence 14) with the implicit causality of a preceding verb (Caramazza et al., 1977).

14. Clinton confessed to Archie because he offered forgiveness.

The participant had to judge whether a pronoun in the because clause referred to the first or second noun in the preceding clause. Consistent with their analysis, Caramazza et al. found that pronoun resolution took longer in the inconsistent sentences than consistent sentences (see also Au, 1986; Ehrlich, 1980; Matthews & Chodorow, 1988; Springston, 1975). Another semantic relation that influences anaphoric resolution is the category link between text elements (Corbett, 1984; Garrod & Sanford, 1977). For example, reading times for sentences containing a definite anaphor denoting a category (e.g., the vehi-
cle) are longer when the antecedent is an atypical member of the category (e.g., a tank) than when it is a typical member (e.g., a bus).

Finally, pronoun resolution is influenced by the situation model of a text. As discussed earlier, the different dimensions of the situation model, such as time, spatial location, and character, may interact in complex ways. For example, after the shift from one situational episode to another, it is difficult to resolve a pronoun that refers to a minor narrative character mentioned in the former episode. However, resolving a pronoun to the main character who was last mentioned in the previous episode does not pose difficulty (A. Anderson et al., 1983). In another study that addresses narrative situation models, people read passages in alternate versions, such as ones that included either 15b or 15b (Morrow, 1985).

15. a. Paul caught the flu and was feeling pretty awful. He told his eldest son Ben to keep the house quiet. He got up from bed to go to the bathroom, irritated by the noise. ( . . . )
   b. That noisy Ben was messing up the kitchen.
   b. Ben was wondering when his father would feel better as he ate in the kitchen.
   c. The floor felt cold on his feet.

The readers generally interpreted the pronoun his in c to refer to the main character (Paul) even in a version of the passage that mentioned the son (Ben) more recently (e.g., sentence 15b). Only when the perspective was changed to that of Ben (15b) was Ben preferred as the referent of his (see also Malt, 1985).

The relative contributions of the surface, textbase, and situational representations to anaphoric resolution have been subjected to empirical scrutiny. Garnham, Oakhill, Ehrlich, and Carreiras (1995) showed that when definite anaphors appear in text shortly after their antecedent, surface (e.g., gender agreement) and situational representations make independent contributions to resolution. This contradicted a linguistic hypothesis that stated that pronoun anaphors are resolved exclusively on the basis of a referential situation model (Sag & Hankamer, 1984). Garnham et al. also reported that, with more intervening text, the contribution of the surface level quickly diminished.

In other instances, the situation model appears to predominate. Consider sequences 16 and 17:

16. I was really frightened by a Doberman. (They are dangerous beasts./ It is a dangerous beast.)
17. Last night we went to hear a new jazz band. (They/it) played for nearly five hours.
The *they* versions of 16 and 17 are ostensibly ungrammatical because the pronouns disagree in number with their antecedents. However, Gernsbacher (1991) proposed that the first sentence of each sequence supports a situation model that can function as a referent for the plural *they*. In 16, a *Doberman* signifies the category of all Dobermans, and in 17, *band* consists of a number of musicians. In fact, people needed less time to read and assigned higher naturalness ratings to the *they* versions of 16 and 17 than the *it* versions (Oakhill, Garnham, Gernsbacher, & Cain, 1992).

**Toward a Model of Anaphoric Resolution.** The present treatment might suggest the following simple analysis of anaphoric resolution: “Anaphors constitute retrieval cues for their referents. Anaphoric resolution is guided by a variety of syntactic, semantic, and discourse relations between anaphor and antecedent.” However, even aside from the aforementioned contributions of situation models to anaphoric resolution, other anaphoric phenomena reveal this analysis to be a considerable oversimplification. First, if anaphors function mainly as retrieval cues for their antecedents, then pronouns should be readily replaceable with corresponding noun phrases. As a result, sentences 18 and 19 should be equally felicitous.

