

Processes of Question Answering

Murray Singer, University of Manitoba

In press, In G. Rickheit, T. Herrmann, & W. Deutsch (Eds.), Psycholinguistik-
Psycholinguistics. Berlin: Walter de Gruyter.

Outline

1. Introduction
2. Processing Stages of Question Answering
3. Parallel Processing In Question Answering
4. Conclusions

1. Introduction

People's ability to answer questions is a central criterion of successful comprehension and learning, arguably even more crucial than paraphrasing and summarization (Lehnert, 1978). Questions and their answers play a central role across discourse domains, including the retrieval of information from texts of different genres, the measurement of people's knowledge, and speakers' and listeners' assessments of one another's respective goals (e.g., Graesser & Murachver, 1985; Lehnert, 1978). In the development of theory in this realm, the similarity between answering Does a giraffe have a long neck? and recognizing or verifying A giraffe has a long neck has been noted (Graesser, Lang, & Roberts, 1991; Reder, 1982; Singer, 1990).

The complexity of question answering is revealed by the following excerpt from a mystery set in ancient Rome.

"I squatted by the body while Hermes held the lantern near the face. Sure enough, the man wore a tunic with a senator's wide, purple stripe. He was middle-aged, bald and beak nosed, none of which were distinctions of note. And he had at least one enemy, who had stabbed him through the heart." (Roberts, 1993)

After reading this passage, readers should readily be able to answer Who held the lantern? and What did the stabbed man wear? The answers are stated explicitly in the passage, and failure to answer them correctly would raise doubt about the reader's attention to or grasp of the story. However, the question answering capability of the reader is much greater. If asked, Who is describing the murder scene?, familiarity with the mystery genre would permit many readers to provide an answer such as "the detective." Likewise, most people would likely provide a sensible answer to What instrument was used to kill the man?, even though the weapon is not identified. If, in contrast, one were asked, Who murdered the man?, one would have to reply "don't know," in view of the sparse

information provided by the passage. How readers decide whether or not they can provide a sensible answer to a question is itself an important issue.

This chapter includes two main sections. Section 2 identifies, and presents evidence about, a series of processing stages that have been proposed to contribute to successful question answering. Section 3 provides an overview of parallel processing models that address the role of interactive processing and global memory search in question answering. First, however, an outline of current discourse comprehension theory is provided.

1.1. Discourse Comprehension and Representation

The present emphasis on answering questions about discourse requires an outline of contemporary theories of comprehension. According to the influential construction-integration theory (CI) of Kintsch (1988, 1998), comprehension proceeds in cycles. During the construction phase of a CI cycle, the current text segment is analyzed into idea units called "propositions." For example, I squatted by the body might be analyzed as: P1 (SQUAT, I), P2 (LOCATION:BY, P1, BODY) (Kintsch, 1974). The propositions are organized into a coherence network that also includes a small number of (a) coherence-preserving inferences (Haviland &

Clark, 1974) and (b) close associates of the explicit text ideas (Swinney, 1979). The network specifies the strengths of the connections among all of these components. Then, during the integration phase, activation is settled in the network, following connectionist principles (Rumelhart & McClelland, 1986). As a result, only those components that are highly interconnected effectively remain in the network. These processes modify the original network to yield a long-term memory (LTM) representation of the text. At the end of each cycle, a small number of highly active propositions are retained in working memory (e.g., Baddeley, 1986) for further processing.

The LTM representation captures multiple levels of discourse, including its surface features, a textbase network of its propositions, and a model of the situation to which the discourse refers (Schmalhofer & Glavanov, 1986; van Dijk & Kintsch, 1983; Zwaan & Radvansky, 1998). The situation model integrates discourse information and general knowledge, and may take a form very distinct from the original text.

2. Processing Stages of Question Answering

Theorists have identified and investigated numerous processing stages of question

processing. This theoretical approach has guided the inspection of many phenomena of text retrieval and question answering.

2.1. Question Encoding

Answering a question first requires the encoding of its meaning. The predominant treatment of this problem involves the propositional analysis of the question statement. Using Kintsch's (1974) notation, the proposition underlying Hermes held the lantern is (HOLD, AGENT:HERMES, PATIENT:LANTERN), in which AGENT and PATIENT name the semantic roles (Fillmore, 1968) of the arguments in relation to the predicate. The propositional representation of the question, Who held the lantern?, can be written as (HOLD, AGENT:?, PATIENT:LANTERN) (Singer, 1990). The notation, AGENT:?, is meant to convey that the questioner wishes to know the identity of the agent of the proposition.

