

3 Cognitive theories in discourse-processing research

Gail McKoon and Roger Ratcliff

Much has been learned in the past decades about how readers comprehend discourse. In large part, advances have come about because empirical methods were developed in the 1980s that allow the examination and separation of online processes, off-line processes, and the memory representations that result from these processes. This chapter reviews these methods and shows what can be interpreted from them. We begin with several general points, continue with a discussion of particular methods, and then review new methodologies that have been developed in the last several years. In the course of these discussions, we use examples from our own research, but many others (including all the authors of the other chapters in this book) have provided similar examples.

1. The first general point is that cognitive processes can be separated into those that occur quickly and automatically and those that occur more slowly and strategically (Posner, 1978). For example, for the sentence "The janitor swept the classroom," a reader might infer that the janitor used a broom, and do so automatically or strategically. Usually, in the field of discourse research, interest has focused on what a reader understands without special, strategic effort. However, in other fields, such as education, it might be strategic effects that are of most interest.
2. The second general point concerns the process by which information is retrieved from memory automatically. Theories over the past several decades have described the process as "resonance" (beginning with Lockhart, Craik, and Jacoby, 1976, and Ratcliff, 1978; first applied in discourse-processing research as the "minimalist hypothesis" by McKoon and Ratcliff, e.g., 1986; 1992).

The notion of resonance is that information retrieval is a fast, passive process (i.e., an automatic one) by which cues in short-term memory interact with all the information in long-term memory in parallel. This fast, easy process accesses all the information in memory, but the degree to which any specific cue in short-term

memory evokes any specific piece of information in long-term memory depends on the strength of the association between them in memory. This strength determines the degree to which the information is evoked. In its essence, fast, passive, parallel retrieval provides information "for free."

For discourse processing research, it is important to stress that resonance-type retrieval can operate both during reading and during memory tests. For the janitor sentence, the contents of short-term memory during reading would be the words of the sentence and their meanings, including the words "swept" and "janitor" and their meanings. The degree to which "broom" became available during reading (and possibly encoded) would depend on the strength of the associations among "broom," "swept," "janitor," and their meanings.

The fact that an item in short-term memory can match information in long-term memory to varying degrees has an important implication for investigations of whether and what kinds of inferences are automatically encoded during reading. In early discourse processing research, questions about inference encoding were almost always phrased in an all-or-none manner. With resonance models, this is no longer appropriate.

The application of the resonance view to memory tests can again be illustrated with the "janitor" sentence. The word "broom" might be tested in single-word recognition, where subjects decide whether test words are old or new according to whether they had appeared in a previously read sentence. If "broom" were presented as a test word, it would make available "The janitor swept the classroom" sentence to the degree that "broom" was encoded as part of its meaning during reading. If "broom" was encoded during reading to a sufficiently high degree, then subjects might respond, in error, that the word "broom" had actually appeared as part of the janitor sentence.

Resonance theories make no distinction between a lexicon of information about words and memory for other kinds of information; a cue evokes all kinds of information at once. In contrast, in research prior to about 1990, it was often said that the processes that identify a word, for lexical decision, for example, can be divided into "pre-lexical" processes and "post-lexical" processes (e.g., Forster, 1981). Pre-lexical processes determine the information about a word that is available from the lexicon and post-lexical processes determine whether and how that information is relevant to whatever language comprehension task is at hand. The moment at which information from the lexicon became available to post-lexical processing was labeled "the magic moment" (Balota, 1990). In this view, the effects

of some variable on pre-lexical processes occur independently of effects on comprehension. For example, the word "swept" in the janitor sentence might facilitate lexical access for "broom" but this would indicate nothing about whether "broom" is inferred as the instrument of "swept."

3. The third general point is that retrieval from long-term memory is context dependent (Tulving, 1974): Memory can never be assessed without taking into account the environment of other cues in which a particular cue to memory is tested. The environment includes other information in short-term memory at the time of the memory test, and it also includes general characteristics of the test situation as a whole. The assumption of context-dependent retrieval has been incorporated into all current models of memory. A correlate of it is that it is never possible to know that some piece of information has not been encoded into memory; there always might be some test environment in which it is (to some degree) retrieved.
4. The fourth general point is that the combination of resonance retrieval and context-dependent retrieval provides an interpretation of priming effects that is quite different from the traditional spreading activation interpretation (but see ACT*, Anderson, 1983). In text processing research, priming has been examined between words, phrases, and sentences, for example, "dog" and "cat" in lexical decision. Ratcliff and McKoon (1988) explain findings like this with a compound-cue model, a model that is an implementation of context-dependent retrieval into global retrieval models (e.g., Gillund and Shiffrin, 1984). In the model, all the contents of short-term memory are combined to match against long-term memory. For "dog" and "cat," responses to "cat" are facilitated because the compound cue "dog cat" makes available the associations between them in long-term memory. The model explains priming data both qualitatively and quantitatively.

The import of Ratcliff and McKoon's model lies, in part, in its sharp contrast with traditional spreading activation accounts of priming by which the prime activates the target before the target itself is presented — "dog" as a test item immediately activates "cat" and so responses to "cat" as a test item are facilitated. In the compound-cue model, there is no advance effect of "dog;" the association can affect performance only when "dog" and "cat" are together in short-term memory.

In discourse research, the compound-cue model applies in three situations. In one, subjects are given a list of texts to read followed by a recognition test, where the tested items can be single words, phrases, or sentences. To the extent that the prime and target are

associated in the encoded representation of a text, their compound should facilitate "old" responses to the target. For the sentence "The janitor swept the classroom," the memory test might be recognition of single words.

Responses to "janitor" immediately preceded by "classroom" would be facilitated to the extent that they were encoded together when the sentence was read. Likewise, if "broom" was encoded during reading, then "janitor" or "classroom" should facilitate "old" responses (i.e., incorrect responses) to "broom."

The second situation in which the compound-cue model applies is that, during reading, the words of a text form compounds of the words and their meanings, and these compounds are matched against memory. For the janitor sentence, the compound would be made up of: "janitor" and its meanings, "swept" and its meanings, "classroom" and its meanings, the meanings evoked by all the pairs ("janitor-swept," "janitor-classroom," and "swept-classroom"), and the meanings evoked by the triple "janitor-swept-classroom." Comprehension of this sentence would be facilitated to the extent that general knowledge provided pre-existing associations among parts of the compound.

