Processing category terms in context:
Instantiations as inferences

PAUL WHITNEY
University of Arkansas at Little Rock, Little Rock, Arkansas

Three experiments were conducted to isolate conditions in which readers make their representations of text more concrete by instantiating (inferring a specific instance of) category terms. In previous research using single-sentence contexts and a modified Stroop task as an on-line measure of comprehension, no evidence for instantiation was obtained. In the present experiments, the modified Stroop task was used to test for instantiation when category terms were referred to anaphorically in subsequent sentences and when the category terms served as the subjects of single-sentence contexts. The pattern of the Stroop effects was dependent upon whether typical or atypical exemplars were suggested by the context. Facilitation of color naming was obtained in the former case, and interference was obtained in the latter case when the instances followed the appropriate biasing contexts. The results highlight the role of foregrounding and context in guiding the use of elaborative inferences.

During the past decade of research on comprehension processes, the notion that the mental representation of text incorporates aspects of the knowledge base of the reader has been particularly influential. The basic premise is that readers actively construct a mental representation that goes beyond the propositional structure of text. However, most of the studies of constructive processing have focused on memory for text; the results of these studies are ambiguous with regard to the use of constructive processes at encoding, because general knowledge can be used at retrieval to supply details that were not originally encoded (see Garnham, 1982). As Spiro (1980) pointed out, there is fairly wide agreement that the mental representation of text involves integration with prior knowledge, but relatively little is known about the initiation and use of constructive processes at encoding. It was toward an understanding of the initiation and use of constructive processes that the present experiments were directed.

Basing their theories on the notion that comprehension is a constructive process, several theorists have proposed that people use sentence contexts and their world knowledge to create a mental model or scenario that represents the events described in a text (Anderson & Shifrin, 1980; Bransford & Johnson, 1973; Johnson-Laird, 1983; Murphy, 1984; Sanford & Garrod, 1981).

A particularly interesting example of this position is the instantiation hypothesis proposed by Anderson and his colleagues (Anderson et al., 1976; Anderson & Shifrin, 1980). The instantiation hypothesis states that, whenever possible, readers use context to encode general terms as particular examples. The evidence for instantiation came from cued-recall studies in which, for example, subjects were better able to recall the statement, "The fish attacked the swimmer," when given *shark* as a cue than when given *fish*. Anderson et al. (1976) also obtained evidence that they interpreted as conflicting with the widely held notion that there are instances of natural categories that serve as best examples, or prototypes, for the category (Rosch, 1975). When sentences were designed to bias encoding of a category term toward an atypical exemplar of the category, Anderson et al. (1976) found that the relevant atypical exemplar served as a better retrieval cue than did a prototypical exemplar. For example, when subjects were asked to recall the statement, "The guest saw the bird that roasted on the grill," the atypical exemplar *chicken* was a more effective cue than was *robin*, which is a more typical bird. Based on their results, Anderson et al. (1976) proposed that words have a family of potential meanings rather than a single meaning. Following this line of reasoning, they suggested that subjects activated schemata based on the context and that comprehension of the sentences involved instantiating a specific meaning on the basis of information in the schemata.

Thus, Anderson et al. (1976) used their results not only to conclude that an important aspect of comprehension is encoding general terms in as specific a manner as possible, but also to draw conclusions about the representation of meaning in memory. This is most clear from their discussion of encoding category terms in context. They contrasted their results with those of Rosch (1975). Rosch's (1973, 1975) studies of categorization indicated that some members of categories are more representative

than others. The internal structure, or organization, of semantic categories seem to be based on a gradient of typicality in which certain members are quite representative of the category as a whole and others are only peripherally related to most of the other members. The gradient of typicality not only is reflected in "goodness of example" ratings that subjects assign to category instances, but also predicts the speed of deciding whether a given instance is a member of the category. Anderson et al. (1976) pointed out that Rosch's conclusions were based on responses to category terms presented in isolation. According to Anderson et al., the internal structure of categories is not fixed, but instead is altered by context. Unfortunately, Anderson et al.'s cued-recall results do not provide unequivocal evidence that readers encode specific instances for category terms in context during comprehension. One reason the results are inconclusive is that it is possible that what is instantiated is not a specific *meaning* but is a specific *referent* for the general term.

Certainly, individuals can understand the meaning of *bird* without assigning to it a particular referent, such as *robin* or *chicken*. The distinction between a word's meaning (intension) and the entities to which it can refer (extension) has long been recognized by philosophers and linguists. This distinction is embodied in Johnson-Laird's (1983) view of constructive processes. Johnson-Laird (1983) proposed that sentences are processed in two stages, characterized by the kind of mental representation used. First, a sentence is encoded into a propositional representation; that is, the information from the sentence is stored in a mental language that is close to the surface form of the sentence. This level of representation involves an understanding of individual words and their relations and is the basic representational format used by most models of comprehension (e.g., Kintsch & van Dijk, 1978). However, Johnson-Laird (1983) argued that a second stage is available, in which the sentence context and inferences based on world knowledge are used to construct a mental model that represents the situation described by the sentence. In this way, a representation of the objects and events mentioned in the sentence is constructed. Both Johnson-Laird (1983) and Garnham (1979) interpreted instantiations as inferences about the referents of general terms based on one's model of the sentence. If instantiations are actually referential inferences, then it is unwarranted to conclude from Anderson et al.'s (1976) data that specific meanings of category terms are instantiated in sentence contexts.

