On-Line Evidence for Elaborative Logical Inferences in Text

R. Brooke Lea
University of Massachusetts at Amherst

A model of propositional-logic reasoning proposed by M. D. S. Braine, B. J. Reiser, and B. Rumain (1984) claims that inferences such as "p or q; not p //: q" are made spontaneously by readers at the moment both premises are available. This claim is inconsistent with some evidence in the text-processing literature that suggests that only those inferences necessary for textual coherence are made spontaneously. In the present study, participants read stories in which a logical inference was not necessary to maintain textual coherence, and inference making was assessed with on-line probes. Two experiments tested logical forms central to Braine et al.'s model, and both indicated that participants were making the logical inferences. Two further experiments replicated this result with stories that did not begin with thematic titles. These findings support Braine et al.'s prediction that some propositional-logic inferences are made routinely in texts that do not require them for coherence.

When people read text, they often must elaborate on the information stated explicitly in the passage by making a variety of inferences. Usually a text will require the reader to fill elliptical gaps, and people's competence in doing so permits some economy in presentation; the writer need not provide all of the intended communication explicitly but can rely on the reader to produce the remainder. A variety of inferences have been investigated in the text-processing literature, such as case inferences (e.g., Corbett & Dosher, 1978), causal relations (e.g., Myers, Shinjo, & Duffy, 1987), trait inferences (e.g., Newman & Uleman, 1989), anaphoric inferences (e.g., E. J. O'Brien, Duffy, & Myers, 1986), goal-related inferences (e.g., Abbott & Black, 1986), thematic inferences (e.g., Kieras, 1985), inferences from story grammars (e.g., Mandler, 1984), inferences from scripts (e.g., Bower, Black, & Turner, 1979), propositional-logic inferences (Lea, O'Brien, Fisch, Noveck, & Braine, 1990), and many others.

After more than 2 decades of work to compile an inventory of the inferences people make during reading, several researchers recently have attempted to provide a large-scale account of text comprehension (e.g., Graesser, Singer, & Trabasso, 1994; Kintsch, 1988; McKeon & Ratcliff, 1992). Important to this work are issues regarding the conditions under which inferences are made. In particular, the time course of inference making has been central to many recent investigations (e.g., Balota, Flores d'Arcais, & Rayner, 1990; Graesser & Bower, 1990; McKeon & Ratcliff, 1992; Singer, Graesser, & Trabasso, 1994). One question that is often posed is which types of inferences tend to be made when the licensing conditions for the inference are met (i.e., when the information required to make the inference is first available) and which are made only when needed for textual coherence or to answer a question. Obviously, the number of potential inferences that could be generated from a given text is uncountable, and although individual differences must exist among readers as to what they choose to elaborate on, the question arises as to whether there exists a subset of inferences that most if not all readers would make given the same piece of text. It generally is acknowledged that inferences required to maintain coherence are made during reading. There is less agreement, however, about when inferences that are not necessary for coherence (i.e., forward or elaborative inferences) are made. Opinions vary considerably, but to the extent that a consensus exists, researchers tend to believe that such inferences are either not made or are only made under specific conditions (Keefe & McDaniel, 1993; McKeon & Ratcliff, 1980, 1986, 1989, 1992; Murray, Klin, & Myers, 1993; Noordman & Vonk, 1992; E. J. O'Brien, Shank, Myers, & Rayner, 1988; Potts, Keenan, & Golding, 1988; Singer, 1988; Singer & Ferreira, 1983; Swinney & Osterhout, 1990). The present research is concerned with precisely this issue, and in the experiments presented here, I attempted to specify when in the course of reading text and under what text conditions propositional-logic inferences are drawn.
Consider the following brief text (from Lea et al., 1990, p. 363) as an illustration of how propositional-logic inferences and other sorts of inferences can be involved in discourse:

Dress-up

(1) Jerry was trying to decide what to wear to meet his fiancee's parents.
(2) "I'd like to wear either my striped or my checkered shirt," he thought.
(3) "If I wear my striped shirt, I'll have to wear my blue pants since they match the shirt," he thought, "but I'm not sure where my blue pants are."
(4) He looked in his closet for his checkered shirt and saw that it was wrinkled, so he decided he could not wear that.
(5a) "So I'd better find my blue pants," he thought.

Note that Sentence 1 presents the theme of the passage. The knowledge structure relating to dressing oneself (e.g., the dressing script) would be elicited by this topic sentence, permitting the author to focus on Jerry's decision about what to wear, rather than on a detailed description of the process of dressing. Social protocol relating to meeting a fiancee's parents might be accessed, thereby allowing a reader to infer that Jerry's goal is to make a good impression, that his plan includes his appearance, and that the instruments involved in his plan include the clothes he will wear. The anaphoric inference that he in Sentence 2 refers to Jerry in Sentence 1 must be made if readers are to appreciate that the person who is deciding about shirts is the same person who is about to meet his fiancee's parents; this reference needs to be maintained throughout the passage. Also required for the comprehension of the story are propositional-logic inferences. Unless the disjunction or in Sentence 2, the conditional if in Sentence 3, and the negation not in Sentence 4 are understood and the appropriate inferences are drawn from them, the story will be difficult, if not impossible, to comprehend. Consequently, these propositional-logic inferences are required to comprehend the story.

Whether a text-based inference is drawn on-line or not often depends on the role the potential inference might play in the coherence of a particular text. For example, consider the difference between a text in which Sentence 6 is followed by Sentence 7a and one in which Sentence 6 is followed by Sentence 7b; (Potts et al., 1988, p. 405):

(6) No longer able to control his anger, the husband threw the delicate porcelain vase against the wall.
(7a) He had been angry for weeks, but had refused to seek help.
(7b) It cost him well over one hundred dollars to replace the vase.