The third situation occurs when items (words, phrases, or sentences) are tested "online," that is, during reading or immediately after it. In this procedure, the test item forms a compound with the text information preceding it. From the compound-cue point of view, information from the text does not activate information in memory about the test item before the test item is presented, again in sharp contrast to a traditional spreading activation mechanism.

Measuring comprehension and memory

There are many empirical paradigms that have been used in the discourse-processing literature, and they vary in the interpretations of data that are possible from them. Here we discuss five of the most popular paradigms.

1. In cued and free recall, subjects are asked to produce all the information that they remember from a text. The first issue with these methods is that subjects' responses may reflect information that was encoded during reading or information that was constructed at the time of the memory test (and in either case, processing might be automatic or strategic). For example, if "The janitor swept the classroom" was a to-be-remembered sentence and the test of memory was cued recall, subjects might produce "broom" in response to "janitor."

This response might be due to "broom" being inferred when the sentence was read and encoded into memory as part of the representation of the sentence, or it might be that subjects connect "broom" to the janitor sentence only when "janitor" is presented as a cue.

In an early demonstration of this problem, Corbett and Doshier (1978) gave subjects lists of sentences to read and remember, and followed them with cues for recall. They found that "hammer" was an equally good cue for the sentence "John pounded the nail with a rock" as for the sentence "John pounded the nail." Either "hammer" was not encoded with "John pounded the nail" or it was encoded with both "John pounded the nail" and "John pounded the nail with a rock." Either way, it cannot be argued that "hammer" was encoded during reading as the instrument of pound.

The second issue with cued and free recall is that subjects can edit their responses, perhaps to make them seem more coherent. Suppose, for example, they recall that something happened, but not to whom it happened. They might attempt to make their recall more coherent by deleting the "something happened" information from it or they might attempt to make their recall more coherent by generating new information that was not part of the to-be-remembered information.

A third issue is that subjects' performance in recall tasks presents a classic case of an item-selection artifact: The subject, not the experimenter, selects what responses to make. A subject might have a perfect verbatim representation of a to-be-remembered text, but still produce only partial information at test (perhaps because responding with full information would require too much time). A subject might remember, for example, everything from a multisentence text but decide to produce only the highlights.

2. Several paradigms have the same problems as cued and free recall. In a "topicality" experiment, subjects are asked to generate a topic sentence for a text they have read, in a multiple-choice experiment, they are given several possible choices to answer a test question, and in a question-answering experiment, they are given open-ended questions. In all three of these paradigms, subjects' responses are likely to reflect information constructed at the time of the test. Also, subjects are usually given as much time as they want to make their responses and that may encourage them to adopt special strategies.

Two other paradigms have an additional problem. In a "close" experiment, subjects are given sentences with blanks in them and for each, they are asked to fill in the blank with a word appropriate to the sentence. Subjects in a "think-aloud" experiment are asked to

talk about what they are thinking as they read through a text. For these paradigms, responses are at least as likely to reflect constructed information as automatically encoded information, and highly likely to reflect special strategies. The additional problem is that they are disruptive to normal reading.

Subjects cannot decide what to say aloud or what to fill into a blank without, at least temporarily and to some extent, losing track of what they are reading.

Finally, empirical measures must have bases in theory, but for none of these paradigms are there well-understood models for how responses are produced. What determines what information subjects produce when they "think aloud"? What are the processes by which a word is generated to fill in a blank? Of all the information from a text that is available in memory, how does a subject pick which to report? One implication of this problem – the lack of theoretical understanding of response mechanisms – is that there is no way to know what to do when the methods give different results. If subjects produce the correct referent of a pronoun when they think aloud but not when they are asked an open-ended question, how can it be decided whether the referent was encoded automatically during reading? This is in strong contrast to the highly developed models for retrieval in paradigms for which responses are fast and automatic.

3. In online tests, a subject's task is usually single-word recognition or lexical decision. Online tests are essential to investigations of reading. They show the information that is available to a subject at particular points in time. However, there are two important issues for the interpretations of online data that are sometimes overlooked. One is that online tests can show what information is available during reading but they cannot show what is encoded into memory. "Broom" might be available during reading, because it is evoked by "The janitor swept the classroom," but not encoded into the representation of the sentence in memory. Knowing what was encoded into memory requires an off-line task. The second issue, mentioned above, is that online tests tap only the interactions of the test item with the text being read; there is no way to separate out the contribution of the test item alone.

To illustrate misconstruals of online data we give three examples, two for lexical decision and one for naming latencies. First, in a classic study, Onifer and Swinney (1981) presented sentences like the two below auditorily. Somewhere during each sentence a string of letters for lexical decision appeared visually. The items of interest were ambiguous words such as "bug." Test words that matched one or

the other of the word's meanings ("insect" or "spy") appeared either immediately after the ambiguous word or several words later. Onifer and Swinney's finding was that response times (RTs) for both words were shorter than control test words at immediate test, but at the later test, RTs were shorter only for the word appropriate to the context of the sentence. In other words, for both sentences below, RTs for both "insect" and "spy" were speeded immediately after "bug," but at the later test, only "spy" was speeded for the first sentence and only "insect" for the second.

For several weeks following the exterminator's visit, they did not find a single bug anywhere in the apartment.

For several weeks following the discovery that they were being watched by the CIA, they kept checking the phone for a bug or a hidden video camera.

The standard and highly influential interpretation of this result was that both meanings of an ambiguous word are activated immediately; context does not operate quickly enough to activate only the appropriate meaning. However, the view that interactions between test word and context determine RTs gives a different interpretation: Responses to "spy" and "insect" are fast with immediate test because they are in short-term memory in a compound cue with "bug." Responses for the inappropriate word are slow later because "bug" is no longer in short-term memory.

In another study with online lexical decision, Nicol and Swinney (1989) examined whether readers understand the implicit objects of verbs. In "the police stopped the boy that the crowd at the party accused of the crime," the object of "accused" is "boy." They found what they took to be a very surprising result – responses to "girl," a strong associate of "boy," were speeded, relative to a control, immediately after "accused" compared to immediately before. They attributed this to readers filling in "boy" as the implicit object. However, again, consideration of interactions between test word and context gives a different interpretation, one demonstrated by McKoon and Ratcliff (1994) with the sentence "The crowd at the party accused the boy," for which responses to "girl" were speeded from before to after "accused" even though "boy" was not an implicit object.