A second factor limiting confidence in Anderson et al.'s (1976) conclusions is their use of a cued-recall task to determine what information was processed during comprehension. Their cued-recall results may reflect inferences made at the time of retrieval rather than inferences made during sentence encoding (Gumenik, 1979).

To test for instantiations during comprehension, a task must be used that assesses what information is activated in memory shortly after sentence processing. Such a test was conducted by Whitney and Kellas (1984), using a modified Stroop (1935) paradigm. Previous studies using this task showed that when target words printed in colored ink were preceded by word or sentence contexts as priming stimuli, longer color-naming latencies were obtained when the target was semantically related to the priming context (Conrad, 1974; Warren, 1972). The amount of color-naming interference, relative to a control condition, reflects the degree to which the information in the target was activated in memory during reading of the priming context.

Thus, the Stroop task allows a direct examination of what information has been accessed during encoding. Whitney and Kellas (1984) had subjects read either a control sentence or a sentence containing a category term biased toward a typical or an atypical exemplar. Significant color-naming interference was obtained for typical exemplar targets, even when preceded by sentences designed to bias interpretation toward atypical exemplars. No color-naming interference was generated to atypical exemplars. Certainly, the subjects were not instantiating the typical exemplars in the atypical-biasing sentences. These results can be explained as the usual pattern of priming from semantic categories. The presence of the category term in the sentence resulted in activation of its meaning. Because the semantic properties of the category overlapped more with the properties of typical than with the properties of atypical exemplars, color-naming interference was generated for the typical exemplars (cf. Rosch & Mervis, 1975; Smith & Medin, 1981). Whitney and Kellas concluded that assigning a referent to a category term is not a routine aspect of sentence encoding. Instead, subjects appear to activate information about the meaning of the category term, and this meaning is more related to typical than to atypical exemplars. In short, the Whitney and Kellas (1984) study was designed to test whether instantiation actually involves selection of a specific meaning, but found instead that instantiation is probably best regarded as an optional inference about what a general term refers to in a given context. These results cannot be attributed to insensitivity of the Stroop task to contextual effects. Several studies have successfully used the task to study the effects of context on accessing word meaning (e.g., Oden & Spira, 1983; Whitney, McKay, Kellas, & Emerson, 1985). However, it is important to note that although there was no evidence for instantiation in the contexts used by Whitney and Kellas (1984), there may be situations in which readers do use instantiation. The previous instantiation results highlight the fact that little is known about the use of constructive processes in comprehension.

The present research was designed to isolate conditions that lead to inferences about the referents of general terms. According to a view of text representation based on mental models, a fundamental goal of comprehension is to use sentence contexts and world knowledge to form as specific and integrated a mental representation of text as possible. A direct consequence of such an approach would seem to be that readers should use constructive processes to instantiate general terms. Unfortunately, evidence is lacking concerning the conditions that engender instantiation

during the encoding of text. However, because mental models are usually postulated to serve as the basis for integration of information across sentences, one may speculate that instantiations might be helpful in aiding sentence integration. A major means for maintaining a sense of coherence in text is the use of anaphoric reference, that is, making reference to a concept introduced earlier in text. If instantiation and other elaborative inferences are fundamental to the formation of a coherent representation of text, then situations involving anaphoric reference might engender such constructive operations.

There are several reasons to suspect that making anaphoric references to general terms might lead to instantiation. There is precedent in the literature for inferences being made in paragraphs, rather than in simple sentences. Although readers do not appear to spontaneously infer the instruments implied by actions in simple sentences (Dosher & Corbett, 1982; Singer, 1979), such inferences are made in some longer contexts in which they aid sentence integration (McKoon & Ratcliff, 1981). Because integrating information from two sentences is easier if the antecedent information is specific rather than general (Garrod & Sanford, 1977), instantiation may be useful in sentence integration. Also, using a pronoun to refer to an antecedent in a text has the effect of foregrounding, or emphasizing, the antecedent (Chafe, 1972; Sanford & Garrod, 1981). Consequently, making a pronominal reference to a general term has the potentially important effect of making the general term the focus of processing. If people do instantiate when an anaphoric reference is made to a general term, it could be because the general term then becomes the linguistically foregrounded topic, to use Chafe's (1972) terminology.

The first of the present experiments investigated the possibility that category terms are instantiated when referred to anaphorically in subsequent sentences. A modified Stroop task was used to determine what information was activated during comprehension, so that constructive rather than reconstructive processing could be studied. As in the previous research, the focus was on making inferences about typical (very representative) and atypical (less representative) members of semantic categories. The inclusion of the typicality variable is important from a methodological standpoint, because an adequate demonstration of a referential inference requires more than obtaining color-naming interference to typical exemplars after typical-biasing contexts. Even if no particular referent is assigned to the category term during sentence processing, simply activating the meaning of the category term will result in Stroop interference when the targets are typical exemplars.

To test whether instantiations are made after integrating sentences during resolution of anaphoric reference, Experiment 1 employed sentence sets in which biasing sentences were followed by sentences making anaphoric references to the category terms. Typical and atypical category members were used as Stroop targets following the biasing sentences. If the specificity of the antecedent

aids in integrating information across sentences (Sanford & Garrod, 1981), it might be expected that subjects will instantiate the appropriate exemplar when given these sentence pairs. This would presumably show up as color-naming interference (relative to control) for typical exemplars after typical-biasing sentences and for atypical exemplars after atypical-biasing sentences.

EXPERIMENT 1

Method

Subjects

The subjects were 48 undergraduates who participated as part of a course requirement for introductory psychology. Two subjects, who showed considerable difficulty in distinguishing the colors, were replaced.