18. Does Bob think that his performance will go well?
19. Does Bob think that Bob’s performance will go well?

It is apparent, however, that sentence 19 is either ungrammatical or refers to two different Bobs (Halliday & Hasan, 1976).

Second, the use of definite noun phrase and pronoun anaphors is interwoven with issues of discourse focus. In this regard, Vonk, Hustinx, and Simons (1992) showed that when people are asked to continue a story with a given word, pronouns prompt people to maintain the current text focus, whereas noun phrases (character names in the Vonk et al. study) initiate a topic shift. Conversely, comic strips that depicted topic continuity favored participants’ descriptions beginning with a pronoun, whereas those depicting topic shifts were described with names (Vonk et al., 1992, Experiment 2).

Furthermore, pronouns are preferred to definite descriptions for referents that form the current text focus (Gordon & Scearce, 1995). In a revealing study of this phenomenon, Almor (1999, Experiment 1) examined sequences like 20 and 21:

20. It was the robin that ate the apple. The bird seemed very satisfied.
21. What the robin ate was the apple. The bird seemed very satisfied.
The grammatical constructions (called *clefts*) of 20 and 21, respectively, designate *robin* and *apple* as the new (or *focused*) elements. People needed less time to read the anaphoric noun phrase *the bird* in 20 than 21. More important, Almor (1999, Experiment 3) observed the opposite pattern when the anaphor repeated the expression of its antecedent:

22. It was the bird that ate the apple. The bird seemed very satisfied.

Earlier, sequence 19 illustrated that the repetition of noun phrases can be grammatically unsuitable. Together these results show that the suitability of noun phrase repetition depends on the current condition of focus. Almor presented these observations in the framework of the proposal that anaphoric phenomena are regulated by the processing cost of identifying the referent and computing the new information signaled by the anaphor. This offers a useful step toward a unified model of anaphoric reference, and hence a deeper understanding of text coherence.

**Inference Processes**

Almost every facet of comprehension is at least partly inferential: Complex inferences are needed to identify the intended meaning of ambiguous words (*The farmer filled the PEN*), the grammatical analysis of sentences (*The old man the boats*), and the thrust of ordinary comments (*Do you have a watch?*). As a result, inference processing has been a major focus of investigation for the past 30 years.

In this section, inferences in the textbase and situation model representations are scrutinized. If from the sentence *Alice unlocked the door* the reader inferred that Alice used a key, the resulting textbase representation might be (UNLOCK, AGENT:ALICE, PATIENT:DOOR, INSTRUMENT:KEY). In contrast, understanding *The lightning struck, The hut collapsed* might require an inference in the causal *situation model* about the relationship between the two events.

**Bridging Inferences**

Of central importance are bridging inferences, which link the current clause to the preceding text. Consider sequence 23:

23. a. The pitcher threw to first base.
   
b. The ball sailed into the field.
In keeping with the given-new analysis, understanding 23b requires identifying the referent of the given information, *the ball*. Sentence 23a does not mention a ball, but the reader can draw on world knowledge to determine that the ball was the object that the pitcher threw. Measures of reading time indicate that the processes of bridging a definite noun phrase (*the ball*) to the prior text are more consuming of cognitive resources than resolving anaphors based on identical phrasing, synonyms, and pronouns (Haviland & Clark, 1974; Lesgold et al., 1979).

Similarly, understanding sequence 24 depends on detecting a causal relation between its two sentences (Black & Bern, 1981).

24. The boy walked over to the refrigerator, bumping a bowl he had left on the table. Suddenly, it fell off the edge and broke.

By identifying the links between the current and prior text, bridging inferences preserve text coherence. In their absence, ordinary text would appear as disjointed as *Alice unlocked the door. The ball sailed into the field*. Evidence derived from numerous measures indicates that bridging inferences routinely accompany comprehension. First, Black and Bern (1981) reported that causally related sentences are more effective recall cues for one another than are similar temporally related sentences (e.g., obtained by replacing *bumping* with *seeing* in sequence 24). They concluded that readers inferentially integrate the sentences of 24.