In addition to recognition and lexical decision, naming latencies have often been used to test for the availability of words during reading. For example, for the "bug" sentence, "insect" and "spy" and their control words would be presented just as for lexical decision except that the readers' task would be to name the words as quickly as possible. If naming latencies showed no difference between "insect"

and "spy" relative to their controls at either of the two test positions, then it would have been concluded that readers did not comprehend the context relevance of one of "bug's" meanings over the other. This conclusion is incorrect because of the scaling problem that naming latencies are much shorter than lexical decision or recognition latencies. This means that effects that are significant in the latter two cases may not be significant for naming latencies. In other words, if an experimental variable does not affect naming latencies, it cannot be concluded that that variable is not effective in comprehension.

4. In off-line tests, some amount of unrelated material is presented between the text to be remembered and the test items. The test is usually recognition, for which the test items might again be single words, phrases, or sentences. Typically, subjects are asked to respond as quickly and accurately as possible. If off-line recognition is to be used to investigate automatic processes, then two requirements must be met to rule out strategic processes: One is that responses must be faster than the amount of time that would be needed for strategic processes, which has been shown in a number of studies to be around 700 ms (e.g., McKoon and Ratcliff, 1989a), and the other is that the probability with which the test items of interest occur must be low so that subjects do not guess the purpose of the experiment and adopt special strategies for it.

Off-line tests are appealing for several reasons. One is that the experimenter chooses the items to be tested; there is no item-selection artifact. Another is that they show what was actually encoded during reading, not just what was available. Still another is that the experimenter can choose what combinations of items to test. For the janitor sentence, the experimenter might want to look at memory for the individual words and so the test would be single-word recognition. The experimenter could also choose conjunctions of words, and so the test would be recognition of phrases or whole sentences.

McKoon and Ratcliff (1986) were the first to demonstrate how off-line recognition could be used to investigate encoded inferences, specifically inferences about what would happen next in a discourse. For example, for the sentence "The cameraman was ready to shoot close-ups when the actress fell from the 14th story," "dead" would be what happened next. The aim of the experiment was to show that "dead" was encoded as an inference when the sentence was read. McKoon and Ratcliff's paradigm met the conditions necessary for testing automatic processes: The experimental test words occurred with a low probability and responses were faster than 700 ms.

There were four conditions in the experiment, two versions of the sentence and two test contexts. One version of the sentence was the one given above. The other was "Suddenly the director fell upon the cameraman, demanding that he get a close-up of the actress on the 14th story," which provides an essential control. It uses words from the sentence that predicts "dead" that might be pre-experimentally related to "dead," but rearranges them into a sentence that does not predict "dead." With this control, it cannot be the case that responses to "dead" reflect pre-experimental associations instead of inferences that were encoded during reading.

The compound-cue view suggests that automatic retrieval for sentences such as the actress one will depend on the context in which the "dead" test words are presented. McKoon and Ratcliff used two contexts; the test words were immediately preceded by a prime from their sentence, "actress" for "dead," or the word "ready" (as a neutral prime).

What McKoon and Ratcliff found was that responses to "dead" were more likely to be "old" (an error) when the prime was "actress" than when it was "ready." It was this retrieval-context effect that led us to describe the "dead" inference as "minimal" (McKoon and Ratcliff, 1992).

Soon after McKoon and Ratcliff's (1986) paper, Potts, Keenan, and Golding (1988) proposed a different mechanism to explain the "actress dead" result. Their idea was that the prime "actress" activates its sentence and then when "dead" is presented, it is checked against the sentence for compatibility. Since "dead" is compatible with the predicting sentence, responses to it tend to be errors. With the neutral prime "ready," the sentence is not activated, and so there are fewer errors.

However, there are two reasons to think that Potts et al.'s hypothesis is not correct. One is that there was not sufficient time in McKoon and Ratcliff's experiment for compatibility checking – the SOA between the prime and the test word was only 200 ms, and subjects were instructed to respond within 600 ms of onset of the test word. The second reason stems from other experiments by McKoon and Ratcliff (1989a). With the same paradigm as McKoon and Ratcliff (1986), they used sentences that should evoke a member of a category, sentences such as "The young attorney wanted to make sure she had fresh juice for breakfast so she squeezed the fruit herself." Responses to "oranges" tended to be errors to the same degree with "ready" as a prime as with "attorney." Because "ready" could not provide access to the "attorney" sentence, Potts et al.'s hypothesis can be rejected.

5. It is often claimed that the time it takes subjects to read a word, sentence, or text gives insight into what they understand from the

sentence. Sometimes reading times are measured by subjects controlling the amount of time they spend reading each word, phrase, or sentence, and sometimes they are measured by eye movements, looking at first-pass reading times, total reading times, and probabilities of regression.

However, reading times do not show unambiguously how readers comprehend texts. Reading times can show "glitches" in processing but not the results of them. The problem is that reading times are subject to speed/accuracy trade-offs. Slowdowns in reading can occur because a subject, facing some difficulty in comprehension, takes the time to understand the textual information fully and correctly, or because the subject slows down enough to appreciate the difficulty but not enough to fully resolve it (McKoon and Ratcliff, 1992, were the first to discuss this issue). Failing to resolve difficulties encountered during reading is probably a characteristic of most everyday reading.

The view that comprehension is often incomplete is nicely illustrated by the "Moses illusion" (Erickson and Mattson, 1981). When asked how many animals of each kind Moses took on the Ark, most subjects easily answer "two," not noticing that the sentence has evoked incorrect information.

The Moses illusion also illustrates another point: The absence of a slowdown does not mean that there was no difficulty in processing. The assumption that all possible comprehension difficulties are reflected in reading time is an assumption that is, and must be, incorrect.

Another issue with reading times is that when they slow for one set of sentences but not a comparison set, then it cannot be determined whether comprehension of one set was facilitated relative to the other or one set was inhibited relative to the other. For the two texts below, Sanford and Garrod (1981) found slower reading times for the "control the class" sentence for the first compared to the second text. They attributed this to inhibition: The "on his way to school" sentence generated the inference that John was a schoolboy, and the mismatch between this inference and John controlling the class was thought to be responsible for the slower reading times. Instead it could be that no inference was generated about John being a schoolboy from the "on his way to school" sentence. Instead, reading times were faster with the second text because comprehension of "control the class" was facilitated by its good match with John teaching math.