Design and Materials

The independent variables were type of prime (control, typical biasing, atypical biasing) and target typicality (typical, atypical). Both variables were manipulated within subjects.

The exemplars used were the same as those used in Whitney and Kellas's study (1984). They were drawn from Uyeda and Mandler's (1980) typicality norms, in which exemplars were rated on a 7-point scale, with a rating of 1 being *most typical*. The typical exemplars had a mean rating of 1.50, and the atypical exemplars had a mean rating of 4.05.

The priming stimuli consisted of 18 sentence sets adapted from Whitney and Kellas (1984). These stimuli were altered from those in the previous research by adding a sentence to each of the primes. The second sentence of each biasing prime made an anaphoric reference to the category term in the first sentence. Thus, each set contained a sentence pair that biased interpretation toward a typical member of the category, and a sentence pair that biased interpretation toward an atypical member of the category. Each set also contained control sentences for each of the biasing primes. An example sentence set is shown in Table 1.

It should be noted that the control condition used here was somewhat more stringent than is usually the case for these kinds of priming studies. Typically, control sentences are devised to be simply irrelevant to each type of target stimulus. In the present study, the control sentences for the typical exemplars were the same as the typical-biasing primes, except another noun was substituted for the critical category term in each prime. The same type of control sentences were constructed for the atypical exemplars. The advantage

Table 1
Examples of Priming and Target Stimuli

Targets			Primes
Typical	Atypical		
car	helicopter	(T)	The reporter went to the vehicle to look for the papers. She hoped they were in it as she had left them.
		(T-C)	The reporter went to the office to look for the papers. She hoped they were in it as she had left them.
		(A)	The reporter in the vehicle looked down on the parade. While she was up in it, she could see all the people.
		(A-C)	The reporter in the window looked down on the parade. While she was up in it, she could see all the people.

Note — *(T)* = *typical biasing; (T-C)* = *control for typical exemplars; (A)* = *atypical biasing; (A-C)* = *control for atypical exemplars.*

of this type of control procedure was that when the prime-target relationship represented an appropriately biased condition (e.g., a typical-biasing prime followed by the typical exemplar target), the control condition to which it was compared used a sentence pair that differed from the biasing prime only in the presence of the category term. When the prime-target relationship was inappropriate (e.g., the typical-biasing prime followed by the atypical exemplar), the control condition involved the sentence pair written for the target exemplar. This sentence was irrelevant to the target and also used a context different from the inappropriate biasing prime. This more typical control procedure is adequate in these conditions, because the inappropriate contexts do not have the potential of directly priming the Stroop targets (see Whitney et al., 1985).

Crossing the three prime types with the two levels of typicality yielded six unique conditions. Thus, the primes and targets were assigned to six lists, such that one prime from each set appeared once in each list, paired with one of the relevant targets. Within each list, each condition appeared three times. As in the previous research, 36 filler trials were added to each list to reduce the number of trials on which the prime and target were related. The purpose of the filler trials was to prevent the subjects from adopting task-specific reading strategies. The filler trials were similar in style to the critical primes, but each trial involved an unrelated noun as the Stroop target.

Each list was divided into six blocks of nine trials. Each block contained three critical trials randomly mixed with six filler trials. Each target stimulus of the critical trials was randomly assigned to three of the five colors used in this experiment (red, blue, green, yellow, and brown). Across lists and conditions, each color appeared with approximately equal frequency.

Apparatus

The experiment was conducted on an Iconix two-channel tachistoscope. The priming sentences were typed in Orator typeface on 12.5×20 cm white cards. The target stimuli were printed on the same type of cards using an 8-mm stencil and colored ink. Each letter of the priming stimuli subtended a visual angle of approximately 0.5°. All stimuli were centered in the viewing field when presented. A microphone placed near the subject was interfaced with a voice-operated relay, so that naming the color of the target word would stop the millisecond timer of the tachistoscope.

Procedure

The task was presented as one involving the ability to divide attention between reading comprehension and an unrelated distractor activity. Subjects were instructed that their primary focus should be on comprehending the sentences in a normal fashion. Then, when the single word appeared, they were to name the color of the target word as quickly and accurately as possible. Ten practice trials were given, all with unrelated primes and targets.

On each trial, the experimenter gave a ready signal and then initiated the trial. The primes were presented for 7 sec, followed by a 500-msec delay before the target word appeared. The color-naming response stopped the timer and removed the target word from view.

To ensure that the subjects would remain oriented to the reading task, a brief recognition test was given after each block of trials. The subjects had to distinguish between phrases they had seen in the previous block and semantically unrelated distractor phrases. Each subject participated in a single session lasting approximately 50 min.

Results and Discussion

Sentence Recognition

The subjects averaged approximately 93% correct responses (hits and correct rejections) on the recognition task. These results indicate that the subjects were carefully reading the priming sentences.

Color-Naming Latencies

The mean response times for each subject (excluding errors) were analyzed in a 3×2 (prime type × target typicality) within-subjects ANOVA. Response times and percentage error rates (naming incorrect color or latency over 2 sec) are shown in Table 2. Error rates were quite low with no discernible trends. All effects reported as significant reached at least the .05 alpha level, unless otherwise noted.

The ANOVA revealed a significant main effect of prime type [$F(2,94) = 6.03$, $MSe = 9,766$] and a significant prime type × target typicality interaction [$F(2,94) = 3.28$, $MSe = 9,295$].[1] The comparisons of interest within this interaction involve whether color-naming times following the biasing primes differ from those following their respective control primes for each level of typicality.