When Sentence 6 is followed by Sentence 7a, the reader is permitted to make the forward elaborative inference that the vase broke. However, when Sentence 7b follows Sentence 6, the inference is forced for the purpose of coherence; that is, the inference (that the vase broke) is required to connect information in Sentence 7b about the husband replacing a vase to information in Sentence 6 about him throwing it against the wall. Thus, Sentence 6 followed by Sentence 7b calls for a bridging inference. Potts et al. (1988) concluded from their results that inferences were made in the bridging versions of their texts but not in the forward-inference versions. Most researchers agree that when an inference is necessary to maintain or achieve textual coherence, it will be made on-line. As the previous example suggests, a type of inference that is not required for coherence can be made so by making its output necessary to bridge a conceptual gap, that is, by making it a bridging inference. Thus, evidence that a type of inference (such as inferences about predictable events) can be made on-line does not necessarily mean that it will be made in circumstances in which it is not necessary for textual coherence. Hence, the question of which inferences are drawn on-line in text comprehension interacts in many cases with the role that inference plays in the text.

Propositional-Logic Inferences in Text

Propositional-logic inferences were recently added to the inventory of inferences made in text comprehension. Lea et al. (1990) found that participants reading text had little trouble making logical inferences of the sort specified by a mental-logic model proposed by Braine, Reiser, and Rumain (1984). The model consists of a set of inference schemas, a reasoning program that operates the schemas, and a set of independently motivated pragmatic principles (e.g., Gricean implicatures; Grice, 1975). The model predicts the inferences people make routinely from the meaning of English-language sentence particles such as if, and, or, and not. It is hypothesized that one function of this logic in everyday life is to integrate information from different propositions in a discourse. For instance, to understand the story about Jerry deciding what to wear, the reader needs to combine the information in Sentences 2 and 4 to make the or-elimination inference that Jerry's shirt choice has been narrowed to the striped one (striped or checkered; not checkered / .-. striped), and then the reader needs to combine that inferred information with the conditional premise in Sentence 3 (if striped then blue pants) to deduce by modus ponens that Jerry would need to find his blue pants. Examples of schemas central to the model are modus ponens (if p then q; p / .-. q), or-elimination (a or b; not a / .-. b), double negation (not not r / .-. r), and not both elimination (not both p and q; q / .-. not p). Braine et al. (1984) described both a direct-reasoning routine and some indirect-reasoning strategies; the direct routine operates spontaneously to draw inferences from given information and is hypothesized to be universal, whereas the strategies component describes some reasoning strategies that may be available to many undergraduates but not to everyone. The work done by Lea et al. and the present research are concerned only with the direct-reasoning routine. (For a more detailed discussion of the inference schemas and reasoning program, see Braine, 1990; Braine & O'Brien, 1991; Lea et al., 1990; or D. P. O'Brien, 1993; for a discussion about how reasoning generally is both logical and pragmatic, see Noveck, Lea, Davidson, & O'Brien, 1991.)

Lea et al. (1990) presented participants with short (five- to six-sentence) stories and asked the participants to indicate whether or not the last sentence made sense in the context of the story. For example, the story presented previously about Jerry deciding what to wear to meet his fiancee's parents was one of the experimental texts, and participants were presented either with Sentence 5a or with the following sentence and

---

1 Other mental-logic models have been proposed (e.g., Rips, 1994), and the present research is not intended to distinguish among them.
LOGICAL INFERENCES IN TEXT

(5b) "So I don't have to find my blue pants," he thought.

Unless the reader has made the logical inferences predicted by Braine et al.'s (1984) model (hereinafter "the model"), it would be difficult to determine whether Sentence 5a or Sentence 5b makes sense in the context of the story. Lea et al. found that participants correctly accepted the appropriate final sentence and correctly rejected logically unsound sentences 94% of the time. In a second task, participants were asked to indicate whether a given piece of information was presented explicitly by the text or whether it had to be inferred. Participants were given three types of information for each story: (a) a paraphrase of information presented explicitly in the text, (b) an inference that followed from the story propositions by application of the schemas of the model, and (c) an inference that was valid in standard logic but would not have been drawn by the model. Participants often (69%) falsely recognized model inferences as having been presented in the text, though they rarely (17%) did so with the foil inferences. Similar results were obtained by Fisch (1991) for texts in which model inferences were not bridging inferences. These results indicate that readers draw the inferences predicted by the model routinely enough that they often do not realize they are making inferences.

As noted above, the question of whether inferences are drawn on-line is often addressed in terms of coherence. Braine et al.'s (1984) model, however, along with other models that discuss reasoning performance in discourse (e.g., Sperber & Wilson, 1986), predicts that when the premises necessary for the execution of a model inference are simultaneously present in the reader's mind (working memory), the inference will be made—regardless of the influence that inference might or might not have on local coherence. The model's claim as it currently stands, therefore, is inconsistent with assumptions often made in the text-processing literature about when nonbridging inferences are made.

Therefore, the present research was designed to discover when model inferences are drawn in text comprehension and under what circumstances they are drawn. The experimental strategy consisted of presenting participants with texts in which model inferences were elaborative inferences that were not necessary for textual coherence and then testing participants through an on-line priming technique to see whether the inferences were in fact drawn spontaneously while participants were reading the text. For example, consider the following sample text:

**What to Eat**

Tony was trying desperately to stick to his diet.

"Well," said his mother Maria, "You can have either bread or corn flakes with breakfast."

Tony seemed to take forever before giving an answer.

"Alright Ma, since I have to skip something, I won't have the corn flakes."