Identifying the element that the questioner wishes to know about (e.g., AGENT:?) is a critical feature of question comprehension. The simple active sentence, Hermes held the lantern, is relatively neutral in its "given-new structure" (Clark & Haviland, 1977): The speaker might either have (a) assumed that the listener knew that Hermes held something (given), and identified the thing as a lantern (new); or (b) assumed that the listener knew

that someone held the lantern, and identified the holder as Hermes. Successful question comprehension and answering, in contrast, depend on the clear distinction between the given, or presupposed, information; and the new, or focal, idea. Wh- questions are unambiguous in this regard. It is clear that Who held the lantern? presupposes that someone held the lantern, and requests the identity of the holder. Many yes-no questions, such as Was the lantern held by Hermes?, also clearly signal which element the speaker is querying. The notation (HOLD, AGENT:HERMES?, PATIENT:LANTERN) conveys that HERMES is the queried concept.

However, identifying the question focus also involves the relevant knowledge possessed by the listener. Why did Susan phone the Prime Minister? appears to focus on Prime Minister; whereas a question with similar grammatical structure, Why did Susan yodel to the dog?, seems to interrogate the role of yodelling. Lehnert (1978) observed that people interpret questions to query their unusual elements (Prime Minister, yodel), which in turn depends on the listener's knowledge.

2.2. Question Categories

Many researchers have compiled inventories of question categories (Graesser & Murachver, 1985; Lehnert, 1978; Trabasso, Secco, & van den Broek,

1984). Some question categories can be linked to analyses of semantic roles and cases (Fillmore, 1968), such as agents, patients, instruments, and locations. Thus, one can ask Who held the lantern?, How was the man stabbed?, and Where was the man stabbed? Other question categories correspond to the discourse relations of causality and temporal organization (Graesser & Murachver, 1985; Trabasso et al., 1984). One can ask about the causal antecedents (Why was the man stabbed?) and consequences of a discourse event and about its significance.

Both wh- and yes-no questions can be formulated about any of these categories (Singer, 1986, 1990). For example, one can ask either What did the man wear? or Did the man wear a tunic?; and either Why was the man stabbed? or Was the man stabbed to prevent him from going to the authorities?. A central difference between the processing of wh- and yes-no questions is that only yes-no questions involve the comparison of the retrieved information and the focal information of the question (Singer, 1984, 1986). These comparison operations will be examined in section 2.5.

Interrogative pronouns and adverbs provide cues about the category of a question. Unambiguous signals of question category are provided by interrogative terms such as who

(agent) and where (location). In other instances, the relation between the interrogative term and question category is more complex. For example, How can introduce the questions How did Bill open the window? and How many oranges did you buy? In general, question category must often be derived from a combination of the interrogative term, syntax, and semantics.

Question categorization has a direct bearing on the processes of searching memory for the queried information. These effects will be considered in the section 2.4.

2.3. Strategies of Question Answering

Strategy selection has been proposed to form a distinct processing stage of question answering and text retrieval (Graesser & Murachver, 1985; Reder, 1987; Singer, 1990). People can evaluate a question by directly retrieving it (Reder, 1982) from a specific knowledge representation; or by judging its plausibility (Camp, Lachman, & Lachman, 1980; Lehnert, 1977; Reder, 1982, 1987, 1988; Singer, 1991a). In the context of the mystery excerpt (section 1.), consider the question, Was the man killed with a dagger? One might either attempt to retrieve the answer from the representation of the passage, which in this case would fail; or to judge its probability with reference to

one's world knowledge about stab wounds.

Answering strategy is guided by factors intrinsic and extrinsic to the question statement (Reder, 1987). The prototype extrinsic factor is the answering task that one adopts, perhaps as a result explicit instructions. Intrinsic factors are illustrated by the relative activation of question statement at the time it is encountered, due either to its high familiarity or a recent presentation.

The effects of several variables on answering strategy were explored by Reder (1982). The subjects read stories; and, in a subsequent test, were instructed either to recognize the test items or judge their plausibility. The test occurred 0 minutes, 20 minutes, or 2 days after reading. The test statements were either of high or medium plausibility with reference to their stories. For example, in the context of the mystery excerpt, The man was stabbed with a dagger might be of high plausibility whereas The man was stabbed with a letter opener might be of medium plausibility. Finally, each test item was either stated or unstated in the story.