John was on his way to school. The bus trundled slowly along the road. He hoped he could control the class today.

John was not looking forward to teaching math. The bus trundled slowly along the road. He hoped he could control the class today.

Questions about comprehension and memory

Our argument in this chapter is that reading comprehension and memory can be fruitfully investigated only when the methods used to investigate them are based on an explicit conceptualization of retrieval processes: Fast automatic processes can be separated from slower strategic ones and, when retrieval is automatic, it is a resonance-type process by which cues in short-term memory are matched passively against all the information in long-term memory in parallel in a context-dependent fashion. In the next sections, we illustrate the impact of this conceptualization of retrieval with a series of examples from our own research, although all of the same points can be made with the research of many others. All of the examples represent inferences of one sort or another: Simple ones such as the referent of a pronoun and more complex ones such as what will happen next in a story.

Available and encoded. If a reader can be said to have understood an inference, then the required information must have been available during reading and the inference must be encoded into memory. McKoon and Ratcliff (1980) and Dell, McKoon, and Ratcliff (1983) showed both for nominal anaphors. They used four-sentence texts, such as the one to follow, in two conditions. In the first, the fourth sentence begins with an anaphor that refers to an entity mentioned in the first sentence. For the following text, readers should infer that "the criminal" refers to the burglar mentioned in the first sentence. In the second condition, the fourth sentence begins with an entity not previously mentioned, "dog."

The burglar surveyed the garage.

The banker and his wife were on vacation. Newspapers were piled at the curb.

The criminal OR a dog slipped away from the streetlamp.

Dell et al. (1983) used an online single-word recognition paradigm. The words of a text were presented one at a time for 250 ms per word (about normal reading time for college students). At any point during a text, a test word could be presented instead of the next word of the text. For the burglar text, for example, "burglar" was presented as a test word immediately after "the criminal" in the first condition and immediately after "a dog" in the second condition. Dell et al.'s hypothesis was that the relation between the criminal and the burglar would lead to facilitation of responses to "burglar," and this is what they found.

Dell et al. also tested the word "vacation" immediately after "the criminal" and "a dog." It was possible that something other than the relation between "the criminal" and "the burglar" was responsible for

the facilitation of "burglar" in the "criminal" condition. Perhaps, for example, "a dog" might have been more difficult to understand for some reason than "the criminal." Counter this possibility, Dell et al. found that RTs for "vacation" did not differ between the two conditions.

Dell et al. also tested the word "garage" immediately after "the criminal" and "a dog" to show that faster responses for "burglar" after "the criminal" were not due solely to the a priori semantic relatedness of "criminal" and "burglar." Connecting "the criminal" to "burglar" when "the criminal" is read should have made not only "burglar" more available but also words directly related to "burglar," such as "garage," and Dell et al. found that it did.

We stress again that facilitation for "burglar" and "garage" does not mean that the connections among "burglar," "criminal," and "garage" were encoded into memory when the criminal sentence was read, only that they became more available when "the criminal" text was read than when "the dog" text was read.

To look at what connections are encoded into memory, McKoon and Ratcliff (1980) used off-line single-word recognition. For the burglar text, the test word "streetlamp" was immediately preceded by the test word "burglar." If the inference that "the criminal" referred to "the burglar" and was encoded when "the criminal" was read, then responses to "streetlamp" should be facilitated in the first condition compared to the second and, again, this is the result that was found.

Available but not relevant. When, during reading, the words and meanings of a text evoke strongly associated information from long-term memory, then that information can be irrelevant to the meaning of the text as a whole. To show the immediate availability of such information, McKoon and Ratcliff (1989b) used the same paradigm as Dell et al., with single-sentence texts of the three kinds shown below. In each case, the test word was "sit," for which the correct response was "new," and it was presented immediately after the final words of the sentences.

1. *After shopping for hours, the grandmother headed for her favorite chair.*
2. *After shopping for hours, the grandmother headed for her favorite store.*
3. *After shopping for hours, the grandmother found the perfect chair.*

Responses to "sit" were slower and less accurate after the first sentence than the second, indicating that it was only the final word, "chair," of the first sentence that evoked "sit." The more interesting finding was that responses to "sit" were as slow and inaccurate after the third sentence as after the first. Consistent with a compound cue, the relation between "chair" and "sit" affected responses even when "sit" was irrelevant to the text.

Connections among the elements of a text. Of all the types of inferences that have been investigated in discourse research, the most common are inferences that connect the elements of a text. The key to predictions about such connections is that the elements of a text that are currently being read evoke other information via resonance-type retrieval.

Consider the text below, used in experiments by McKoon and Ratcliff (1980).

Early French settlements in North America were strung so thinly along major waterways that land ownership was not a problem. The Frenchmen were fur traders, and, by necessity, the fur traders were nomads. Towns were few, forts and trading posts were many. Little wonder that the successful fur trader learned to live, act, and think like an Indian. Circulation among the Indians was vital to the economic survival of the traders.

To fully comprehend this text, a reader needs to infer the connection between "circulation among the Indians was vital" and "the fur traders were nomads." This connection is not explicit in the text – it is left unsaid that it was the fur traders for whom circulation was vital. However, if readers do infer this, then "circulation among the Indians was vital to the fur traders" should evoke earlier-mentioned information about the fur traders, for example, that they were nomads. In contrast, "land ownership was not a problem" should not directly evoke earlier information about the fur traders because the land ownership information is a general fact about early settlements.

To test whether readers infer the connection between "circulation among the Indians was vital" and "the fur traders were nomads," McKoon and Ratcliff used an off-line test. The test items were phrases from the text and for each, subjects were asked to decide whether it was true or false according to a previously read text. The manipulation was one of priming. The target test item was "the fur traders were nomads." If the test item that immediately preceded it was "circulation among the Indians was vital," responses were faster and more accurate than if the preceding test item was "land ownership was not a problem." This result is especially noteworthy because the French settlements text is fairly complex, yet the combination of inferred information and resonance-type retrieval leads to the encoding of appropriate connections.