Participants read the text sentence-by-sentence on a computer screen at their own pace and were asked for a lexical decision on a word semantically associated with the proposition inferred. (In this case, the inference was *bread*, and BUTTER was the associate.) If participants spontaneously made the inference bread when they read that Tony decided against corn flakes, then words associated with the inference would have been activated and lexical decisions on such words should have been faster as compared with a control condition in which the inference was not made. Control (no-inference) versions of the stories were identical except that the second premise was not asserted in the final sentence. For the preceding example, the control version of the story ended with the sentence,

He wondered if he could talk her into more helpings of corn flakes.

Note that the second (logical) premise, not corn flakes, was absent from the control sentence, thereby depriving readers of the information necessary to make the inference that Tony would have bread.

Experiment 1 had two parts, both designed to investigate whether adult participants would draw an inference from one of the model's core schemas (or-elimination, i.e., a or b; not a / ∴ b) at the moment the second of two premises was presented and understood by the participant. Because context-checking can be a problem with lexical-decision measures (Potts et al., 1988), an additional condition was run using a naming task, which is not subject to context-checking effects (Potts et al., 1988). The condition using the lexical-decision task is reported as Experiment 1a, and the naming-task condition is reported as Experiment 1b. Each story began with a theme-setting title; Experiment 2 presented participants with the same stories used in Experiment 1, but without the titles provided in Experiment 1. As previously discussed, current research on forward inferences indicates that these elaborations may only be made in rather constrained circumstances, especially those in which the inference concept is in focus at the time of test. Experiment 2, therefore, attempted to generalize the results of Experiment 1 to situations in which the focusing influence of a title was absent and began to explore the role that reading focus might play in engaging the model's reasoning program. For present purposes, reading focus is defined as those aspects of the text on which a reader chooses to concentrate his or her processing resources. The model's prediction regarding when inferences will be made does not invoke goal setting, reasoning strategies, or attention to certain features; the only condition specified is that the requisite premises should be in working memory at the same time. Because only one sentence separated the two premises in the present texts, it was assumed that when participants were reading the second premise, propositions from the first premise were still available (Kintsch & van Dijk, 1978). The model, therefore, would not make differential predictions between Experiments 1 and 2; in each experiment, the two premises for or-elimination were equally available. Thus, according to the model, the thematic contribution of the title should not have affected participants' inference making. Experiment 3 investigated a second core schema from the model (modus ponens,
i.e., if $p$ then $q$; $p / \vdash q$ in the same manner used in Experiment 1a, and Experiment 4 tested the modus ponens stories, again with the goal of determining whether the inference effects can be generalized to texts without titles.

### Experiment 1a

Both parts of Experiment 1 investigated or-elimination ($a$ or $b$; not $a / \vdash b$) in texts that contained a thematic title (see Appendix A for examples of the stories). The texts were each four sentences long and were presented in the following order: title, scene-setting sentence, first premise, filler sentence, second premise (or control sentence), lexical decision probe, comprehension statement. The title and first sentence provided a theme and some of the context that would be present in an ordinary text. A filler sentence followed the first premise to ensure that any activation of the probe item at the time of test was due to inference rather than to residual activation from the explicit presentation of its associate in the first premise (i.e., BUTTER should have been activated because bread was inferred at the end of the test, not because bread was read at the beginning). Thus, the filler sentence allowed activation from the explicit presentation of the to-be-inferred proposition to decay.

Inference versions of the stories contained the second premise in the last sentence, immediately before the lexical-decision probe. The no-inference versions were identical except that, instead of asserting the second premise, key words from that proposition were simply mentioned in the final sentence; this was done to control for semantic priming. In the aforementioned example, the control sentence mentioned corn flakes without providing the information that Tony had decided not to eat them, as was true in the inference version.

Note that the final sentences for each version did not lead to different contexts; in the story about Tony, both final sentences described Tony’s thoughts about eating corn flakes. Thus, context-checking, which can be a concern with the lexical-decision task, is avoided. I return to the issue of context-checking in the Discussion for Experiment 1a.

Finally, participants were asked to respond true or false to a comprehension question about the text. Participants were led to believe that the experiment was about comprehension and that these questions constituted the critical stimuli, when in fact they served only to ensure that the participants read the texts carefully.

Note that unlike the stories used by Lea et al. (1990), these texts contained information that led to inferences that were not bridging inferences. That is, the reader need not have made the inference that Tony decided to have bread for breakfast to have maintained a coherent representation of the text.

### Method

**Participants.** Forty-seven undergraduate psychology students participated to fulfill part of a course requirement in Introductory Psychology at New York University. All participants were native speakers of English.

**Materials and experimental design.** Each participant read 60 stories: 22 were inference stories, 22 were control (no-inference) stories, and 16 were filler stories. No-inference stories differed from their inference counterparts only in that they did not assert the second premise; instead, key words from the second premise proposition were simply mentioned. Of the 22 stories of each type, 12 contained probes that were English words, and 10 contained probes that were nonwords. The comparison of interest was the lexical-decision latencies for the 12 words that followed inference stories versus the lexical-decision latencies for the 12 words that followed no-inference stories. Nonword stories were included so that the proportion of word and nonword responses was approximately equivalent; 47% of the probes were nonwords. The associates used for the lexical-decision task were obtained from two lists of associative norms (Jenkins, 1970; Keppel & Strand, 1970). Approximately one third of the associates were based on opposite relations (e.g., cold-hot); the remaining two thirds were based either on synonyms (e.g., insect-bug) or on general semantic relatedness (e.g., cheese-crinker).

In addition to the 44 stories that contained logical premises ($22$ inference + 22 no-inference), there were 16 filler stories that were designed around the same form but contained no logical particles. These stories had been developed for use in an experiment about trait inferences and were included in this research to help disguise the focus of the experiments. Half of the filler stories contained word probes and half contained nonword probes. All stories were followed by a comprehension question; the right answer was true for half the stories and false for half the stories. The comprehension questions were given to encourage readers to take the texts seriously and to read them carefully.