Reder (1982) reasoned that direct retrieval ought to be more efficient than plausibility judgment only at short delays, when the verbatim details of the passage are still relatively intact. Consistent with this proposal, correct

recognition took less time than plausibility judgments in immediate testing, but the opposite was true after a two-day delay (see also Singer, 1979). Other features of the data likewise suggested that, independent of the instructions, people favour a retrieval strategy at short delays and plausibility at long delays. In the plausibility instruction condition, for example, judgment time should be the same for stated and unstated test items, on the assumption that the test item is evaluated only with reference to relevant knowledge. However, in immediate testing only, plausibility judgment time was .93 sec less for stated than unstated test items. This indicated that, contrary to their instructions, the plausibility judges had retrieved the stated test items (see also Reder, 1987).

Associations have sometimes been drawn between (a) direct retrieval and the search of a message representation, and (b) plausibility judgment and evaluation on the basis of pertinent world knowledge (e.g., Lehnert, 1977, p. 57; Yekovich & Walker, 1986). This relationship probably holds for question answering about stories that refer to stereotypical situations, such as going skiing or doing the laundry (Yekovich & Walker, 1986). However, there is evidence that either strategy

can be applied to either type of knowledge representation (Singer, 1991a; see also Lehnert, 1978; Lorch, 1981).

2.4. Memory Search in Question Answering

2.4.1. Fact retrieval: The fan effect.

After the encoding and categorization of the question and strategy selection, memory may be searched for the requested information.

Important clues about the character of these processes stem from the well-known fanning paradigm (Anderson, 1974, 1976). In this procedure, subjects memorize a list of facts, illustrated by set (1):

- (1) a. A pilot is in the garage.
- b. A doctor is in the store.
- c. A doctor is in the bank.
- d. A tailor is in the store.

Recognition time for test items varies with the total number of facts in which the concepts of the test item participated. The test item A pilot is in the garage is a 1-1 fact, because pilot and garage each participated in exactly one fact (1a). Likewise, according to this scheme, A doctor is in the bank is a 2-1 test item, and A doctor is in the store is a 2-2 test item. Of these three test items, recognition time is least for (1a) and greatest for (1b) (Anderson, 1974). This outcome is called the fan effect because answer time is regulated by

the number of links fanning out from each concept in a network that interrelates the facts. One explanation of fan effect takes the form of the representational and processing assumptions of the ACT model (Anderson, 1976, 1983, 1993), which will be examined in section 3.1.

2.4.2. Question answering, the fan effect, and focused memory search. Anderson's (1976) fact retrieval procedure was extended to question answering (Singer, Parbery, & Jakobson, 1988). The subjects memorized complex facts such as (2a), and, one day later, answered questions such as (2b) and (2c).

(2) a. The teacher watered the peas, the corn, and the lettuce with the hose.

b. Did the teacher water some corn?
(patient focus)

c. Did the teacher use a hose?
(instrument focus)

In one experiment, each question presupposed the participation of the agent (e.g., teacher) and focused on either on the patient or instrument case. Each question could then be categorized in terms of the number of concepts that had been learned in the queried case and the other non-presupposed case. Following this scheme, (2b) is a 3-1 question: It asks about the accuracy of the patient, corn, and fact (2a)

included three patients and one instrument. Question (2c), in contrast, was a 1-3 item. Singer et al. (1988) reported systematically lower answering times for 1-3 questions than 3-1 questions, even though both types involved exactly four related concepts. This indicated that people can focus their memory search on the queried case, such as the instrument in (2c). These results were similar to focused-search results in the realm of questions about taxonomic categories (McCloskey & Bigler, 1980) and theme-related actions (Reder & Anderson, 1980).

2.4.3. Strategies of fact retrieval. In section 2.3, it was proposed that question statements may be evaluated by retrieving them from a knowledge representation (be it the discourse textbase or general knowledge) or by a plausibility judgment. The extrinsic and intrinsic factors that guide the choice of answering strategy exert parallel effects in fact retrieval. For example, Reder and Anderson's (1980) subjects learned sets of theme-related facts such as (3).

(3) a. Alan bought a ticket for the 10:00 A.M. train.

b. Alan heard the conductor call, "All aboard."

c. Alan arrived on time at Grand

Central station.

In an inspection of the extrinsic factor of test composition, some test lists included distractor items that were related to the learned sets (e.g., Alan checked the railway schedule); and, in other lists, all of the distractor items were about unrelated themes (e.g., Alan lifted car with the jack). Reder and Anderson reasoned that, in the presence of related distractors, a correct decision depends on the direct retrieval of the test item, which permits its content to be scrutinized. This would result in the usual fan effect. With unrelated distractors, in contrast, it is not necessary to compare the test fact with all of the facts learned about the character--rather, one can accept a test fact as long as it is consistent with a theme associated with that character. Consistent with this analysis, a significant fan effect was measured only in the presence of related distractors. Other results confirmed that people perform plausibility judgments in the fanning paradigm when instructed to do so (Reder & Ross, 1983) and at long test delays (Reder & Wible, 1984). These outcomes converge with the findings about answering strategies.