Pronoun resolution without pronouns. Perhaps the most powerful of the demonstrations of resonance-type retrieval during reading has been provided in studies by Gerrig and McKoon and colleagues (Gerrig and McKoon, 1998; 2001; Greene, Gerrig, McKoon, and Ratcliff, 1994; Love and McKoon, 2011; McKoon, Gerrig, and Greene, 1996). In these studies, we have used texts such as the one below that has three parts:

INTRODUCTION MENTIONS THREE CHARACTERS

Jane was dreading dinner with her cousin, Marilyn. She complained loudly to her roommate, Gloria. "Every time I go to dinner at my cousin's I get sick. Gloria asked, "Why did you agree to go?" Jane said, "Because I'm too winpy to say no." Jane went off to have dinner.

MIDDLE PART DOES NOT MENTION COUSIN MARILYN OR JANE

Gloria decided to cook something nice for herself for dinner. "As long as I'm alone," she thought, "I'll eat well." Gloria searched her refrigerator for ingredients. She found enough eggs to make a quiche.

CONCLUSION

Gloria was still up when Jane arrive home about midnight. Gloria asked Jane, "Did she make the evening unbearable?"

We labeled the first sentence of the conclusion the "reunion" sentence. The idea was that information in this sentence would evoke, via resonance-type retrieval, information from the introduction sentences. In the example above, "Gloria was still up when Jane arrived home about midnight" would evoke the information that Jane and Gloria were roommates, that Jane had a cousin named Marilyn, that Jane was going to dinner with her cousin, and so on. The hypothesis was that evoking this information would increase the availability of "cousin Marilyn." To demonstrate this, Gerrig et al. used an online, single-word recognition paradigm. The test word "cousin" was presented immediately before or immediately after the reunion sentence. The result was a speed-up in RTs from the first of these test points to the second, relative to a control condition. Moreover, when "cousin" was tested immediately after the sentence with the pronoun ("Did she make . . ."), there was no further decrease in RTs. It is this finding that we labeled "pronoun resolution without pronouns."

One important comment about Gerrig et al.'s experiments is that, almost certainly, the idea to look for pronoun resolution without pronouns would not have happened without the resonance-based retrieval framework. Previously, it had been assumed that pronouns find their antecedents by searching backward through a text. The resonance view eliminated backward search as the main process by which referents of pronouns are automatically found.

The results of Gerrig et al.'s experiments are an especially compelling use of the resonance-based framework. The reunion sentence evoked "cousin" via a fast, passive process, and an explicit pronoun did not further increase "cousin's" availability. The configuration of cues at the test point immediately after the reunion sentence (the information in the reunion sentence and the test word "cousin") served to draw

together appropriate portions of the text; the pronoun in the second sentence of the conclusion was not essential.

Integrating text information and general knowledge. Connections among text elements can also depend on information that is evoked from long-term memory about general knowledge of the world. Allbritton, McKoon, and Gerrig (1995) provided a demonstration of this with narratives that evoked common metaphor-based schemas. The two narratives below illustrate the two conditions of the experiment. The hypothesis was that in the first narrative, "the city's crime epidemic" would evoke a schema for "epidemic" as a metaphor. The final sentence, "public officials desperately looked for a cure," is consistent with this metaphor, and so it should be closely connected to "the city's crime epidemic." In the second narrative, "the city's crime epidemic" would still evoke the epidemic-as-metaphor schema, but the intervening information between this and the final sentence moves the narrative to a nonmetaphoric use of "cure" in the final sentence.

In Allbritton et al.'s experiment, the paradigm was a priming manipulation with test statements presented for true/false judgments. "Public officials desperately looked for a cure" was the target test statement, and it was immediately preceded by "The city's crime epidemic was raging out of control." Consistent with the hypothesis, "true" responses to the target were faster for the schema-matching version than the mismatching version.

METAPHOR-MATCHING

The most recent crime statistics confirmed what New Yorkers had suspected. All major categories had increased significantly from last year. The city's crime epidemic was raging out of control. Extra police patrols had been ordered, but they had little effect. If anything, they seemed to aggravate the problem. Patrols in problem areas only inflicted more violence on neighboring areas. Soon, the violence began to infect even "safe" neighborhoods. Public officials desperately looked for a cure.

METAPHOR-MISMATCHING

The most recent crime statistics confirmed what New Yorkers had suspected. All major categories had increased significantly from last year. The city's crime epidemic was raging out of control. Though badly needed, police patrols in the city could not be increased. A new and virulent strain of pneumonia was plaguing the force. Almost a third of the department was infected already. The disease had struck at the worst possible time. Public officials desperately looked for a cure.

In another experiment that looked at the interactions among textual information and real-world knowledge, McKoon, Ratcliff, and Seifert (1989) used narratives that expressed schemas such as "going to the beach," "going to a restaurant," and "going shopping." The two narratives below both instantiate "going to the beach."

1. *Linda decided to skip work on Thursday and go to the beach. At the beach, Linda found the parking lot to be surprisingly full for a weekday, but she eventually found a spot. The beach, too, was crowded, but Linda was still able to spread her towel in a dry place close to the water. Not wanting to get a sunburn, Linda put on some sunscreen. After lying on her towel for some time, Linda was getting hot so she decided to take a dip, and dove into the refreshing water. Although she usually enjoyed the power of her executive secretary position, today she was happy not to be at work. After a short swim, Linda toweled off and packed up her things for the long walk to the car.*
2. *Because the sun was shining so brightly, Nancy decided to spend the day by the sea. When she had gotten to her favorite seaside spot, Nancy parked her car under a tree. Nancy walked quickly over the hot sand until she found an empty space where she could lay her blanket. Hoping to add some color to her pale skin, Nancy splashed on some baby oil. The sun was very strong, so Nancy decided to get up and go for a swim. Nancy slowly strolled out into the cool ocean. Her hobby was bird-watching, so she watched the cliffs above her for nesting swallows. When she finally felt water-logged, she headed back to her blanket. She dried off for a while in the warm sun and then dressed for the trip home.*

In McKoon et al.'s experiment, subjects were given forty-two narratives to read, and among them (widely separated in the list of forty-two) were two narratives for each of twenty-one schemas. After the forty-two stories, there was a list of 216 test phrases given for true/false judgments. There were three conditions of interest:

1. SAME SCHEMA, SAME NARRATIVE
*found an empty space for her blanket
slowly strolled out into the ocean*
2. SAME SCHEMA, DIFFERENT NARRATIVE
*spread her towel in a dry place
slowly strolled out into the ocean*
3. DIFFERENT SCHEMA, DIFFERENT NARRATIVE
*looked over the wine list and ordered chablis (from a restaurant-schema narrative)
slowly strolled out into the ocean*

The hypothesis for the experiment was that schema-related information in one narrative would evoke schema-related information in another narrative when the two narratives shared the same schema. Consistent with this, responses to "slowly strolled out into the ocean" were faster in

the second condition than the third. In fact, responses in the second condition were just as fast as in the first. The conclusion is that information in the narratives evoked general knowledge about a relevant schema, and when two narratives evoked the same schema, they became associated in memory.