Second, reading time has been used to probe the bridging of the following sequences:

25. Tony’s friend suddenly pushed him into a pond. He walked home, soaking wet, to change his clothes.
26. Tony met his friend near a pond in the park. He walked home, soaking wet, to change his clothes.

According to rating data, the sentences of 25 are causally closer than those of 26. Reading time for the second sentence of such sequences systematically increased with causal distance (Keenan, Baillet, & Brown, 1984; Myers et al., 1987). This outcome was interpreted to reflect the readers’ execution of bridging inference processes.

Other investigations of bridging inference have used stimulus materials such as the following:

27. The spy threw the report in the fire. The ashes floated up the chimney.
28. The spy threw the report in the fire. Then he called the airline.
Both 27 and 28 permit the inference that the spy burned the report, but only in 27 is this inference needed to bridge the ashes to the first sentence. People need more time to answer Did the spy burn a report? after sequence 28 than 27. Furthermore, answer time for 27 is indistinguishable from that observed when the first sentence explicitly states that the spy burned the report (Singer & Ferreira, 1983). Likewise, naming time for a word representing the central inference (BURN) is similar in an explicit condition and after the bridging sequence 27, but longer after sequence 28 (Potts et al., 1988). These results favor the conclusion that bridging inferences reliably accompany comprehension. The status of inferences such as those underlying sequence 28 are examined later in the section on Elaborative Inferences.

**Theoretical Interpretations.** The fundamental conclusion that bridging inferences routinely accompany comprehension is relatively uncontroversial. However, the development and comparison of two contemporary theories of comprehension and inference—memory-based text processing and constructionism—have focused on more subtle aspects of the computation of bridging inferences. The central feature of the memory-based analysis is that the current clause, plus additional ideas carried over in working memory, function implicitly as memory cues (e.g., McKoon, Gerrig, & Greene, 1996; O’Brien, Lorch, & Myers, 1998). In this capacity, the current clause induces the activation of prior text and relevant world knowledge by means of passive resonance processes (Ratcliff, 1978) on the basis of surface form or semantic similarities. Consistent with this analysis, the current text can restore, to working memory, matching concepts from earlier in the text and ideas associated with those concepts. A strong version of the memory-based processing analysis is that comprehension computations are restricted to (a) ideas carried over in working memory, (b) current text, and (c) ideas that have resonated to the current text (Albrecht & Myers, 1995). The memory-based analysis receives support from the findings that people’s access to a text idea is regulated by its recency and degree of elaboration in the text and its similarity to the contents of working memory (Albrecht & Myers, 1995; O’Brien & Albrecht, 1991; O’Brien, Albrecht, Hakala, & Rizzella, 1995).

The competing constructionist theory has as its central principle that readers engage in a search after meaning (Bartlett, 1932; Bransford et al., 1972; Graesser, Singer, & Trabasso, 1994). Two assumptions of constructionism are that readers (a) monitor coherence at all levels of text representation, and (b) routinely seek explanations for text outcomes such as physical effects and intentional actions. The thrust of these assumptions is that the computations of comprehen-
sion do not strictly proceed from lower representations, such as surface form, to higher ones, such as the situation model. In fact, situational inferences may sometimes take precedence (Barton & Sanford, 1993; Sanford & Garrod, 1998).

Constructionist theorists have proposed that causal situation models take the form of networks of text events interconnected by links of physical, motivational, and psychological causation. Consistent with this analysis, memory for text and importance ratings of text ideas are influenced by the degree of connectedness of these ideas and by their appearance on the main causal chain underlying the text (Trabasso, Secco, & van den Broek, 1984; Trabasso & Sperry, 1985). The need to inferentially link a text statement to other, causally related ideas, is associated with higher reading times for those statements (Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990). That result converges with the finding that achieving a complete understanding of a sentence entails the validation, with reference to world knowledge, of candidate causes of a text outcome (Singer, Halldorson, Lear, & Andrusiak, 1992).