2.4.4. Question search procedures and focused memory search. Section 2.4.2. presented evidence

that people can use information about the type of link between concepts to focus their memory search on a queried category. There is another research tradition which has examined people's versatile use of both the type and direction of conceptual links to answer questions (e.g., Goldman, 1985; Graesser & Murachver, 1985; Graesser et al., 1991; Graesser, Robertson, & Anderson, 1981; Trabasso et al., 1984). This analysis has emphasized text statement categories, such as the physical events and actions of stories; plus directional links that connect those categories, such as links of reason, consequence, and enablement (Graesser, 1981; Schank & Abelson, 1977). Consider sequence (4):

- (4) a. The burglar climbed the drainpipe.
- b. He pried the window with a crowbar.
- c. The lock broke.
- d. The burglar opened the window.
- e. He entered the house.

Of these statements, (4c) is an event and the others are actions. Sentence (4a) enables (4b) (a forward enablement link), (4b) causes (4c), and (4c) enables (4d). However, (4b) is also directly connected to (4d) by a forward reason link--the burglar pried the window in order to open it. Likewise, (4e) is the reason for (4d) (Graesser & Murachver, 1985).

Graesser proposed that each combination of an interrogative term plus a story statement category is associated with a distinct question search procedure (Graesser & Clark, 1985; Graesser & Murachver, 1985; Graesser et al., 1981). Combining the interrogative why with a story event (e.g., Why did the lock break?) asks about the cause of the story statement. Why-action questions, in contrast, query a character's reason for performing the action. The search procedure for a combination corresponds to the tracing of a particular link type in a particular direction. Why-action questions require the tracing of a Reason link in the forward direction. Thus, Why did the burglar pry the window? is reasonably answered, "To open it." Conversely, how-action questions require the tracing of a Reason link in the backward direction. How did the burglar open the window? can be answered, "He pried it." A corollary is that satisfactory answers may be derived from a combination of legal paths. Thus, Why did the burglar pry the window? can be answered either "To open the window" (one forward Reason link) or "To enter the house," (two forward Reason links) (Goldman, 1985; Graesser & Murachver, 1985).

This analysis was supported by the results that a high proportion of people's answers to

questions are consistent with the proposed question search procedures, and that people give higher "goodness-of-answer" ratings to those answers that conform with the analysis than to those that do not (Graesser & Murachver, 1985; Graesser et al., 1981). This proposal also addresses an issue raised in section 2.2.-- namely, that interrogative terms do not bear a one-to-one relationship with question categories.

2.4.5. Feeling-of-knowing and question search processes.

The decisions of whether and how long to search for a queried fact are governed by the familiarity of the question. As a result, people will execute a lengthy search for a question involving familiar concepts, such as What is the capital of Australia?, but will answer "don't know" rapidly for the unfamiliar question, What is the capital of Zaire?

(Collins, Brown, & Larkin, 1980; Costermans, Lories, & Ansay, 1992; Glucksberg & McCloskey, 1981; Nelson & Narens, 1980; Reder, 1987). A person's assessment of question familiarity has been called the feeling-of-knowing (Hart, 1967).

In her analysis of answering strategies, Reder (1987) proposed that the feeling-of-knowing of a question statement is a function of its familiarity and the recency with which it was encountered. She characterized both

familiarity and recency as intrinsic factors that influence the duration of question search. In one experiment, Reder "primed" some questions by preexposing their concepts to the subjects, in a preexperimental task. Then, some of the subjects were instructed to answer each question; whereas others were told simply to estimate, as quickly as possible, whether they believed that they could answer the question. The estimators were influenced by the priming manipulation--their estimation difference between easy and difficult questions was smaller for primed than unprimed items. The "answerers," in contrast, were not affected in this way by the priming manipulation.

2.5. Question Comparison

Answering the yes-no question, Did the man wear a tunic?, depends on the comparison of the focal question concept, tunic, and the retrieved information. Question comparison processes have been studied with reference to verifying sentences like An elephant is not small. Clark and Chase (1972) proposed that this test sentence is propositionally encoded as NOT(SMALL, ELEPHANT) and is compared with the corresponding general knowledge (LARGE, ELEPHANT). The modifier NOT of NOT(SMALL, ELEPHANT) mismatches the implicitly affirmative modification of (LARGE, ELEPHANT); and the

concepts SMALL and LARGE also mismatch. The first mismatch results in the change of a response index from its initial value "yes," to "no"; and the second mismatch causes the response to revert back to "yes," the correct answer.