As expected, when the atypical exemplars served as targets, color-naming times did not differ following typical-biasing primes and control primes ($t < 1$), but significant color-naming interference was obtained following atypical-biasing primes [$t(94) = 3.46$]. Surprisingly, the response latencies for typical exemplar targets showed a different direction of the Stroop effect. When typical exemplars followed typical-biasing primes, the latencies were shorter than those after control primes. This effect narrowly missed standard significance levels [$t(94) = 1.98, p < .07$]. The latencies following atypical biasing versus control prime did not differ significantly ($t < 1$).

These results are very different from those obtained by Whitney and Kellas (1984) using single-sentence versions of these same stimuli. In the previous study, atypical-biasing primes led to color-naming interference for typical targets, indicating a semantic priming effect from the category term but no context-sensitive inference. The present results show that atypical exemplars are activated in memory when the category term in an atypical-biasing context is referred to in a subsequent sentence. There is also evidence for a Stroop effect for typical exemplars following typical-biasing primes, but this effect is in the direction of facilitation.

The facilitation effect was unexpected but not unprecedented. Several other reports of facilitation exist in the literature. For example, Dosher and Corbett (1982) found no Stroop effect when implied instruments followed sentence primes, except when the subjects were instructed to consciously generate the implied instruments. Under these conditions, a Stroop facilitation effect was obtained. Further evidence concerning the facilitation effect is presented in Experiment 2. The key result of the present

Table 2
Mean Color-Naming Latencies (in Milliseconds)
and Percentage Error Rates (in Parentheses)
by Prime Type and Typicality, Experiment 1

Target	Prime Type		
	Control	Typical Biasing	Atypical Biasing
Typical	969 (6)	930 (3)	965 (4)
Atypical	956 (4)	961 (1)	1,024 (1)

experiment is that both types of Stroop effects indicate context-sensitive processing of the category terms. The evidence supports the hypothesis that subjects instantiate category terms under conditions involving anaphoric reference.

EXPERIMENT 2

The purpose behind Experiment 2 was to attempt to replicate the results of Experiment 1 and to provide some insight into the Stroop effects obtained. The modified Stroop paradigm has usually been used to determine what information is automatically activated in memory when a prime is presented. The mechanism of the effect is somewhat controversial, but is usually explained in terms of response competition between the tendency to pronounce the word as opposed to stating the ink color. When the distractor word in the Stroop task has been activated in memory because of the prime, the word is delivered more strongly (or quickly) to the output system and competes with saying the ink color (Posner, 1978; Posner & Snyder, 1975). Clearly, there is more than automatic activation of information related to the prime involved in the facilitation effect. In addition to Dosher and Corbett's (1982) findings, Neill (1978) reported facilitation in the traditional Stroop task of naming the ink color of words that are themselves color terms. He found facilitation when the correct response from the previous trial matched the distractor word on the current trial. It seems that it is the conscious generation of the distractor word before its presentation that leads to the facilitation effect. The mechanisms for this effect is not well understood. One possibility is that even in the control (unrelated) condition, there is a baseline level of interference due to modest response competition between the word and the ink color. When the word that will serve as the Stroop target is actively processed before its presentation as the colored word, it may be processed so quickly or effortlessly when it appears as the Stroop word that response competition is reduced below the baseline level. Thus, color-naming facilitation may represent a condition of little or no interference relative to the control condition. It also has been suggested that generation may lead to a short refractory period or may affect the number of activated responses (see Dosher & Corbett, 1982; Neill, 1978).

In any case, the mechanism of either type of Stroop effect is of less interest here than are the conditions under which they will be reliably obtained. Interference is obtained when there is overlap between the information activated during processing of the prime and the obligatory semantic processing of the distractor word. Facilitation is apparently obtained when processing of the prime results in explicit, active processing of the Stroop word before it is presented.

Based on this interpretation of the Stroop effect, it seems that in Experiment 1, atypical exemplars were activated in memory when atypical-biasing primes were read. When typical-biasing primes were read, the typical exemplars

were consciously generated. This interpretation is based mainly on the similarity between the marginal Stroop facilitation effect obtained here and the one obtained by Dosher and Corbett (1982). To provide additional evidence that facilitation occurs when the target exemplars are processed explicitly in the priming context, a condition was included in Experiment 2 in which the pronouns involved in the anaphoric references were replaced by the intended exemplars. This should indicate whether Stroop facilitation occurs when subjects actually encode the intended exemplar in reading the anaphores involving a pronominal reference.

The logic behind Experiment 2 was the same as for the previous one. The comparisons of interest involve whether, for each type of exemplar, color-naming times after biasing primes differ from color-naming times after control primes for the implicit anaphores (involving pronominal reference). These comparisons provide a test of the instantiation hypothesis, by indicating whether the pattern of Stroop effects is sensitive to biasing context. For the explicit anaphores, these same comparisons serve as a test of whether Stroop facilitation is obtained when instantiations are explicitly encoded. An extended set of materials was developed to provide more evidence for the reliability of the results.

Method

Subjects

The subjects were 84 undergraduates participating to fulfill their introductory psychology research requirement. An additional 30 subjects participated in a normative task on the stimuli. All subjects were native English speakers.

Materials

The category names and exemplars used in this experiment were taken from normative data collected by Uyeda and Mandler (1980) and by McFarland, Duncan, and Kellas (1978). The ratings of the typical exemplars used ranged from 1.00 to 2.00, with a mean of 1.47. Ratings for the atypical exemplars ranged from 3.05 to 5.79, with a mean of 4.08.