The representation of the situation models of goal and cause that underlie text has provided a field for comparing memory-based text processing and constructionism. Several findings have been invoked as support for memory-based processing. In one study, people read passages such as one describing Mary as having to make an air reservation but getting sidetracked (Albrecht & Myers, 1995). Later she is described as going to bed—an action that is inconsistent with the goal of booking a flight. However, the detection of this inconsistency, as indexed by the reading time for the second event, depended on surface overlap between the two critical clauses. In a similar vein, readers overlooked a nearby cause for a father becoming angry (“broken window”) when the text provided a different, highly elaborated cause (“lost keys”) earlier (Rizzella & O’Brien, 1996, Experiment 2a). In both studies, the authors concluded that the results were consistent with the memory-based explanation and inconsistent with the constructionist explanation because the absence of surface overlap (Albrecht & Myers, 1995) and presence of a competing, elaborated cause (Rizzella & O’Brien, 1996) prohibited the detection of text causes.

In another study, the description of a character eating a cheeseburger provided access to the concept VEGETARIAN even when the character had been described as no longer being a vegetarian or even never having been one (O’Brien, Rizzella, Albrecht, & Halleran, 1998; see also Gerrig & McKoon, 1998). Consistent with memory-based processing, this result indicates that the semantic relation between CHEESEBURGER and VEGETARIAN is sufficient to reinstate VEGETARIAN as a candidate cause of the character eating a cheeseburger.
TARIAN to working memory even when it has little bearing on the present context. Gerrig and McKoon (1998) claimed that the retrieval of text ideas not germane to the current causal structure of the text (e.g., VEGETARIAN in this example) contradicted constructionism.

However, consistent with constructionism, there have been numerous demonstrations that readers detect cause and goal relations that span moderate text distances. In these studies, either the causal conditions have been compared with control conditions that are matched for the degree of surface overlap (Long, Golding, & Graesser, 1992; Singer & Halldorson, 1996; Suh & Trabasso, 1993; van den Broek & Lorch, 1993) or surface overlap between the current outcome and its distant cause was absent (Richards & Singer, 2001). In another study, the amount of intervening text did not influence the detection of the relation between a text outcome and its antecedent cause (Lutz & Radvansky, 1997). This outcome was argued to challenge a tenet of memory-based text processing.

The distinction among these and other theoretical analyses (see Graesser et al., 1994, for a review) remains controversial. However, consideration should be given to the possibility that memory-based and constructionist processing constitute complementary rather than competing analyses. The impact on comprehension of passive resonance processes has been effectively scrutinized in the framework of the memory-based analysis, but those processes are not inherently inconsistent with the constructionist theory (Graesser et al., 1994). Conversely, the detection of complex, situational relations need not be prohibited in the memory-based framework (Rizzella & O’Brien, 1996).

**Elaborative Inferences**

Earlier it was recounted that people correctly answer *Did the spy burn the report?* faster after sequence 29 than sequence 30 repeated here. This indicates that readers inferentially bridge the sentences of 29.

29. The spy threw the report in the fire. The ashes floated up the chimney.

30. The spy threw the report in the fire. Then he called the airline.

The implication that the spy burned the report is also carried by sequence 30, but text coherence does not depend on that inference. Ideas that are strongly implied by a discourse context but do not bear on coherence are called *elaborative inferences*.

Considerable evidence based on measures of cued recall (Corbett & Dosher, 1978), answer times (Singer, 1980), and speeded judgments
of single words (McKoon & Ratcliff, 1986; Potts et al., 1988) has indicated that elaborative inferences, perhaps counter to intuition, do not reliably accompany comprehension. These studies scrutinized inferences ranging from implied roles, such as the participation of the police when a *burglar was arrested*, to causal predictions, such as the breaking of a delicate vase when it is described as being dropped. These results have been understood in terms of the observation that every message suggests so many plausible inferences that to compute, during comprehension, all of them would overwhelm the available cognition resources: In other words, a computational explosion would result (Charniak, 1975; Rieger, 1975).