This analysis was extended to question answering about discourse (Singer, 1984, 1986). Consider sentence set (5):

(5) a. The doctor ate the chicken on the patio.

b. The doctor ate the fish on the patio.

c. The doctor ate on the patio.

d. Did the doctor eat some chicken?

Question (5d) can reasonably be answered "yes," "no," and "don't know" in the contexts of (5a), (5b), and (5c), respectively. Singer (1984, 1986) proposed that question (5d) is encoded as (EAT, AGENT:DOCTOR, PATIENT:CHICKEN?). Search then focuses on the queried case, the patient (McCloskey & Bigler, 1980; Singer et al., 1988). A preliminary search (Reder, 1987) reveals whether the antecedent representation (5a, 5b, or 5c) includes any information in the queried case. Sentences (5a) and (5b) include patients, which then must be retrieved and compared with the focal question component. When (5d) follows (5c), in contrast,

the preliminary search reveals (5c) to include no information in the patient case. This permits the response index to be immediately changed to "don't know." Consistent with this hypothesis, answer times are consistently faster in the "don't know" than the "no" condition (Singer, 1984, 1986).

3. Parallel Processing In Question Answering

The stage analysis of question answering is justified because some answering processes, such as retrieval, arguably must precede others, such as comparison. As a research strategy, stage analysis has permitted the clarification of numerous answering processes. However, a strictly serial analysis of question answering is inconsistent with the emphasis, in contemporary cognitive theory, on parallel processing (e.g., Rumelhart & McClelland, 1986). In this regard, Robertson, Weber, and Ullman (1993) reasoned that the appearance of an interrogative word at the beginning of a sentence should initiate retrieval of the queried concept in parallel with the encoding of the sentence. In fact, they measured longer reading times for sentences beginning with interrogatives than for control declarative sentences (see also Graesser et al., 1991). Therefore, two theoretical models that accommodate the parallel processes of question

answering will next be outlined and compared. The merits of such models will then be identified.

3.1. Parallel Processing Answering Models

The TEXTR model of text retrieval (Kintsch, 1998; Singer & Kintsch, 2000) combines existing successful theories of (a) text comprehension, namely the construction-integration (CI) theory (Kintsch, 1988, 1998) and (b) recognition, namely SAM (Gillund & Shiffrin, 1984). As described in 1.1., the CI model addresses the construction of a discourse representation in long-term memory (LTM). According to TEXTR, questions are initially processed in the same manner as the rest of the text (Lehnert, 1978). As a result, the question itself modifies the LTM discourse representation and becomes a part of it. Then, the familiarity of the question is assessed using the computations of the SAM recognition model. Because the question has been integrated with the discourse representation, the familiarity computation takes into account the connection strength between the question statement and every other element in the representation. Comparing a test item with the entire contents of memory is called global matching, and is characteristic of modern theories of memory (Clark & Gronlund, 1996). Finally, familiarity is converted to a

yes-no decision by comparing it to a response criterion. Simulation analyses revealed that TEXTR can predict complex patterns of people's "yes" replies to questions as a function of retrieval variables such as assigned task (recognition, verification), test delay, question type (explicit, paraphrase, implicit), and the importance, in the text, of the questioned idea (Kintsch, 1998; Singer & Kintsch, 2000).

Consider next the application of the well-known ACT theory (Anderson, 1976, 1983, 1993; Anderson & Reder, 1999b) to question answering and fact retrieval. According to ACT, the activation of a test fact, such as A lawyer is in the park, is derived from its resting activation level plus the sum of the strengths of its links to its concepts (lawyer, park). These link strengths are determined by a person's exposure to the facts in learning, according to a connectionist learning rule (Anderson & Reder, 1999b). In particular, the more facts that a concept has participated in, the weaker its links to those facts. At test time, one category of link (e.g., person) may be more heavily weighted than another (e.g., location) (Anderson & Reder, 1999b). Fact activations are ultimately converted to predicted judgment times using an exponential

formula--greater activation results in faster judgments. The ACT model can account for (a) the basic fan effect (section 2.4.1), (b) the similar magnitude of the fan effect for target and distractor items (Anderson, 1974, 1976), and (c) numerous other fact retrieval phenomena (Anderson & Reder, 1999b).

Parallel processing occurs in TEXTR when activation is settled in the coherence network during the integration phase of comprehension. Then, a question accumulates familiarity in parallel from all elements of the LTM discourse representation. In ACT, activation likewise spreads in parallel from all of the concepts in a text fact, until a fact is retrieved from memory.

3.2. Comparing the Models

There appear to be several fundamental differences between TEXTR and ACT. However, they may reflect differences in the current implementations of the models rather than their inherent properties.