To summarize, participants read 60 stories, 44 of which contained at least one logical premise and 16 of which did not. Thirty-two of the 60 stories contained word probes, and the remaining 28 contained nonword probes. Participants read 2 practice stories before they began the experiment.

Participants were randomly assigned to one of two groups. Each participant in Group A read stories 1–12 in their inference version and stories 13–24 in their no-inference version; participants in Group B read the no-inference versions of the first 12 stories and the inference versions of the second 12. Thus, the within-subjects factor (inference set) and the between-subjects factor (group) yielded a split-plot, Group × Inference Set design. I predicted that the main effect for inference set (inference vs. no-inference) would be significant but that neither the main effect for group (A vs. B) nor the interaction would achieve significance.

**Procedure.** The stories were presented to participants in a random order. A story trial began with the title presented in the center of the screen. When participants were ready to begin reading the story, they pressed the space bar, which removed the title and replaced it with the first sentence of the story. Participants were asked to read the stories (sentence by sentence) at their own pace; when they read and understood the sentence on the screen, they pressed the space bar and moved on to the next sentence. Sentences replaced each other on the screen; no more than one sentence was present on the screen at a time. When participants read the sentence immediately preceding the probe, their button press cleared the screen and activated the lexical-decision probe, which was presented in all capital letters with an asterisk at either side. Participants understood that they should indicate as quickly and accurately as possible whether the probe was an English word or not. The 7/ key on the keyboard functioned as the yes key, and the equally accessible 2 key served as the no key. Participants

---

2 I am grateful to James Uleman for providing the filler texts.
were instructed to keep a finger on these two keys at all times. After the lexical decision, a sentence appeared on the screen and participants were asked to indicate whether the sentence followed from or was true of the story they just read. Participants’ response to this comprehension statement advanced them to the next story. Of the 60 comprehension statements, 11 were the propositions that could be deduced from the premises in the stories, 22 were negations of something presented explicitly in the text, and 27 were simple (nonlogical) elaborations from the text.

Results

The dependent measure of principal interest was the participant’s reaction time to the lexical-decision probes from inference stories and their no-inference counterparts. Several manipulations on the data were made before testing the difference between the two groups. (The data-preparation procedure was adopted from Uleman, Hon, Roman, & Moskowitz, in press.) First, any latencies associated with wrong lexical decisions were discarded. Then the natural logs of the reaction times were computed in an effort to normalize the data and minimize extreme points. Next a bivariate regression was performed in which reaction time was analyzed as a function of ordinal sequence of story presentation. This regression was done for each participant while ignoring all within-subject conditions and was performed to eliminate any effect due to the order of story presentation. The residuals from this regression were used to examine the outliers for each participant, and any data points that were further than three standard deviations from the mean for that participant were eliminated. The remaining residuals were then submitted to a second pass of this regression analysis, and any additional points beyond three standard deviations were removed. This trimming procedure removed less than 1% of the scores.

Mean response time for targets associated with no-inference texts was 1,196 ms (2.7% errors); the response time for targets associated with inference texts was 1,133 ms (2.8% errors). These means represent the response times after they were transformed to logs, averaged by participant, and then transformed back. Note that variance that was due to order of story presentation and was later eliminated by the regression was not removed from these means; the subsequent analyses of variance (ANOVAs) were performed on residualized data that did not contain that variability. In what follows, $F_1$ refers to tests against an error term based on subject variability, and $F_2$ refers to tests against an error term based on item (story) variability. As the means illustrate, when probes followed inference stories, they were identified more quickly than when they followed no-inference stories. A $2 \times 2$ mixed design ANOVA indicated that this difference was reliable by subjects and by items, $F_1(1, 45) = 10.07, p < .01, MSE = 3.54$; $F_2(1, 22) = 12.29, p < .01, MSE = 2.63$. On average, participants got 83% of the comprehension questions right.

The lexical-decision latencies associated with nonword targets were analyzed to investigate whether a response-speed advantage was afforded to targets simply because they followed an inference story (and was not due to any inference sanctioned by the story). If latencies for nonword targets that followed inference stories were significantly faster than latencies for those that followed no-inference stories, then the parallel finding with word targets reported in this experiment would be impeached. An inference versus no-inference analysis for nonwords, however, yielded $F < 1$ for this experiment. Thus, there is no reason to suspect that spillover effects with either type of material can account for the results of Experiment 1a.

Discussion

These results are consistent with the model’s prediction that inferences such as or-elimination are made spontaneously by readers at the moment both premises are simultaneously available. Lexical-decision targets were identified more quickly as words when they followed stories that permitted the reader to infer a semantic associate of that word than when they followed an otherwise identical story that did not sanction that inference. This differential priming strongly suggests that participants were in fact drawing the propositional-logic inference available from the information in the story, that they were doing so immediately following the presentation of the second premise—and that these inferences were being made in the absence of any demand for local coherence. As noted previously, the texts did not require readers to draw the inference to obtain a coherent representation of these texts. Thus, these results support the model’s prediction about the conditions under which the or-elimination inference will be made.

As mentioned earlier, one issue often raised in priming research that uses the lexical-decision task is context-checking. Context-checking is a concern because some studies have shown that the probability that the prime and target are related can affect lexical-decision latencies (e.g., Balota & Chumbley, 1984; Forster, 1981). Thus, if the target is differentially related to the two versions of the text, then the lexical decision can be promoted in the version that is more similar to the target (usually the inference version). The problem is that this kind of postlexical process may yield results indistinguishable from those produced by priming due to inferences.