There is, however, accumulating evidence that certain elaborative inferences accompany comprehension. It is instructive to consider some prominent examples. First, people appear to encode implied semantic features that are especially relevant to text. For example, people verify *Tomatoes are red* faster than *Tomatoes are round* after reading a text that emphasizes the color of tomatoes and vice versa for one that emphasizes the shape of tomatoes (McKoon & Ratcliff, 1988).

Second, people elaborate category terms to their specific members. Consider sentence 31:

> 31. Julie was convinced that spring was near when she saw a cute red-breasted bird/robin in her yard.

When sentence 31 uses the category term *bird*, people generate the elaborative inference ROBIN as is evidenced by reading times for a continuation sentence (Garrod, O’Brien, Morris, & Rayner, 1990; McKoon & Ratcliff, 1989, Experiment 2; O’Brien, Shank, Myers, & Rayner, 1988), speeded recognition (McKoon & Ratcliff, 1989), and cued recall (R. Anderson, Pichert, Goetz, Schallert, Stevens, & Trollip, 1976; see also Dubois & Denis, 1988; Whitney, 1986).

Third, it was mentioned earlier that people do not reliably draw elaborative inferences about implied roles. In one study (Lucas, Tanenhaus, & Carlson, 1990), however, people heard sequences such as 32 over headphones.

> 32. a. There was a broom in the closet next to the kitchen.
   b. Bill swept the floor every week on Saturday.

The participants were required to make lexical decisions about implied instrument words such as *broom*, which for this sequence appeared on a screen coincident with the end of the spoken word *week*. Lexical decision time was lower for implied instruments than for con-
trol words. Thus, although an inference about a broom does not routinely accompany the comprehension of 32b, it does when the relevant concept has been made available by the text (see also McKoon & Ratcliff, 1992).

Consistent with the latter observations, readers may also encode implied semantic roles into text representations. In this regard, Mauner, Tanenhaus, and Carlson (1995) proposed that sentence 33a is grammatically more suitable than sentence 33b.

33. a. The game show wheel was spun to win a prize.
b. The game show wheel spun to win a prize.

Rationale clauses such as to win a prize require the participation of an agent. The short passive construction The game show wheel was spun implies the involvement of an agent, but the intransitive The game show wheel spun does not. Consistent with this analysis, in people’s examination of to win a prize, reading time was longer and the number of “does not make sense” responses greater in sentence 33b than 33a.

Fourth, there is evidence that people draw elaborative inferences about text themes. For example, 1 second after encountering the last word of The townspeople were amazed to find that all the buildings had collapsed except the mint, people made faster lexical decisions about earthquake than when the same word followed a control sentence (Till, Mross, & Kintsch, 1988).

Elaborative Inferences and the Construction-Integration Analysis. The construction-integration model (Kintsch, 1988) provides a framework for understanding the circumstances in which elaborative inferences reliably accompany comprehension. Recall that the reader initially constructs a coherence network consisting of explicit text propositions and their close associates, coherence preserving inferences, and thematic generalizations. During integration, activation accumulates in those elements of the coherence network that are highly interconnected. Consider, in this context, some of the classes of elaborative inferences just examined. On reading sequence 34, the concept RED might be constructed as an associate of TOMATO.

34. The still life would require great accuracy. The painter searched many days to find the color most suited to use in the painting of the ripe tomato.
However, RED would also bear links to COLOR, the compound concept RIPE TOMATO, and other text ideas. As a result, it would remain highly activated after the integration process. ROUND might be similarly constructed after reading 34, but, in the absence of connection with many text ideas, would likely not survive integration.

Likewise, the sequence There was a broom in the closet next to the kitchen. Bill swept the floor every week on Saturday explicitly denotes BROOM. The joint presence in working memory of BROOM, SWEEP, and FLOOR might be sufficient to result in the encoding of inferential links among them.