The priming stimuli consisted of 24 sentence sets including 16 of the sets from Experiment 1. Two versions of each set were constructed. In the implicit versions, the anaphoric references involved pronouns (as in Experiment 1). In the explicit versions, the pronoun was replaced by the intended exemplar (e.g., "The reporter went to the vehicle to look for the papers. She hoped they were in the car as she had left them"). The control primes were the same for the implicit and explicit versions. Thus, the implicit control primes differed from the biasing primes only in the use of the category term. The explicit control primes differed from their biasing primes only in the use of the category term and the presence of the exemplar in the second sentence.

To check whether the implicit sentence pairs actually biased the intended exemplars, and did so equally for the typical and atypical conditions, 30 subjects were asked to generate the instances that they thought were suggested by the context sentences. Two booklets were constructed for collecting normative data by using one biasing sentence pair from each set in each booklet. Every sentence pair was followed by a question asking what instance of the category term fit the context. Each booklet had 12 typical-biasing and 12 atypical-biasing sentence pairs in a random order. The booklets were duplicated, and each booklet was presented to 15 subjects. The subjects were asked to record their choice of exemplar on a separate answer sheet.

All sentence pairs used in this experiment had the intended exemplar generated by at least 50% of the subjects, with no other single answer given by more than 25% of the subjects. For the typical-biasing sentence pairs, the intended exemplar was supplied an average of 65% of the time. Atypical-biasing pairs led to the intended response 67% of the time. A comparison of the percentages of correct answers for the typical- and atypical-biasing sentences revealed no evidence of differences in the effectiveness of the context for the two types of sentences ($t < 1$).

Design

The independent variables manipulated were anaphore type (implicit, explicit), prime type (typical biasing, atypical biasing, control), and target typicality (typical, atypical). The latter two variables were manipulated within subjects.

For each anaphore type, the primes and targets were assigned to six lists in a fashion similar to that used in Experiment 1. Within each list, each condition appeared four times. Thus, there were 24 critical trials in each list. An additional 24 filler trials were added to each list.

For presentation purposes, each list was divided into six blocks of eight trials. In each block, there were four critical trials randomly interspersed with four filler trials. The order of block presentation was randomized for each subject. Colors (red, blue, green, and brown) were assigned to Stroop targets such that one trial of each condition used each color.

Thus, there were six lists of stimuli for each anaphore type. Subjects were randomly assigned to one list of either the explicit or implicit stimuli.

Apparatus

The experiment was conducted on an Iconix three-channel tachistoscope. The priming sentences were typed on 15×23 cm white cards in Orator typeface. Target stimuli were printed on 15×23 cm cards using an 8-mm stencil and colored ink. Each letter of the priming stimuli and target words subtended a visual angle of approximately 0.3° and 0.4°, respectively. Target words were printed so that they were centered in the viewing field. As before, when the subject named the target color, a microphone, interfaced with a voice-operated relay, stopped the millisecond timer of the tachistoscope.

Procedure

Each subject was tested individually in a single session of 12 practice and 48 experimental trials. Primes and targets were unrelated on all practice trials. The instructions were the same as in the previous experiment. For each trial, the experimenter gave a ready signal and then initiated the trial. After a 250-msec delay, a fixation point was replaced by the sentence pair (priming stimulus). The first word of the sentence pair was aligned with the fixation point. Each prime appeared for 7 sec. After the termination of the prime, there was a 500-msec delay before the presentation of the target.

In this experiment, each recognition test consisted of phrases taken from four of the experimental trials randomly mixed with four distractor phrases, similar in style but unrelated in meaning. Two subjects who averaged less than 75% correct responses on the recognition test and 3 subjects who had considerable difficulty in distinguishing the colors on the practice trials were replaced. The experiment required approximately 45 min per subject.

Results and Discussion

Sentence Recognition

The subjects' performance on the recognition phase indicated that the priming sentences were being read quite carefully. The subjects averaged 91% correct responses (hits and correct rejections).

Color-Naming Latencies

As indicated earlier, the results of greatest interest to the instantiation hypothesis are the pattern of the effects at each level of typicality in the implicit-anaphore condition. The purpose of the explicit condition was to give evidence as to whether Stroop facilitation should be obtained from explicit encoding of the Stroop target during sentence comprehension. Accordingly, separate 3×2 (prime type × typicality) within-subjects ANOVAs were conducted for each anaphore type on the mean color-naming latencies (excluding errors) for each subject in each condition. Each ANOVA was followed by a test of the pairwise comparisons on the mean color-naming times for the biasing primes versus their controls. Subject errors included making an incorrect color-naming response or taking longer than 2 sec to respond. Less than 1% of the trials were excluded for latencies of more than 2 sec.

Implicit-anaphore analysis. Mean response times (excluding errors) and error rates for each condition are shown in Table 3. The error rates were uniformly low and no reliable trends were observed. In the ANOVA on response times, neither main effect reached significance, but there was a significant prime type × target typicality interaction [$F(2,82) = 3.57$, $MSe = 4,244$]. When the Stroop targets were typical exemplars, there was significant facilitation for typical-biasing primes compared to control primes [$t(82) = 2.81$]. Color-naming times following atypical-biasing primes did not differ from those following control primes ($t < 1$). When atypical exemplars served as targets, color-naming times did not differ following typical-biasing and control primes ($t < 1$), but there was color-naming interference after atypical-biasing primes in comparison to control primes [$t(82) = 2.18$]. These effects replicate the pattern obtained with the smaller stimulus set used in Experiment 1.