Context-checking was relatively easy to control for in the present research because, unlike many pragmatic inferences, these logical inferences are not driven by context. This point is perhaps best made by Lea et al.’s (1990) Experiment 3 in which participants were presented with abstract premises ($ps$ and $qs$) and asked to write down whatever followed from them. In that situation, participants were deprived of content and context, yet 95% of the time they produced the modus ponens and or-elimination inferences. Clearly, context is not needed to trigger these inferences. Therefore, the inference/no-inference manipulation in the present research did not require text versions that described different contexts. Hence, context-checking is an unlikely source of the priming effect evident in the data from Experiment 1a. However, to rule out context-checking empirically, I performed a replication using naming instead of lexical-decision latencies as the dependent measure. Because the naming task does not involve a decision process, naming times are thought to be immune to postlexical context-checks that can affect lexical-decision latencies (see, e.g., Keenan, Golding, Potts, Jennings, & Aman, 1990, for a discussion of this issue).
Experiment 1b

Method

Participants. Twenty-four undergraduate psychology students participated to fulfill part of a course requirement at the University of Massachusetts. All participants were native speakers of English.

Materials. The only differences in materials from Experiment 1a were the targets; in accordance with text comprehension research that has used the naming task to detect forward inferences (e.g., Keefe & McDaniel, 1993; Murray et al., 1993; Potts et al., 1988), the words probes for experimental stories in Experiment 1b consisted of the inference concept itself, rather than an associate. For example, in the story about Tony deciding what to eat for breakfast, BREAD was used as the naming target, not the associate, BUTTER, as in Experiment 1a. Targets for the filler stories were words taken from those stories' first sentences or last sentences; thus, across the 60 stories, word probes were taken equally from the beginning, middle, and end of the passages, thereby depriving participants of any systematic pattern from which to predict the targets.

Procedure. The stories were presented to participants in one of six random orders. When participants read the sentence immediately preceding the probe, their key press blanked the screen, and a row of Xs appeared on the screen for 500 ms, followed by the probe. Participants named the target as quickly as possible. The naming times preceding the probe, their key press blanked the screen, and a row of Xs appeared on the screen for 500 ms, followed by the probe. Participants named the target as quickly as possible. The naming times consisted of the latency between the presentation of the target and the onset of the utterance. All other aspects of the procedure were the same as in Experiment 1a.

Results

The data for Experiment 1b were treated in the same manner as the data for Experiment 1a. Less than 6% of the data points were excluded as outliers. The mean naming time for no-inference targets was 526 ms; the mean naming time for inference targets was 507 ms. These means indicate that response times for targets in the inference stories were on average faster than those in the no-inference stories. A mixed ANOVA showed that this difference was significant both by subjects, \( F_1(1, 22) = 11.65, p < .01, MSE = 1.61 \), and by items \( F_2(1, 22) = 6.27, p < .05, MSE = 3.13 \). On average, participants got 86% of the comprehension questions correct.

Discussion

The naming results clearly replicate the lexical-decision findings of Experiment 1a; when participants were presented with a word to pronounce, their response times were reliably faster when that word could be logically deduced from the text than when it could not. These data show that the inference effect is sufficiently robust that it can be replicated with a different dependent variable. Most important, however, this result from a naming task rules out a context-checking explanation for the lexical-decision data from Experiment 1a and for the remaining lexical-decision data presented in this article.

Experiment 2

Some recent research has suggested that thematic focus can play a role in determining whether or not inferences are drawn by readers (e.g., Noordman, Vonk, & Kempff, 1992; Seifert, 1990; Trabasso, van den Broek, & Liu, 1988; Vonk & Noordman, 1990). For example, Noordman et al. found that participants could be induced to make a causal inference only if the inference was made relevant to the purpose of the reader. They manipulated relevance to the reader by providing information at the beginning of each text that focused attention on the inference in question.

In Experiment 1, story titles were included to provide participants with some of the thematic richness that clearly is missing in four-sentence stories but that is present in most ordinary texts. Previous research has shown that titles can promote the integration of information and has yielded evidence both from memory tasks (e.g., Kozminsky, 1977; Stern, Dahlgren, & Gaffney, 1991) and from on-line measures (e.g., Garrod & Sanford, 1983; Smith & Swinney, 1992). Given the current evidence that elaborative inferences are made only under rather constrained circumstances (e.g., Murray et al., 1993; Potts et al., 1988) and particularly in situations when the inference is in focus at the time of test or is tagged as especially relevant by the experimenter, it is important to generalize the elaborative-inference results from Experiment 1 to texts in which the focusing influence of a title is absent. Consequently, participants in Experiment 2 read passages that did not begin with titles. Therefore, Experiment 2 began to explore the issue of which factors that encourage deep reading can be dropped without affecting the spontaneity of deductive inference making.

Method

Participants. Fifty-eight undergraduate psychology students participated to fulfill part of a course requirement in Introductory Psychology at New York University. All participants were native speakers of English.

Materials. The materials differed from those used in Experiment 1 in that each story began with the words NEXT STORY rather than with the title that had been used in Experiment 1. Otherwise, the materials and design were identical to those in Experiment 1a.

Procedure. There were no differences in procedure from Experiment 1a.

Results

The data for Experiment 2 were treated in the same manner as the data for Experiment 1. As in the previous experiment, less than 1% of the data points were excluded as outliers. The mean response time for no-inference targets was 1,170 ms (0.87% errors); the response time for inference targets was 1,151 ms (1.2% errors). These means indicate that response times for targets in the inference stories were on average faster than those in no-inference stories. A mixed ANOVA showed that this difference was marginally significant by subjects, \( F_1(1, 56) = 3.50, p = .066, MSE = 3.18 \), and was not reliable by items, \( F_2(1, 22) = 2.10, p = .162, MSE = 2.59 \). Participants in Experiment 2 got 86% of the comprehension questions correct.