3.2.1. Memory matching. TEXTR uses global memory matching (section 3.1.): Its familiarity values are based on the connection strengths between a question and all of the elements of the discourse representation. In contrast, the activation of a fact in ACT is derived only from its links to its component facts (Anderson &

Reder, 1999b), a local memory match (Clark & Gronlund, 1996). However, ACT could readily be modified to incorporate global matching. If two unrelated facts, such as A lawyer is in the park and A teacher is in the church, appeared consecutively during learning, it is plausible that each would have links of at least modest strength to the other's concepts. The computation of ACT activation could then incorporate these extra-fact connections.

3.2.2. Learning. As discussed in 3.1., ACT connection strengths are derived using a connectionist learning rule. The situation seems quite different in TEXTR, because connection strengths in the coherence networks are fixed by the experimenter. This is consistent with the practice in construction-integration modelling (e.g., Kintsch et al., 1990). However, in TEXTR, question activation is governed not by the initial coherence network but by the resulting LTM discourse representation. The connection strengths in the TEXTR LTM representation are determined by the reader's experience, such as the number of cycles during which a given proposition is processed. As in ACT, this experience is captured by connectionist learning rules (Kintsch & Welsch, 1991).

3.2.3. Representations. TEXTR is predicated

on a multilevel discourse representation whereas, at least for the fanning paradigm, ACT posits a simple, well-defined representation. A recent debate about ACT's representations focused on the finding that when people learned facts linking inanimate objects and locations, such as The desk is in the lobby, a fact retrieval test revealed that the number of objects appreciably influenced judgment time but the number of locations did not (Radvansky, Spieler, & Zacks, 1993). Radvansky (1999; Radvansky et al., 1993) attributed this differential fan effect to the representations that were constructed during comprehension: He proposed that when a place (lobby) was associated with several objects, the subjects represented a location-based situation model, in which the objects occurred in a single location (the lobby). In contrast, an object name (e.g., desk) associated with three locations is suggestive of three distinct desks. The differential fan effect was attributed to this difference in the representations.

Anderson and Reder (1999b), in contrast, denied the necessity of positing situational representations: Rather, they analyzed these results in terms of differential weightings, within the ACT model, for object links and location links (see 3.1.). However, there is no

reason, in principle, that ACT processes could not be applied to representations quite different from the typical fanning network (e.g., Anderson & Reder, 1999b, Figure 1). Even for fact lists, it is plausible that people would derive different representations depending on whether or not the list suggested that each instance of an object referred to the same entity.

3.3. Significance of the Parallel Processing Analyses

Parallel processing analyses of question answering have several benefits. First, they can address processing interactions among the stages of question answering, defined in terms of the mutual influence of higher and lower level processes (Rumelhart & McClelland, 1986). Empirical evidence reveals numerous interactions of this sort. For example, memory retrieval has been shown to occur concurrently with question encoding (Robertson et al., 1993). Likewise, when people answered intermixed questions about recent and earlier stories, there was evidence that different strategies were applied to the immediate and delayed questions (Reder, 1988; Singer, Gagnon, & Richards, 2000). This strongly favours the mutual influence of strategy selection and memory search upon one another over the serial execution of strategy

selection and then memory search.

Second, these parallel processing models have the capacity to accommodate additional variables in this realm. Consider the impact of test composition on question strategies. Test sets that include (a) distractors that are related to the original items (Lorch, 1981; Reder & Anderson, 1980; Singer, 1991b) or (b) many test items that appeared verbatim among the original stimuli (Reder, 1987) favour the answering strategy of direct retrieval over plausibility judgment. Test composition could be incorporated to the computations of both TEXTR and ACT by comparing strong (related) and weak (unrelated) connection strengths between the learned material (a text or fact set) and the distractor items (Anderson & Reder, 1999a; Gillund & Shiffrin, 1984; Reder & Anderson, 1980).

As a corollary, parallel processing models also offer alternative ways to address a given variable. In TEXTR, for example, differences between recognition and verification answering tasks can be treated either as a difference (a) between the response criterion associated with the tasks (e.g., Kamas, Reder, & Ayers, 1996; Miller & Wolford, 1995; Singer & Kintsch, 2000) or (b) in the relative weighting assigned to the surface, textbase, and situational discourse

representations (Hasher & Griffin, 1978).