Degrees of availability. In the resonance retrieval framework and the memory models from which it was derived, pieces of information evoke each other to varying degrees. Some pieces of information may be so strongly evoked that they become encoded into the representation of a text, and others so weakly evoked that they play no part at all in comprehension.

This view of memory suggests that the referents of anaphors may not be uniquely identified during reading. A pronoun, for example, might be uniquely identified only if the referent it evoked was higher in strength than any other possible referents. Greene, McKoon, and Ratcliff (1992) showed how this might occur with simple texts such as the one below.

Mary and John were doing dishes after dinner. One of them was washing while the other dried. Mary accidentally scratched John with a knife and then she dropped it on the counter.

At the end of this text, Mary and John have both been explicitly mentioned twice and so they may be equally salient (with perhaps a slight edge to Mary as the subject of the last sentence). Greene et al. used the same paradigm as the Dell et al. (1983) study described above. The test words of interest were the two possible referents of the pronoun in the last sentence plus another word from the text that was not connected directly to the two characters. For the text above, these words were "Mary," "John," and "dishes." When these test words were presented at the end of the final sentence, responses to "Mary" and "John" were faster than responses to "dishes," but they themselves were equally fast. The interpretation of this result is that, at the time of the test, Mary and John were equally salient and so "she" did not differentially evoke Mary over John. A crucial feature of the design of this experiment is the use of the test word "dishes." Without it, the failure to find a difference between "Mary" and "John" would be simply a null hypothesis.

If information is evoked during reading to varying degrees, then the natural question is what variables can make some entity in a text more or less available than another. One such variable is the syntactic position of a concept in a discourse. As an example, consider "John smeared the wall with paint" and "John smeared paint on the wall." "Wall" is said to be more salient in direct-object position than object-of-preposition position and so, it is hypothesized, "John smeared the wall with paint" implies that

the whole wall was affected by the painting activity whereas "John smeared paint on the wall" allows the wall to be only partially painted.

To obtain empirical evidence about how syntactic positions affect saliency, McKoon, Ratcliff, Ward, and Sproat (1993) used texts with two versions:

The librarian was furious when she got to work today. Somebody had inserted some magazines inside some newspapers late last night.

OR

The librarian was furious when she got to work today. Somebody had inserted some newspapers inside some magazines late last night.

In an off-line, single-word recognition experiment, responses to "magazines" were faster when it was in direct object position than when it was in object-of-preposition position.

Concepts in a text can also be made more salient by their relations to general knowledge. A narrative from McKoon and Ratcliff (1992) had the two versions below. The hypothesis was that "picking a flower for someone" is particularly salient because it corresponds to well-known schemas about gifts (e.g., the giver seeks to please the receiver, the receiver will likely thank the giver, etc.). "Smelling a flower for a moment" does not correspond to a well-known schema.

A girl was enjoying the warm spring weather. She walked up to the entrance of a park and bent down to an ornamental display to pick a flower for her sister.

OR

bent down to an ornamental display to smell a flower for a moment.

Then she walked into the park and down to a small stream where some ducks were feeding. She smiled to see seven tiny ducklings trailing behind their mother.

McKoon and Ratcliff (1992) tested the hypothesis with an online, single word recognition test in which a test word appeared at the end of the text (immediately after "mother" for the text above). There were two possible test words, "flower" and "display." For "flower," responses were faster with the "picking flower" text than with the "smelling flower" text, indicating that "flower" was indeed more salient when it was picked. Responses to "display" were also faster with the "picking" version than the "smelling" version, indicating that the "picking" version brought not only "flower" into increased salience but also information connected to it.

Glenberg et al. (1987) gave a different interpretation of the result for "flower." Our texts were modifications of the ones they used. For the "flower" text, the words "to an ornamental display" were words that we added. Glenberg et al. claimed that shorter RTs for "flower" were the result

of readers encoding a complete, real-life model of the situation described by the text, a model in which the flower would still be in possession of the girl at the end of the "picking" version but not the "smelling" version. However, this interpretation is ruled out by McKoon and Ratcliff's finding that "display" also had faster responses after the "picking" version.

Finally, we mention one more demonstration of the effect salience can have on the degree of match between a test word and textual information. Gerrig, Love, and McKoon (2009) hypothesized that small mysteries about an entity in a text could make the entity more salient. If, for example, a text mentioned a person named "Judy" but did not specify anything about her role in the narrative, then the identity of Judy would present a small mystery to the reader. Supporting Gerrig et al.'s suggestion, responses to "Judy" as a test word in an online, single-word recognition experiment were faster when her role was not explained than when it was ("the principal Judy"). In another chapter of this volume, Gerrig and Wenzel give a complete discussion of the importance of small mysteries like "Judy."

Retrieval context. One demonstration of the power of context was provided by McKoon and Ratcliff (1995) in a lexical decision experiment. They showed that the standard priming effect between highly associated words such as "close far" could be eliminated by context. Priming was observed when the other pairs in a list had the same relation, opposites, as "close far" (e.g., "broad narrow") but not when the other pairs had a different relation (e.g., "cold snow," "sour lemon," "blue sky"). We conclude from this that even such strong associates as "close far" can be overridden by context.

In discourse research, context effects are well-appreciated. For example, in the McKoon and Ratcliff (1986) study described above, the most important finding was a context effect – subjects tended to make more errors when the to-be-inferred word was tested with a prime from its sentence than when the prime was "ready."

A similar result was obtained for interactions of general knowledge with textual information (McKoon and Ratcliff, 1988). The sentences "The still life would require great accuracy. The painter searched many days to find the color most suited to use in the painting" are more about tomatoes being red than round. Full understanding of the text should include the "red" information but, as with the "dead" inference, we found that the "red" information was encoded only minimally. When test sentences that were consistent with the meaning of a text, like "tomatoes are red," were given for off-line true/false judgments, responses were facilitated only when the test sentence was immediately preceded in the test list by other information from the same text (e.g., "The still life would require great accuracy").