Lexical-decision latencies associated with nonword targets were analyzed to investigate any possible spillover effects. Again, the main effect of inference-set for nonwords produced
an $F < 1$. Thus, there is no reason to suspect that spillover effects were involved in the results.

A three-way ANOVA (Inference-Set x Group x Experiment) was computed for Experiments 1a and 2 to investigate whether there was a reliable difference between the priming effects in the title experiment compared with those in the no-title experiment. The Experiment x Inference Set interaction, however, failed to reach significance, $F(1, 101) = 1.45, p > .20, MSE = 3.34$; $F(1, 22) = 2.25, p > .10, MSE = 2.50$. None of the other interactions or main effects of interest were significant except, as expected, for inference set: $F(1, 101) = 13.32, p < .001, MSE = 3.34$; $F(1, 22) = 11.81, p < .01, MSE = 2.72$.

Discussion

The analyses of Experiment 2 data indicate that, although priming of inference targets is indicated by the inference group mean, the differences from no-inference controls are not as reliable as in Experiment 1. However, when the data from Experiments 1a and 2 are compared, the absence of interaction between inference set and experiment indicates that priming (and therefore inference making) were not significantly different between Experiment 1a (title) and Experiment 2 (no title). Moreover, the combined data yield a strong main effect for inference set, which, together with the lack of interactions, suggests that, overall, participants were making the inferences and that removing the titles did not lead to appreciably different results. I revisit the question of what contribution the title makes to inference making after presenting the data from Experiments 3 and 4.

Experiment 3

The model of natural propositional logic proposed by Braine et al. (1984) specifies several inference forms that people make spontaneously in everyday situations. The present line of inquiry would be greatly strengthened by the investigation of more than one of these forms. Experiment 3 investigated modus ponens (e.g., if p then q; p -> q) with the same overall design and strategy used in Experiment 1a (see Appendix B for examples).

Method

Participants. Fifty-four undergraduate psychology students participated to fulfill part of a course requirement in Introductory Psychology at New York University. All participants were native speakers of English.

Materials. A total of 72 stories were presented in Experiment 3. They were designed in the same way as the stories in Experiments 1a and 2 were. In an effort to obtain more measures per participant, 4 more stories were used in the present experiment than were used in the previous experiments. Thus, 28 stories were written around the modus ponens form. Four example stories are presented in Appendix B. Forty-four filler stories (28 nonword and 16 nonlogic) were generated to balance the proportion of word-nonword stories and logic-nonlogic stories. Each story used in Experiment 3 began with a title.

Procedure. There were no differences in procedure from Experiment 1a.

Results

The data were treated in the same manner as in the previous experiments. Less than 1% of the data points were excluded as outliers. The mean response time for no-inference targets was 1,165 ms (0.79% errors); the response time for inference targets was 1,134 ms (0.66% errors). These means indicate that response times for targets in the inference stories were, on average, faster than those in no-inference stories. A mixed ANOVA indicated that this difference was statistically reliable by subjects and by items, $F(1, 52) = 8.66, p < .01, MSE = 3.28$; $F(1, 26) = 4.51, p < .05, MSE = 2.04$. On average, participants in Experiment 3 got 86% of the comprehension questions correct.

An analysis of lexical-decision latencies associated with nonword targets revealed that targets following no-inference stories were, on average, slightly faster than those following inference stories. This discrepancy is in the opposite direction from any sort of spillover hypothesis that could account for the results.

Discussion

The results of Experiment 3 indicate that lexical-decision targets were identified significantly faster when they followed stories that permitted participants to use modus ponens to infer a semantic associate of that word than when they followed control versions of the stories. These results, therefore, agree with those of Experiments 1 and 2 and indicate that the inference forms specified in the model are made spontaneously by readers at the moment the necessary premises are simultaneously available and that they are made in the absence of any demand for local coherence.

Experiment 4

Experiment 2 investigated whether evidence for inference making from the or-elimination form in Experiment 1 could be reproduced with stories that did not furnish the thematic focus provided by a title. The results of that experiment suggested that there was still evidence of priming, even though the effect was diminished somewhat when the titles were removed. Experiment 4 was conducted to investigate whether the inference effect for modus ponens could be generalized to stories that did not begin with a title.

Method

Participants. Fifty-six undergraduate psychology students participated to fulfill part of a course requirement in Introductory Psychology at New York University. One participant was replaced because she exhibited an extremely slow reading rate and poor comprehension. All participants were native speakers of English.

Materials. The materials differed from those used in Experiment 3 in that each story began with the words NEXT STORY rather than with the title that had been used in Experiment 3. Otherwise, the materials and design were identical.

Procedure. There were no differences in procedure from the previous experiments.
Results

The data received the same treatment as in the previous experiments. Less than 1% of the data points were excluded as outliers by the trimming procedure. The mean response time for no-inference targets was 1,123 ms (0.50% errors); the response time for inference targets was 1,108 ms (0.75% errors). These means indicate that response times for targets in the inference stories were on average faster than those in no-inference stories. A mixed ANOVA showed that this difference was significant by subjects, $F_1(1, 54) = 4.25, p < .05, MSE = 2.31$, though not by items, $F_2(1, 26) < 1, MSE = 2.46$. Participants, on average, got 86% of the comprehension questions correct in Experiment 4.

The difference between the priming effects in Experiment 3 and 4 was assessed by looking at the Experiment $\times$ Inference Set interaction, which was not significant, $F_1(1, 106) < 1, MSE = 2.78; F_2(1, 52) = 1.00, MSE = 2.59$. The main effect for inference set was significant by subjects, $F_1(1, 106) = 12.90, p < .001, MSE = 2.78$, and was marginal by items, $F_2(1, 26) = 3.77, p = .063, MSE = 1.91$. None of the other interactions or main effects of interest were significant.