There is experimental evidence consistent with this analysis of elaborative inference processing. Consider sentence 35:

35. After standing through the three-hour debate, the tired speaker walked over to his chair.

One-quarter second after reading sentences such as 35, Keefe and McDaniel’s (1993) participants viewed a word that represented an elaborative implication of the sentence, such as sat. Naming time for the test word was about the same when it was implied and explicitly stated in the sentence, but longer when it followed a control sentence with similar wording but quite different meaning. This outcome suggests that the readers had drawn an elaborative inference about sitting. However, when just one additional sentence intervened between sentence 35 and the test word, naming time in the inference condition resembled the control condition rather than the explicit condition. This result pattern generally coincides with the construction-integration notion that ideas which initially appear in the coherence network may not endure integration. Immediate probing might suggest that an elaborative inference has accompanied comprehension, but it might not long survive in its competition with other encoded ideas.

MEMORY FOR TEXT

To benefit from understanding a text, one generally must remember it. The quality and quantity of people’s memory for discourse provides a reflection of the mental processes of comprehension, representations that result, and changes in those representations over time. Text memory has been examined in terms of question answering, sentence recognition and verification, and free and cued recall of text. These tasks
provide complementary and converging evidence about retrieval from text.

Stages of Question Answering

A prevailing research approach to question answering has been the identification of the processing stages that contribute to successful answering. First, questions must be encoded to their propositional form. For example, *What did Alice paint?* would be denoted as (PAINT, AGENT:ALICE, PATIENT:?). The related yes–no question *Did Alice paint the roses?* would be encoded as (PAINT, AGENT:ALICE, PATIENT:ROSES?). These two questions, respectively, request (a) the identity of the painted object, and (b) the accuracy of ROSES. This notation highlights the particular importance, in question answering, of distinguishing between the given and new information (Clark & Haviland, 1977). For *What did Alice paint?*, it is given that Alice painted something. The question requests the identity of the new information—namely, the painted objects.

Next, the answerer must categorize the question. Questions may interrogate the semantic roles of a question statement, such as its agent, instrument, or location (e.g., Fillmore, 1968). Other questions address complex relations between the question statement and related ideas, including relations of cause (*why*?), reason (*how*?), and time (*when*?) (Graesser & Murachver, 1985; Lehnert, 1978; Trabasso et al., 1984). Both wh- (*Who painted the roses?*) and yes–no (*Were the roses painted by Alice?*) questions may be formulated about any of these categories (Singer, 1990). The correlation between interrogative term and question category is imperfect: For example, *why* can query either a reason (*Why did Alice paint the roses?*) or a cause (*Why were the roses red?). Therefore, question categorization typically requires syntactic and semantic analysis by the answerer.

Strategies of Question Answering. Strategy selection has been proposed to form a distinct stage of question answering. To answer a question about a text, one can attempt to retrieve the question statement from memory or judge its plausibility (Camp, Lachman, & Lachman, 1980; Lehnert, 1977; Reder, 1982). For example, to answer *Did Alice paint the roses with a brush?*, one could retrieve that statement from one’s memory of the story or evaluate it with reference to one’s knowledge of painting.

Strategy selection is guided by factors (a) extrinsic to the question such as task instruction, and (b) intrinsic to the question such as its current activation. The current activation of a question might be due to
the recency of encountering the question statement (Reder, 1987). In one study, participants read stories and subsequently answered questions about them immediately or 2 days later under instructions either to retrieve or judge the plausibility of the questions (Reder, 1982). The influence of instruction on strategy adoption is reflected in Table 3.1. Mean correct answer time was lower for the retrievers than for the plausibility judges in immediate testing, when the verbatim details of the story were still available, but the reverse pattern was found after 2 days. However, delay not only influenced answer times, but also strategy: Answer time under the plausibility instruction was longer in immediate than delayed testing. This suggests that, contrary to their instruction, the plausibility judges were attempting to retrieve the question statements in immediate testing. In general, people are versatile in their adjustment of answering strategy. They can adjust their strategy for each question on the basis of (a) advice (Reder, 1987), and (b) whether the question represents a recent or much earlier story (Reder, 1988; Singer & Gagnon, 1997).