Summary. Overall, the general cognitive principles listed at the beginning of this chapter and their application to methodologies designed to investigate discourse processing have led to many intriguing findings that might otherwise have gone unnoticed. We have illustrated this here with findings that were surprising when they were introduced to the field, for example, pronoun resolution without pronouns, schema-related connections from one otherwise unrelated narrative to another, subtle effects of syntactic salience, and the context effects that are interpreted as reflecting minimal encoding.

New directions

Reading comprehension research such as that reviewed here has focused mainly on college students, but reading comprehension is a highly important issue for other populations. Older adults need to understand, for example, medical and legal information, and high-school dropouts need to understand the information on GED tests. The general question is whether and how discourse processing differs between these populations and college students. We first describe a method, a computational model, for comparing data between populations and then apply that method to inferences of the "actress-dead" kind. The studies we review compared older adults (sixty-five to seventy-five year olds) to college students (McKoon and Ratcliff, 2013).

Comparisons of performance between old and young face two crucial problems. One is that older adults often set more conservative speed/accuracy criteria than young adults, that is, they are more concerned to avoid errors. The second is that older and young adults have different baseline levels of performance: older adults' RTs tend to be considerably slower. These two differences mean that older adults' and young adults' performance cannot be directly compared. For example, suppose "robin" was presented for lexical decision. Older adults' accuracy might be better than young adults' because their lexical knowledge for "robin" is better, or their knowledge might be worse and their higher accuracy the result only of more conservative speed/accuracy criteria. For RTs, older adults might be slower than young adults because their knowledge is worse or because their criteria are more conservative.

When subjects make two-choice recognition decisions or lexical decisions, the question of interest is often whether the information on which older adults base their decisions (e.g., lexical information about "robin") is of the same, better, or worse quality than that of young adults. To answer this question requires a computational model that can separate out speed/accuracy differences from quality-of-information differences.

The model we have used is Ratcliff's diffusion model (1978; Ratcliff and McKoon, 2008). With the model, we have found that in some (although not all) memory and perceptual tasks, the quality of the older adults' information is as good as young adults'. The reason older adults are slower is usually due to more conservative speed/accuracy criteria (and also to slowdowns in processes outside those of interest, such as encoding a stimulus or executing a response). In the next section, we explain how the diffusion model separates information quality from speed/accuracy criteria.

The diffusion model

In the model, evidence about a stimulus accumulates over time from a starting point (z) to one or the other of two criterial amounts, or boundaries, one for each choice. The higher the quality of the evidence, the higher the rate at which it is accumulated. The rate of accumulation is called drift rate, v . Stimuli that differ in difficulty (e.g., in lexical decision, low-frequency versus high-frequency words) differ in drift rates. A response is executed when the amount of accumulated evidence reaches a boundary, either zero for a negative response or a for a positive response. The processes outside the decision process (e.g., encoding, response execution) are combined into a single parameter of the model that has mean duration T_{er} ms. Noise (within-trial variability) in the accumulation of evidence from the starting point to the boundaries results in processes with the same mean drift rate terminating at different times (producing RT distributions) and sometimes terminating at the wrong criterion (producing errors).

The values of drift rates, the nondecision component, and the boundaries are assumed to vary from trial to trial. The assumption of across-trial variability is required if participants cannot accurately set these parameters at the same values from trial to trial. Across-trial variability in drift rate is assumed to be normally distributed with SD η , across-trial variability in the nondecision component is assumed to be uniformly distributed with range st , and across-trial variability in the starting point is assumed to be uniformly distributed with range sz .

The diffusion model is designed to explain all the aspects of two-choice data – accuracy, mean correct and mean error RTs, the shapes and locations of RT distributions, and the relative speeds of correct and error responses. Explaining all of these data simultaneously puts powerful constraints on the model. The model also can reconcile seemingly contradictory results for accuracy and RTs. For example, item recognition data show large increases in RTs with age coupled with small

changes in accuracy or no changes in accuracy at all. The RT data suggest large decrements with age, whereas the accuracy data suggest only small decrements. The diffusion model reconciles these seemingly inconsistent results by mapping the two dependent variables onto the same underlying decision process.

Associations. To illustrate application of the diffusion model, we review a comparison (McKoon and Ratcliff, 2012) of two ways of measuring the strength of the association between two items in long-term memory. In McKoon and Ratcliff's study, older and young subjects were given pairs of words to learn. Sometimes the test items were single words presented for recognition and the strength of the association between the two words of a pair was measured by priming (one member of a pair immediately preceding the other). Other times, to test associative recognition, the test items were pairs of words and subjects were asked to decide whether the words had been studied in the same or different pairs (all the words in a pairs test list had been on the list of to-be-learned pairs). The degree to which same-pair responses were facilitated over different-pair responses was the measure of associative strength. This study was the first to address the question of whether priming in single-word recognition depends on the same information in memory as associative recognition.

Performance on associative recognition has usually been measured in terms of accuracy, and priming effects in item recognition have usually been measured in terms of RTs. This is because the effects of priming on accuracy tend to be small and the effects on RTs tend to be large, whereas for associative recognition, the same-versus different-pair effects are large in accuracy and small in RTs.

Accuracy and RTs for the single-word and pairs-recognition tasks cannot be directly compared because they are measured on different scales, accuracy on a probability-correct scale and RTs on a time scale. Neither RT nor accuracy can be used alone as the basis of a model of performance. A model built solely on accuracy data would almost surely be invalidated by RT data, and a model built solely on RTs would almost surely be invalidated by accuracy data. Using the diffusion model allows accuracy and RTs to be mapped onto the same metrics: drift rates, speed/accuracy criteria, and nondecision times. Another difference between the two tasks is that in single-word recognition, the correct response to both primed and unprimed test words is "yes," but in associative recognition, the correct response to same-pair tests is different from the correct response to different-pair tests. Just as with accuracy and RTs, the model allows the different responses to be measured on the same metric.