Discussion and Further Analyses

The results of Experiment 4 indicate that words associated with the consequent proposition of modus ponens inferences were identified faster when they followed stories that permitted inferences than when they followed otherwise identical control stories that did not. Analyses comparing Experiment 3 with Experiment 4 failed to reveal a difference in the priming effects between the two studies. This result, together with the large main effect for inference set produced by the combined data from the two experiments, suggests that participants were making the modus ponens inference while reading these stories and fails to yield evidence that the titles played a significant role in that result.

To investigate whether there was significant priming in the absence of story titles, I combined the data from the two no-title studies (Experiments 2 and 4) in a single ANOVA, which revealed a statistically reliable difference of inference set for subjects, $F_1(1, 112) = 7.80, p < .01, MSE = 2.70$. The comparable difference for items was not significant, $F_2(1, 50) = 1.87, MSE = 2.46$. Thus, as a set, the no-title data by subjects yield significant evidence of priming.

The potential title effect was pursued by comparing the combined data from the two title experiments with the combined no-title data. The resulting Title $\times$ Inference Set interaction also failed to reach significance, $F_1(1, 211) = 2.34, p > .10, MSE = 2.99; F_2(1, 48) = 3.15, p = .082, MSE = 2.55$. The combined data from the four lexical decision experiments yielded a strong result for inference set: $F_1(1, 211) = 26.66, p < .001, MSE = 2.99; F_2(1, 48) = 15.69, p < .001, MSE = 2.28$. Thus, when combined into a single analysis, the data from all four lexical-decision experiments yielded a straightforward result: (a) There was no reliable effect of title, and (b) the evidence for priming (and therefore inference making) was easily significant. These conclusions hold both for analyses by subject and by item.

General Discussion

Overall, the results of these four experiments support the prediction that the propositional-logic inferences specified in the model are made spontaneously by readers at the moment both premises are simultaneously available. Data from both the title and the no-title experiments produced significant results, though the no-title effects were slightly smaller.

The question arises as to what role the titles played, if any, in the priming effects (and therefore in inference making) investigated in the present research. A possible role for titles is suggested by the somewhat smaller difference between the inference and no-inference means in the no-title experiments as compared with their title counterparts. The statistical analyses reported earlier, however, argue in three ways that participants were in fact drawing inferences even when the stories did not have titles: (a) Experiment 4 (modus ponens, no title) produced a significant difference between inference and no-inference stories; (b) an analysis of the no-title data revealed a highly reliable ($p < .01$) difference between inference and no-inference conditions; and (c) three separate analyses failed to detect a difference between the priming effects in the title versus the no-title experiments. Finally, as noted previously, when the data from all four lexical decision experiments were combined into a single analysis, the main effect for inference set yielded an F greater than 25 (and Experiment 1b's naming data replicated this significant main effect). It appears, therefore, that participants in these experiments were making the inferences under investigation and that the presence or absence of titles did not play a significant role in this result.

The present data, however, do not rule out the possibility that discourse factors other than the availability of premises may be involved in producing model inferences. Other investigations have found that the extent to which an inference concept is supported by the surrounding text, or is in focus at the time of test, is critical at least to detecting the inference experimentally and is perhaps critical to the inference being made at all (e.g., McKoon & Ratcliff, 1990; Murray et al., 1993; E. J. O'Brien et al., 1988; Till, Mross, & Kintsch, 1988). Therefore, it would seem reasonable to expect that whether or not the reasoning program of Braine et al.'s (1984) model is engaged, and an inference produced, may well depend on some discourse factors—even if thematic focus as directed by titles is not reliably one of them. A major goal for future research is to better specify the relationship between the model and discourse elements that are required to engage the reasoning program.

What else could account for the results reported here? If participants were not systematically making inferences more often in inference stories than in no-inference stories, what else could have produced the results in these experiments? As discussed earlier, the naming data of Experiment 1b rule out
context-checking as a viable explanation for the present data. Perhaps participants adopted strategies based on patterns in the materials. For example, one might inquire about the role played by the comprehension questions. These questions were included to ensure that participants were reading and comprehending the texts; without them, participants in a required experiment might not read the texts at all. The questions also served to focus participants' attention on the stories and to encourage them to process the meaning of the stories deeply. As mentioned earlier, the collection of questions participants read did not follow any particular pattern; approximately one fourth of the questions asked about something stated explicitly in the text; about one fourth of the questions required a logical inference; and the remaining questions (approximately 50%) were simple elaborations from the text. Thus, it is unlikely that the comprehension questions specifically primed participants to look for logical inferences. Another possible source of reading strategies might be the number of stories that permitted a logical inference. If the majority of texts require the reader to make a particular kind of inference, then a concern would be that participants might notice the pattern and adopt strategies accordingly. In the present work, however, none of the texts required the readers to make a logical inference, and fewer than half of the stories (approximately 40%) even allowed for such an inference. In sum, the experimental materials provided few systematic cues to promote strategic processing.

Participants made remarkably few errors on the lexical-decision task; about 1% of their responses over the four experiments were errors. Such low error rates suggest that participants may have been trading off speed for accuracy and may help explain why the response times were on average somewhat slower than is typical for lexical decisions and why accuracy was somewhat higher. For example, in the forward-inference study conducted by Potts et al. (1988), lexical-decision latency averaged over predicted inference and control conditions was 975 ms with 5.4% errors, compared with an average of 1,148 ms and 1.3% errors in the present study. Thus, although participants were somewhat slower than average in their response times, they also were notably more accurate.