4. Conclusions

Advances in the study of question answering have stemmed from a combination of serial and parallel processing theories. Serial analyses have identified and clarified the stages of question answering. Parallel processing analyses have incorporated fundamental cognitive principles such as interactive processing and global memory search. Progress in this field will likely continue to accumulate from a blend of these two approaches. At the same time, many other subtleties of question answering need to be addressed. (a) For example, in section 2.1, it was explained that people interpret the unusual element of a question, such as Why did Susan YODEL to the dog?, as the queried concept. (b) People's answers depend on their perception of the knowledge that they share with the questioner. Thus, one's answer to Where is the Empire State Building? will be different if the question is posed in Paris or in mid-town Manhattan (Norman, 1973). (c) People's replies are likewise affected by pragmatic factors such as politeness (Clark & Schunk, 1980). To reply with a simple "no" to Are there any interesting sights in this city? would be rather blunt. Phenomena of this sort offer major challenges in the refinement of question answering theory.

Murray Singer, University of Manitoba, Winnipeg

References

- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. Cognitive Psychology, 4, 451-474.
- Anderson, J. R. (1976). Language, memory, & thought. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard.
- Anderson, J. R. (1993). Rules of the mind. Hillsdale, NJ: Erlbaum.
- Anderson, J. R., & Reder, L. M. (1999a). Process, not representation: Reply to Radvansky (1999). Journal of Experimental Psychology: General, 128, 207-210.
- Anderson, J. R., & Reder, L. M. (1999b). The fan effect: New results and new theories. Journal of Experimental Psychology: General, 128, 186-197.
- Baddeley, A. D. (1986). Working memory. Oxford: Clarendon Press.
- Camp, C. J., Lachman, J. L., & Lachman, R. (1980). Evidence for direct-access and inferential retrieval in question-answering. Journal of Verbal Learning and Verbal Behavior, 19, 583-596.
- Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. Cognitive Psychology, 3, 472-517.
- Clark, H. H., & Clark, E. V. (1977). Psychology and language. New York: Harcourt Brace Jovanovich.
- Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. Freedle (Ed.), Discourse production and comprehension. Hillsdale, NJ: Erlbaum.
- Clark, H. H., & Schunk, D. (1980). Polite responses to polite requests. Cognition, 8, 111-143.
- Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. Psychonomic Bulletin and Review, 3, 37-60.
- Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference understanding. In R. Spiro, B. Bruce, & B. Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, NJ: Erlbaum.
- Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 55-72.
- Fillmore, C. J. (1968). The case for case. In E. Bach & R. T. Harms (Eds.), Universals of linguistic theory. New York: Holt, Rinehart and Winston.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both

- recognition and recall. Psychological Review, 91, 1-67.
- Glucksberg, S., & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. Journal of Experimental Psychology: Human Learning and Memory, 7, 311-325.
- Goldman, S. R. (1985). Inferential reasoning in and about narrative texts. In A. Graesser & J. Black (Eds.), The psychology of questions. Hillsdale, NJ: Erlbaum.
- Graesser, A. C. (1981). Prose comprehension beyond the word. New York: Springer-Verlag.
- Graesser, A. C., & Clark, L. F. (1985). The structures and procedures of implicit knowledge. Norwood, NJ: Ablex.
- Graesser, A. C., Lang, K. L., & Roberts, R. M. (1991). Question answering in the context of stories. Journal of Experimental Psychology: General, 120, 254-277.
- Graesser, A. C., & Murachver, T. (1985). Symbolic procedures of question answering. In A. Graesser & J. Black, The psychology of questions (pp. 15-88). Hillsdale, NJ: Erlbaum.
- Graesser, A. C., Robertson, S. P., & Anderson, P. A. (1981). Incorporating inferences in narrative representations: A study of how and why. Cognitive Psychology, 13, 1-26.
- Hart, J. T. (1967). Memory and the memory-monitoring process. Journal of Verbal Learning and Verbal Behavior, 6, 685-691.
- Hasher, L., & Griffin, M. (1978). Reconstructive and reproductive processes in memory. Journal of Experimental Psychology: Human Learning and Memory, 4, 318-330.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 13, 512-521.
- Kamas, E. N., Reder, L. M., & Ayers, M. S. (1996). Partial matching in the Moses illusion: Response bias not sensitivity. Memory & Cognition, 24, 687-699.
- Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, NJ: Erlbaum.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95, 163-182.
- Kintsch, W. (1998). Comprehension. New York: Cambridge University Press.
- Kintsch, W., & Welsch, D. M. (1991). The construction-integration model: A framework for studying memory for text. In W. E. Hockley & S. Lewandowsky (Eds.), Relating theory and data: Essays on human memory in honor of Bennet B. Murdock (pp. 367-385). Hillsdale, NJ: Erlbaum.
- Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. Journal of Memory and Language, 29, 133-159.
- Lehnert, W. (1977). Human computational question answering. Cognitive