Memory Search in Question Answering. Memory search in question answering has been clarified by J. Anderson’s (1974, 1976) fanning paradigm. In this paradigm, the participants learn facts like A teacher is in the garage, An architect in the park, and A teacher is in the church. In testing, recognition time varies directly with the number of facts in which the concepts of the test fact participated. For example, answer time is longer for A teacher is in the garage than for An architect is in the park because teacher occurred in two facts and architect in only one. This effect is attributed to weaker concept–fact links for concepts that have participated in more facts (J. Anderson & Reder, 1999; but see Radvansky, 1999, for a different view).

An important qualification of the fan effect is that people can focus memory search on a queried category. In one study, people learned that characters like concepts in two categories: For example, the dentist likes five (specified) cities and one animal. When asked, Does the

<table>
<thead>
<tr>
<th>Strategy</th>
<th>0 min</th>
<th>20 min</th>
<th>20 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieval</td>
<td>2.51</td>
<td>2.65</td>
<td>2.85</td>
</tr>
<tr>
<td>Plausibility judgment</td>
<td>3.21</td>
<td>2.68</td>
<td>2.61</td>
</tr>
</tbody>
</table>
dentist like giraffes?, answer time is mainly determined by the number of animals that the dentist likes (McCloskey & Bigler, 1980). Similarly, people can focus their memory search on an interrogated semantic role (Singer, Parbery, & Jakobson, 1988).

Answering questions about text clearly requires focused memory search: Why did Alice paint the roses? and How did Alice paint the roses? demand different answers. Graesser (Graesser & Clark, 1985; Graesser & Murachver, 1985) proposed that each combination of text-statement and question categories is associated with distinct memory search procedures. For example, why-action questions (Why did Alice paint the roses?) require the tracing of reason links between facts in the backward direction, resulting in answers such as, “She wanted to help the gardeners.” Evaluation of this proposal revealed that a large proportion of people’s answers to questions about narratives comply with the analysis, and that people assign higher goodness-of-answer ratings to answers that conform with the analysis than those that do not (Graesser & Murachver, 1985).

In some instances, the focused search of memory may reveal that the text representation includes no information in the queried category. For example, the memory search initiated by Did Alice paint the roses with a roller? may reveal no information about the instrument that Alice used. Frequently, this circumstance permits a rapid indication that the answerer does not know the answer (Collins, Brown, & Larkin, 1980; Costermans, Lories, & Ansay, 1992; Glucksberg & McCloskey, 1981; Nelson & Narens, 1980; Singer, 1984). More generally, there is evidence that people can assess the availability of pertinent information before they retrieve it from memory (Reder, 1987).

Parallel Processes in Question Answering

The inspection of question-answering stages such as encoding, categorization, strategy selection, and memory search carries the suggestion that the stages are executed serially. Intuitively, however, it seems likely that the appearance of who at the outset of Who painted the roses? would permit the question to be categorized, and the search for an agent to begin before the question was completely encoded. In a test of a hypothesis of this sort, Robertson, Weber, and Ullman (1993) determined that reading time for sentences beginning with interrogative terms is longer than for control declarative sentences. They concluded that the memory search initiated by the interrogative word occurred in parallel with question encoding and slowed the latter process. These observations are generally consistent with proposals of parallel processing in highly interconnected representations.
3. TEXT COMPREHENSION

Such processing is highly interactive in the sense that low- and high-level processes mutually influence one another. In the example of Robertson et al. (1993), it is presumed that there is interaction among encoding, categorization, and memory retrieval. Likewise, Reder’s (1988) finding that the recency of the question statement affects strategy selection strongly suggests that memory retrieval (which provides the index of recency) and strategy selection influence one another. Contemporary question-answering theories exhibit many of the properties of connectionist computational models (Rumelhart & McClelland, 1986). According to the ACT-R model of fact retrieval (J. Anderson, 1993; J. Anderson & Reder, 1999), for example, the connection strengths between a fact and its constituent concepts are determined by a connectionist learning rule. During subsequent testing, activation is spread in parallel from all of the concepts of the test fact (e.g., A teacher is in the garage) until a fact is retrieved.