McKoon and Ratcliff's (2012) application of the diffusion model showed that associative recognition and priming in single-word recognition depend on the same information in memory and that this is true for both older and young subjects. This is a conclusion that could not have been drawn without the model. The findings were, first, that there were significant correlations between associative-recognition drift rates and priming drift rates, for both older and young adults; second, that drift rates for priming and drift rates for associative recognition tracked each other as a function of age; and third, that drift rates for priming and associative recognition tracked each other as a function of subjects' IQ. The significant correlations between drift rates for the two tasks show that subjects who do well on one of the tasks also do well on the other, which supports the hypothesis that they rely, at least to some extent, on the same information in memory. The correlations with IQ are what would be expected: higher IQ subjects remember information better.

In the discourse processing literature, same-different pair recognition has not been used. However, it offers a new way to measure the connections encoded between pieces of text information, and this measure can be compared to priming. The question is whether the two measures lead to the same conclusions about discourse processing, or different ones. If they are different, then empirical explorations of the differences may provide new, strong tests of what is going on in discourse processing.

Actress-dead inferences. To show application of the diffusion model to discourse processing research, we replicated McKoon and Ratcliff's 1986 experiment with actress-dead kinds of inferences and off-line single-word recognition, and we compared older adults' performance (ages sixty-five to ninety) to college students' (McKoon and Ratcliff, 2013). For the "dead" test words, we found that the students were considerably faster than the older subjects, but not more accurate. As previously described, with these data alone, there are several possible interpretations: The older subjects were slower because the quality of the evidence about "dead" on which they based their decisions was worse, because they set more conservative speed/accuracy criteria, because they had shorter nondesideration times, or some combination of these. They might have been equally accurate because the quality of their information was as good, because they set their criteria more conservatively, because their nondesideration processes were the same as young subjects', or some combination of these.

When we used the model to explain the data, there was a clear and compelling result: The quality of the information about the actress being dead (drift rate in the model) was just as good for the older adult's as for the young. In other words, they understood what would happen to the

actress just as well as the young adults. The older adults were slower only because they set their speed/accuracy criteria farther apart and their nondesideration times were longer.

Conclusion

All of the theories described in this chapter have led to interpretations of empirical methods and results that have defined discourse processing research since the early 1980s. It is from this foundation that we can ask new questions about what readers understand from discourse, questions such as the fascinating ones raised and illustrated by Gerrig and Wenzel in their chapter.

We especially appreciate the new issues that can be addressed by computational modeling. To our knowledge, the McKoon and Ratcliff (2013) experiment described here is the first application of a sequential sampling model to investigations of language comprehension for older adults. In a similar manner, we are currently extending the model to adults who are learning to read. We firmly believe that using the model will, in the near future, allow investigations of the degree to which many other populations understand and remember all sorts of textual information.

REFERENCES

- Allbritton, D. W., McKoon, G., & Gerrig, R. (1995). Metaphor-based schemas and text representations: making connections through conceptual metaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 612-25.
- Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge: Harvard University Press.
- Balota, D. A. (1990). The role of meaning in word recognition. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (eds.), *Comprehension Processes in Reading* (pp. 9-32). Mahwah, NJ: Erlbaum.
- Corbett, A. T., & Doshier, B. A. (1978). Instrument inferences in sentence encoding. *Journal of Verbal Learning and Verbal Behavior*, 17, 479-91.
- Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior*, 22, 121-32.
- Erickson, T. D., & Mattson, M. E. (1981). From words to meaning: a semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, 20, 540-51.
- Forster, K. I. (1981). Frequency blocking and lexical access: one mental lexicon or two? *Journal of Verbal Learning and Verbal Behavior*, 20, 190-203.

- Gerrig, R. J., Love, J., & McKoon, G. (2009). Waiting for Brandon: how readers respond to small mysteries. *Journal of Memory and Language*, 60, 144-53.
- Gerrig, R., & McKoon, G. (1998). The readiness is all: the functionality of memory-based text processing. Invited article, *Discourse Processes*, 26, 67-86.
- (2001). Memory processes and experiential continuity. *Psychological Science*, 12, 81-5.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1-67.
- Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, 26, 69-83.
- Greene, S. B., Gerrig, R. J., McKoon, G., & Ratcliff, R. (1994). Unheralded pronouns and the management of common ground. *Journal of Memory and Language*, 33, 511-26.
- Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 266-83.
- Lockhart, R. S., Craik, F. I. M., & Jacoby, L. (1976). In J. Brown (ed.), *Recall and Recognition*. London: Wiley.
- Love, J., & McKoon, G. (2011). Rules of engagement: incomplete and complete pronoun resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 874-87.
- McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: some consequences of memory based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 919-32.
- McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior*, 19, 668-82.
- (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 82-91.
- (1988). Contextually relevant aspects of meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 331-43.
- (1989a). Inferences about contextually defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1134-46.
- (1989b). Assessing the occurrence of elaborative inference with recognition: compatibility checking vs. compound cue theory. *Journal of Memory and Language*, 28, 547-63.
- (1992). Inference during reading. *Psychological Review*, 99, 440-66.
- (1994). Sentential context and online lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1239-43.
- (1995). Conceptual combinations and relational contexts in free association and in priming in lexical decision. *Psychonomic Bulletin & Review*, 2, 527-33.
- (2012). Aging and IQ effects on associative recognition and priming in item recognition. *Journal of Memory and Language*, 66, 416-37.
- (2013). Aging and predicting inferences: a diffusion model analysis. *Journal of Memory and Language*, 68, 240-54.

- McKoon, G., Ratcliff, R., & Seifert, C. (1989). Making the connection: generalized knowledge structures in story understanding. *Journal of Memory and Language*, 28, 711-34.
- McKoon, G., Ratcliff, R., Ward, G., & Sproat, R. (1993). Syntactic prominence effects on discourse processes. *Journal of Memory and Language*, 32, 593-607.
- Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18, 5-20.
- Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: effects of frequency of meaning and contextual bias. *Memory & Cognition*, 9, 225-36.
- Posner, M. I. (1978). *Chronometric Explorations of Mind*. Mahwah, NJ: Erlbaum.
- Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inference: lexical decision versus naming. *Journal of Memory and Language*, 27, 399-415.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, 95, 385-408.
- (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Computation*, 20, 873-922.
- Sanford, A. J., & Garrod, S. C. (1981). *Understanding Written Language*. New York: Wiley.
- Tulving, E. (1974). Cue-dependent forgetting. *American Scientist*, 62, 74-82.