Explicit-anaphore analysis. Response times and error rates are shown in Table 4. Again, error rates were low with no reliable trends observed. In the ANOVA on response times, both the main effect of prime type and the prime type × target typicality interaction narrowly

Table 3
Mean Color-Naming Latencies (in Milliseconds) and Percentage Error Rates (in Parentheses) by Prime Type and Typicality, Experiment 2, Implicit Anaphores

Target	Prime Type		
	Control	Typical Biasing	Atypical Biasing
Typical	944 (3)	904 (4)	936 (4)
Atypical	923 (4)	934 (1)	953 (4)

Table 4
Mean Color-Naming Latencies (in Milliseconds) and Percentage Error Rates (in Parentheses) by Prime Type and Typicality, Experiment 2, Explicit Anaphores

Target	Prime Type		
	Control	Typical Biasing	Atypical Biasing
Typical	951 (2)	915 (5)	934 (2)
Atypical	935 (4)	937 (2)	897 (4)

missed standard levels of significance [$F(2,82) = 2.70$, $MSe = 5,772$, $p = .07$, and $F(2,82) = 2.62$, $MSe = 7,011$, $p = .08$, for the main effects and the interaction, respectively]. The typicality main effect did not approach significance. The same four planned, pairwise comparisons were performed as in the implicit-anaphore analysis. Again, two effects were significant, but both were in the direction of facilitation (relative to control). For typical targets, there was facilitation after typical-biasing primes [$t(82) = 2.02$], and likewise, there was facilitation for atypical exemplars after atypical-biasing primes [$t(82) = 2.08$]. These two significant effects are consistent with the interpretation that color-naming facilitation results from explicit processing of the prime-target relationship before presentation of the prime (cf. Dosher & Corbett, 1982).

Overall, the results of this experiment support the idea that when reading the typical-biasing sentence pairs, subjects explicitly generated the typical exemplars, leading to color-naming facilitation when the exemplar was presented as the Stroop target. When reading the atypical-biasing sentences, subjects activated information in memory that led to color-naming interference when the atypical exemplar was presented.

The difference in the direction of the Stroop results for the typical and atypical exemplars, after their respective biasing primes, would seem to bear out the contention that making inferences need not be an all-or-none process (see McKoon & Ratcliff, 1981). Theorists attempting to account for text-based inferences have generally posited three isolable stages (cf. Chafe, 1972; Clark, 1978; McKoon & Ratcliff, 1980). A concept that was presented earlier in a text must first be accessed in memory. Then the activated concept must be brought into working memory (or a short-term buffer), containing the information currently being processed. Finally, the concept is linked to the information being processed. This integrated structure constitutes the mental representation of the text. These stages can easily be applied to elaborative inferences as well. For a concept to be added to one's representation of text, the concept must first be activated in memory and then actively processed in working memory in order to add it to the currently available elements of the text.

According to this framework, the concept relevent to the inference can be activated in memory without connecting the concept with the information in the text. Activating the implied referent for a category term would by itself lead to Stroop interference. Going beyond activation and integrating the activated referent into the representation of the text could lead to the facilitation effect. However, the important point to bear in mind is that both types of Stroop effects obtained indicate that subjects were instantiating the category terms. The results cannot be due to direct priming of the target exemplars by words in the context because, except for the category terms, the sentences were the same for the control and the biasing primes. In addition, the modified Stroop task

is not subject to backward priming effects (Warren, 1972) that have limited the interpretability of some studies using the lexical decision task (Koriat, 1981).

EXPERIMENT 3

It was noted earlier that pronominalization of a concept is a method of foregrounding, that is, making the concept the focus of processing (Chafe, 1972). This raises the possibility that subjects instantiate when resolving an anaphoric reference because the category term is emphasized. Although Whitney and Kellas (1984) found no evidence for instantiation in single sentences, the category terms were in the predicates of the sentences. If it was the foregrounding that the categories received in the previous experiments that led to instantiation, then subjects should instantiate in single-sentence contexts when the categories are emphasized by being placed in the subject position of the sentences.

Method

Subjects

The subjects were 42 undergraduate students drawn from the same subject pool used for the previous experiments. An additional 30 subjects participated in the normative task on the stimuli.

Materials

Again, 24 sentence sets were used as priming stimuli. The implicit-anaphore sentence pairs used in Experiment 2 were revised so that each prime consisted of a single sentence with the category term in the subject position of the sentence (e.g., "The vehicle allowed the reporter to look down on the parade.").

Normative data on the effectiveness of the sentence contexts was collected in the same way as in the previous experiment. Another group of 30 subjects participated in the normative procedure but, in addition to being asked to generate the exemplar suggested by the context, subjects also circled a number from 1 (very easy) to 5 (very difficult) indicating how easy it was to generate an instance for the sentence. Because the semantic contexts that biased interpretation were the same as in the previous experiment, it was expected that similar consistency in generating the appropriate instances would be obtained. This was certainly the case. All of the intended exemplars were provided at least 50% of the time. For the typical-biasing sentences, the intended exemplar was supplied 68% of the time. For the atypical-biasing sentences, the intended exemplar was given 65% of the time. Further evidence that the contexts were equally biasing came from the ratings on the ease of generating the instances. The mean rating of ease of generation for the correct responses was 1.83 for the typical-biasing sentences and 1.75 for the atypical-biasing sentences. No significant differences between the typical- and atypical-biasing contexts were obtained for either the percentage of correct generations or rating of ease of generation (both ts < 1).

Design and Procedure

The independent variables manipulated were prime type (typical biasing, atypical biasing, control) and target typicality (typical, atypical). Both were within-subjects factors.