Recall and Recognition of Text

People frequently strive to remember text in circumstances in which they are not directly queried about it. The most open-ended text-retrieval task is termed free recall. Some robust phenomena were revealed in studies of students’ recall of newspaper articles. Kintsch and van Dijk (1978) classified their participants’ recall responses as reproductions, which corresponded to propositions expressed directly in the original article, and reconstructions, which represented sensible guesses but did not have counterpart propositions in the text. Figure 3.1 reveals that text recall becomes systematically more reconstructive with increased test delay; after 6 weeks, the two classes account for approximately the same proportion of all recall responses. Singer (1982) measured a similar pattern—one that was hardly distinguishable between reading in the laboratory and in the natural setting (i.e., before experimental recruitment). Singer also reported that the participants reproduced 26% of 124 text ideas in immediate recall and only 6% after 6 weeks.

The convergence of reproduction and reconstruction in text recall is further clarified by people’s text recognition. Text recognition is influenced by the relationship between the text and the test sentence—the test sentence can express an explicit text idea, paraphrase the text, represent a reasonable inference based on the text, or constitute an inaccurate distractor. In immediate recognition, the acceptance rate for explicit items is considerably higher than for paraphrases and inferences. With increased delay, the explicit acceptance rate does not decline much, but the rates for paraphrases and inferences rises toward
the explicit rate, and even the acceptance of distractors increases somewhat (Kintsch et al., 1990; Reder, 1982). This profile is illustrated in Fig. 3.2.

This striking pattern can be explained as follows. In immediate testing, participants have reasonably detailed representations of the surface form, textbase, and situation model at their disposal. The recognition rate reflects the number of levels with which the test item is consistent. Thus, explicit items, which are consistent with all three levels, are accepted the most often, followed by paraphrased items, which do not match the surface form, but are consistent with the textbase and situation model. Finally, inferences are accepted the least often because they match neither the surface nor textbase representations and are consistent with the situation model only (Kintsch, Welsch, Schmalhofer, & Zimny, 1990; Schmalhofer & Glavanov, 1986). With increased delay, the verbatim form of the message is all but forgotten. Likewise, there is considerable loss of the idea content (textbase) of the text over time (Kintsch & van Dijk, 1978; Singer, 1982). Therefore, delayed recognition increasingly depends on the robust situation model.
As a result, the relative recognition advantages of explicit and paraphrased test items are lost.

The complex profile in text recognition meshes with the proposal, scrutinized earlier, that text comprehension results in multiple levels of representation. It also meshes with the convergence, with increased delay, of reproduction and reconstruction in text recall (see Fig. 3.1). In recall, like recognition, the participant must increasingly rely on the situational representation as the retention interval grows. Insofar as the situation model represents the blending of text information and general knowledge, many delayed recall responses are reconstructions—they do not correspond to a text proposition.

**CLOSING COMMENTS**

We have reviewed the psychological theories and methods of psychological text comprehension research. Text comprehension proceeds in cycles and involves the simultaneous construction of three levels of representation: surface structure, textbase, and situation model. Key among these levels is the situation model. The situation model is a mental representation of the state of affairs denoted by a text. In most cases, the purpose of text comprehension is the construction of a situa-
tion model. Several decades of research have revealed text comprehension to be a highly complex cognitive process that involves most areas of cognition: perception, working memory, long-term memory, problem solving, and imagery.

Research is currently underway to study the neural mechanisms that subserve text comprehension. Furthermore, the traditional assumption that language comprehension involves the disembodied construction of abstract propositional networks is being challenged by recent proposals and supporting evidence that language comprehension involves analog, perceptual representations that reflect how we, as humans, interact with our environments (Barsalou, 1999; Glenberg, 1997; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002). It is to be expected that these developments will make fundamental contributions to the understanding of what goes on in our mind/brains when we process the black marks on white.

REFERENCES


3. TEXT COMPREHENSION


3. TEXT COMPREHENSION


3. TEXT COMPREHENSION


3. TEXT COMPREHENSION