- Science, 1, 47-73.
- Lehnert, W. G. (1978). The process of question answering. Hillsdale, NJ: Erlbaum.
- Lorch, R. F. Jr. (1981). Effects of relation strength and semantic overlap on retrieval and comparison processes during sentence verification. Journal of Verbal Learning and Verbal Behavior, 20, 593-611.
- McCloskey, M. & Bigler, K. (1980). Focused memory search in fact retrieval. Memory & Cognition, 8, 253-264.
- Miller, M. B., & Wolford, G. L. (1999) Theoretical commentary: The role of criterion shift in false memory. Psychological Review, 106, 398-405.
- Nelson, T. O., & Narens, L. (1980). Norms of 300 general information questions: Accuracy of recall, latency of recall, and feeling of knowing ratings. Journal of Verbal Learning and Verbal Behavior, 19, 338-368.
- Norman, D. A. (1973). Memory, knowledge, and the answering of questions. In R. Solso (Ed.), Contemporary Issues in Cognitive Psychology: The Loyola Symposium. Washington DC: V. H. Winston & Sons.
- Radvansky, G. A. (1999). The fan effect: A tale of two theories. Journal of Experimental Psychology: General, 128, 198-206.
- Radvansky, G. A., Spieler, D. H., & Zacks, R. T. (1993). Mental model organization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 55-72.
- Reder, L. M. (1982). Plausibility judgments versus fact retrieval: Alternative strategies for sentence verification. Psychological Review, 89, 250-280.
- Reder, L. M. (1987). Strategy-selection in question answering. Cognitive Psychology, 19, 90-134.
- Reder, L. M. (1988). Strategic control of retrieval strategies. In G. Bower (Ed.), The psychology of learning and motivation, Vol. 22, (pp. 227-259). San Diego: Academic Press.
- Reder, L. M., & Anderson, J. R. (1980). A partial resolution of the paradox of interference: The role of integrating knowledge. Cognitive Psychology, 12, 447-472.
- Reder, L. M., & Ross, B. H. (1983). Integrated knowledge in different tasks: The role of retrieval strategy on fan effects. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 55-72.
- Reder, L. M., & Wible, C. (1984). Strategy use in question answering: Memory strength and task constraints on fan effects. Memory & Cognition, 12, 411-419.
- Robertson, S. P., Weber, K., & Ullman, J. (1993). Parallel question parsing and memory retrieval. Journal of Memory and Language, 32, 155-168.
- Rumelhart, D. E., & McClelland (1986). Parallel distributed processing: Explorations in the microstructure of cognition, (Vol. I). Cambridge, MA: Bradford Books.

- Schank, R. C., & Abelson, R. (1977). Scripts, plans, goals, and understanding. Hillsdale, NJ: Erlbaum.
- Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situation representations. Journal of Memory and Language, 25, 279-294.
- Singer, M. (1979). Temporal locus of inference in the comprehension of brief passages: Recognizing and verifying implications about instruments. Perceptual and Motor Skills, 49, 539-550.
- Singer, M. (1984). Toward a model of question answering: Yes-no questions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 285-297.
- Singer, M. (1986). Answering wh- questions about sentences and text. Journal of Memory and Language, 25, 238-254.
- Singer, M. (1990). Answering questions about discourse. Discourse Processes, 13, 261-277.
- Singer, M. (1991a). Independence of question answering strategy and searched representation. Memory & Cognition, 19, 189-196.
- Singer, M. (1991b). Question answering strategies and conceptual knowledge. Bulletin of the Psychonomic Society, 29, 143-146.
- Singer, M., Gagnon, N., & Richards, E. (2000). Question answering strategy: The effect of mixing test delays. Manuscript in preparation.
- Singer, M., & Kintsch, W. (2000). Text retrieval: A theoretical exploration. Manuscript submitted for publication.
- Singer, M., & Jakobson, S. J. (1989). Focused search of semantic cases: The effects of question form and case status. Memory & Cognition, 17, 265-273.
- Singer, M., Parbery, G., & Jakobson, L. S. (1988). Focused search of semantic cases in question answering. Memory & Cognition, 16, 147-157.
- Swinney, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. Journal of Verbal Learning and Verbal Behavior, 18, 645-659.
- Trabasso, T., Secco, T., & van den Broek, P. (1984). Causal cohesion and story coherence. In H. Mandl, N. Stein, & T. Trabasso (Eds.), Learning and comprehension of text (pp. 83-111). Hillsdale, NJ: Erlbaum.
- van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.
- Yekovich, F. R., & Walker, C. H. (1986). Retrieval of scripted concepts. Journal of Memory and Language, 25, 627-644.
- Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123, 162-185.