The assignment of priming stimuli to lists and of targets to colors was the same as in Experiment 2. All aspects of the procedure were the same, except that, because the primes were single sentences, the subjects were given 4 sec to read each one. The same kind of recognition task was given to check comprehension. One subject

in this experiment was replaced for failure to achieve 75% correct responses on the recognition phase, and 1 was replaced for problems with distinguishing the colors. The experiment required approximately 30 min to complete for each subject.

Results and Discussion

Sentence Recognition

As in the previous experiment, the subjects performed well on the recognition task. The subjects averaged 92% correct responses.

Color-Naming Latencies

A 3 × 2 (prime type × target typicality) within-subjects ANOVA was performed on the mean response times for each subject in each condition (excluding errors). The mean response times and error rates for each condition are presented in Table 5. No reliable trends in error rates were observed. Again, less than 1% of the trials were excluded for response times over 2 sec. In the ANOVA on response times, there was a significant main effect of prime type [$F(2,82) = 8.00$, $MSe = 8,663$], which was qualified by a significant prime type × target typicality interaction [$F(2,82) = 4.44$, $MSe = 5,887$]. The pattern of the pairwise comparisons of interest within this interaction replicated that of the implicit-anaphore analysis of the previous experiment. For typical-exemplar targets, color-naming times were shorter after typical-biasing primes than after control primes [$t(82) = 3.64$]. For the atypical exemplars, there was color-naming interference after atypical-biasing primes in comparison to control primes [$t(82) = 2.63$]. There was no color-naming interference or facilitation for typical exemplars after atypical-biasing primes or for atypical exemplars after typical-biasing primes (both $ts < 1$).

Presenting the category terms in the subject positions of the sentences led to the same pattern of results as when an anaphoric reference was made to the category terms. Again, the pattern of Stroop effects was dependent on the biasing context. Even in the atypical-biasing contexts, there was more to comprehension than simply activating the meaning of the category terms. Otherwise, there would have been interference for typical exemplars after atypical-biasing contexts.

GENERAL DISCUSSION

The primary purpose of the present research was to isolate conditions in which readers infer a specific referent for a category term. Evidence from the modified Stroop

Table 5
Mean Color-Naming Latencies (in Milliseconds) and Percentage Error Rates (in Parentheses) by Prime Type and Typicality, Experiment 3

Target	Prime Type		
	Control	Typical Biasing	Atypical Biasing
Typical	941 (4)	880 (5)	957 (1)
Atypical	916 (5)	926 (4)	963 (4)

task indicated that category terms were instantiated in biasing contexts when referred to anaphorically and when used as the subject of a sentence. Surprisingly, in all experiments, the direction of the Stroop effects was dependent on the typicality of the biased exemplar. There was facilitation of color naming for typical exemplars after typical-biasing contexts. These Stroop effects were measured relative to control conditions that were the same as the biasing sentences except for the critical category terms. Consequently, the results cannot be attributed to direct priming of the Stroop targets by verbal associates in the context.

Taken together, the results of the present experiments suggest that whether a referent for a general term is inferred or not would seem to depend on the role the term plays in the text. Anaphoric reference was investigated to determine whether readers instantiate in order to facilitate text integration. Placing the category terms in the subject position was used to test the hypothesis that instantiation occurs when the category term is the focus of processing. The contextual conditions used in each of the experiments are complementary if viewed from the standpoint of their function in constructing mental models. Sentence topics presumably received special processing emphasis because they are more likely to be repeatedly referred to and to serve as a central concept to be integrated with successive information (Sanford & Garrod, 1981). When a concept that was not previously introduced is referred to by a pronoun, it is foregrounded and receives emphasis in processing. The results converge on the notion that it is a concept's status as the focus of processing that leads to instantiation. If a mental model of text is being constructed, then it makes sense to give whatever is the focus of processing the greatest specification in the model. It is less important to elaborate concepts that are not the focal point of interpretation, which is usually the case when the category term is in the predicate.

It should be acknowledged that instantiation and other elaborative inferences are likely to be affected by the reader's particular goals as well as by the role played by the concepts of interest in the given context. In fact, by designating the formation of a model-based representation of text as an optional stage of processing, Johnson-Laird (1983) affirms that text can be read in a nonconstructive way (e.g., when skimming). However, it seems probable that, across most situations, a fundamental goal of reading is comprehension through the construction of an elaborated representation. In the present study, care was taken not to elicit use of task-specific reading strategies so that the results would have the widest possible generality regarding reading category terms in context. It can be concluded that if subjects are asked simply to read for comprehension, instantiations are made when the category terms are the focus of processing.

An additional finding of interest in the present research was the difference in the Stroop effects obtained with instantiation of typical and atypical exemplars. Although further research is needed to clarify the mechanism of the

facilitation effect, there is evidence from Experiment 2 and prior studies (Dosher & Corbett, 1982; Neill, 1978) that facilitation is obtained when the target word is consciously preprocessed. The explicit condition in Experiment 2 showed that it is possible to obtain the facilitation effect with atypical exemplars. Thus, there must be a difference in the way typical-biasing and atypical-biasing primes are processed.

It is interesting to note that the difference in the direction of the Stroop effects was obtained despite equating the strength of the typical- and atypical-biasing contexts as much as possible. However, normatively equating context strength by asking subjects to generate the implied referents cannot guarantee that subjects will treat the contexts equivalently in a more normal reading situation. In the normative task, the subjects were reading for the purpose of instantiating category terms. In reading the sentences in the experiments, the subjects' comprehension resources were certainly less specifically directed toward instantiation. The subjects may have been more conservative about actually generating instantiations to the more unusual atypical-biasing contexts.

This study started with the question of what conditions engender the use of constructive processes at encoding. The results provide the beginning of an answer to that question by highlighting the importance of the sentence topic and the complexity of the processes behind elaborative inferences, such as instantiation. It is of interest to compare these results on instantiation with results concerning the conditions that engender other kinds of elaborative inferences. Probably the most frequently investigated elaborative inference involves activating the instruments implied by actions (e.g., Dosher & Corbett, 1982; McKoon & Ratcliff, 1981; Singer, 1979). Studies of instrumental inference have generally found that such inferences are limited to cases in which they are useful for integrating propositions within paragraphs. The present results on inferences about implied referents indicate that they not only are useful for sentence integration, but also are made even in brief single-sentence contexts as long as the inference involves a foregrounded term. Thus, the present findings converge with other evidence indicating that the sentence topic is accorded a special processing status (cf. Garrod & Sanford, 1981; Greenspan & Segal, 1984; Hornby, 1972).

REFERENCES

Anderson, R. C., Pichert, J. W., Goetz, E. T., Shallert, D. L., Stevens, K. W., & Trollip, S. R. (1976). Instantiation of general terms. *Journal of Verbal Learning & Verbal Behavior*, 15, 667-679.
Anderson, R. C., & Shifrin, Z. (1980). The meaning of words in context. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 331-348). Hillsdale, NJ: Erlbaum.
Bransford, J. D., & Johnson, M. D. (1973). Considerations of some problems in comprehension. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.
Chafe, W. (1972). Discourse structure and human knowledge. In J. B.

Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge* (pp. 41-69). Washington, DC: Winston.
Clark, H. (1978). Inferences in comprehension. In D. LaBerge & S. Samuels (Eds.), *Basic processes in reading, perception and comprehension* (pp. 243-263). Hillsdale, NJ: Erlbaum.
Conrad, C. (1974). Context effects in sentence comprehension: A study of the subjective lexicon. *Memory & Cognition*, 2, 130-138.
Dosher, B. A., & Corbett, A. T. (1982). Instrument inferences and verb schemata. *Memory & Cognition*, 10, 531-539.
Garnham, A. (1979). Instantiation of verbs. *Quarterly Journal of Experimental Psychology*, 31, 207-214.
Garnham, A. (1982). Testing psychological theories about inference making. *Memory & Cognition*, 10, 341-349.
Garrod, S., & Sanford, A. J. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning & Verbal Behavior*, 16, 77-90.
Garrod, S., & Sanford, A. J. (1981). Bridging inferences and the extended domain of reference. In J. Long & A. Baddley (Eds.), *Attention and performance IX* (pp. 331-346). Hillsdale, NJ: Erlbaum.
Greenspan, S. L., & Segal, E. M. (1984). Reference and comprehension: A topic-comment analysis of sentence-picture verification. *Cognitive Psychology*, 16, 556-606.
Gumenik, W. E. (1979). The advantage of specific terms over general terms as cues for sentence recall: Instantiation or retrieval? *Memory & Cognition*, 7, 240-244.
Hornby. P. A. (1972). The psychological subject and predicate. *Cognitive Psychology*, 3, 632-642.
Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.
Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
Koriat, A. (1981). Semantic facilitation in lexical decisions as a function of prime-target association. *Memory & Cognition*, 6, 372-375.
McFarland, C. E., Jr., Duncan, E. M., & Kellas, G. (1978). Isolating the typicality effect in semantic memory. *Quarterly Journal of Experimental Psychology*, 30, 251-262.
McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning & Verbal Behavior*, 19, 668-682.
McKoon, G., & Ratcliff, R. (1981). The comprehension processes and memory structures involved in instrumental inference. *Journal of Verbal Learning & Verbal Behavior*, 20, 671-682.
Murphy, G. L. (1984). Establishing and accessing referents in discourse. *Memory & Cognition*, 12, 489-497.
Neill, W. T. (1978). Decision processes in selective attention: Response priming in the Stroop color-word task. *Perception & Psychophysics*, 23, 80-84.
Oden, G. C., & Spira, J. L. (1983). Influence of context on the activation and selection of ambiguous word senses. *Quarterly Journal of Experimental Psychology*, 35A, 51-64.
Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 56-86). Hillsdale, NJ: Erlbaum.
Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 111-144). New York: Academic Press.
Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.
Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.
Sanford, A. J. & Garrod, S. C. (1981). *Understanding written language*. New York: Wiley.
Singer, M. (1979). Processes of inference during sentence encoding. *Memory & Cognition*, 7, 192-200.
Smith, E. E., & Medin, D. L. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
Spiro, R. J. (1980). Constructive processes in prose comprehension and recall. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theo-*

retical issues in reading comprehension (pp. 245-278). Hillsdale, NJ: Erlbaum.

STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

UYEDA, K. M., & MANDLER, G. (1980). Prototypicality norms for 28 semantic categories. *Behavior Research Methods & Instrumentation*, **12**, 587-595.

WARREN, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology*, **94**, 90-100.

WHITNEY, P., & KELLAS, G. (1984). Processing category terms in context: Instantiation and the structure of semantic categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **10**, 95-103.

WHITNEY, P., McKAY, T., KELLAS, G., & EMERSON, W. A., JR. (1985). Semantic activation of noun concepts in context. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **11**, 126-135.

NOTE

1. In all three experiments reported here, the critical prime × typicality interaction showed results that were parallel to the subjects' analyses when tested using materials as the random factor.