The issue of speed–accuracy trade-off raises the question of whether the difference in response times between inference conditions in this research might be due to differential response trade-offs for the two conditions. That is, many more errors on targets following inference stories might indicate that participants had used a response strategy for inference targets that evoked a different speed–accuracy trade-off (one that favored speed over accuracy) than the one they had used for no-inference targets. If this is true, then quicker latencies for inference targets could be due to speed–accuracy trade-off and not to priming, as is claimed. Error rates in the present research, however, are both extremely low and in the opposite direction for such a trade-off; over the four lexical-decision experiments, there was an average of 1.1% errors on inference targets as compared with an average of 1.3% for no-inference targets. Furthermore, the naming results of Experiment 1b show that the inference effect can be obtained with a dependent measure that does not involve a decision process. Thus, the results of these experiments cannot be explained by a speed–accuracy trade-off.

As discussed earlier, other mental-logic models have been proposed, and the present study was not designed to distinguish among them. The data presented here are compatible with any mental-logic model that defines a circumscribed set of inferences that includes modus ponens and or-elimination (e.g., Rips, 1994), provided that the model predicts that the inferences will be made spontaneously in discourse processing. However, not all of the present data are compatible with the current mental-models account of propositional reasoning (Johnson–Laird, Byrne, & Schaeken, 1992). Specifically, mental-models theory would make different predictions for the modus ponens stories than would the mental-logic models proposed by Braine and colleagues (e.g., Braine & O'Brien, 1991; Braine et al., in press; Braine et al., 1984), and by Rips. According to the psychological algorithm of the mental-models theory, there should be no difference in activation of the inference concept between the control and inference versions of the stories, because either version produces a model containing that concept at the time of test. Thus, mental-models theory would predict no difference between inference and no-inference stories. As we have seen, however, the reliable difference between inference and control versions of the modus ponens stories in Experiments 3 and 4 is inconsistent with that prediction.

Two reasons are often cited as to why forward inferences are not made during reading: (a) It is not efficient to infer what will likely be given in the text; and (b) there is the risk, and the associated costs, that one will draw the wrong inference (Murray et al., 1993). Logical inferences—as opposed to most pragmatic inferences—are deterministic rather than probabilistic in that if the premises are true, then the conclusion (inference) must be true. Thus, the second reason cited above may not hold, or hold as strongly, for logical inferences such as modus ponens and or-elimination; the measure of certainty afforded by logical inferences may explain why the present experiments yielded evidence for forward inferences whereas some investigations of pragmatic inferences have not.

The inferences studied here were based on premises explicitly presented by the text, but that would not always be the case in ordinary discourse. Consider the following example:

Bob was asking Jane about her new car.
“Which type of transmission did you get?” asked Bob.
“Well, you know I wanted to get a manual transmission, but my husband doesn’t know how to drive stick, so we couldn’t get the manual transmission.”

Inference: Jane got an automatic transmission.

In this case, Jane is able to rely on Bob’s knowledge that car transmissions are either manual or automatic so that she can focus on what she did not get, and perhaps her disappointment about that, and be assured that she has provided sufficient information to answer Bob’s question (cf. Sperber & Wilson, 1986). In this example, therefore, the major premise is imported from world knowledge. The model does not distinguish between information presented explicitly in a discourse and information that is accessed from memory; as long as the premise information is somehow available, the model predicts that the inference will be made.
The experiments presented here follow from investigations into the role of propositional-logic inferences in text comprehension that were begun by Lea et al. (1990). That study established that readers of short texts have little or no difficulty correctly making the inferences predicted by the model and that the inferences are drawn routinely enough that, immediately after reading the stories, readers often find it difficult to distinguish between information that has been presented explicitly in the text and information they have inferred. The present research extended those findings and conclusions by adopting a methodology that could investigate on-line several essential components of Braine et al.'s (1984) model, namely, the direct-reasoning routine and two of the schemas it operates. As described earlier, the direct routine works spontaneously to draw inferences and is hypothesized to be universally available. However, the conclusions from the present article are not specific to Braine et al.'s model because the aspects of the model specifically relevant to these experiments are shared with some other mental-logic models (e.g., that of Sperber & Wilson, 1986, and the recent model of Rips, 1994)—although, as far as I know, Rips has not discussed a possible application of his model to text comprehension. This study is the first on-line test that readers of text make the modus ponens and or-elimination inferences immediately on encountering the requisite premises in a purely elaborative situation. Consequently, the results presented here constitute the first on-line empirical support for that prediction regarding inference making in text. This article also addresses issues important to current work in discourse processing.

Specifically, the question of when an inference is made in the course of text comprehension is central to several theories of discourse processing. For example, McKoon and Ratcliff have advanced a minimalist hypothesis according to which readers make relatively few inferences while they read (McKoon & Ratcliff, 1992). Under the minimalist hypothesis, in the absence of special strategies, readers make inferences under only two conditions: when the inference is required for local coherence or when the inference is based on information that is readily available. On the one hand, the logic model's condition that both premises are simultaneously present in working memory is consistent with the readily available criterion of the minimalist position. Thus, the results of the present experiments would be compatible with minimalist expectations. On the other hand, however, it seems strange to say that elaborative logical inferences in locally coherent passages are predicted by the minimalist theory simply because the premises are readily available; presumably, there are uncountable potential inferences based on readily available information that are not made in any given text, and the minimalist approach alone provides no means to predict which will be made and which will not. Hence, the minimalist hypothesis alone could not have predicted these results.

In contrast to the minimalist position is a constructionist theory recently defined by Graesser et al. (1994). Briefly, the constructionist position depicts a relatively active reader who attempts to construct a meaningful representation of the text that (a) addresses readers' goals, (b) is coherent both locally and globally, and (c) explains why actions, events, and states are mentioned in the text. The underlying principle of this view is search (or effort) after meaning by comprehenders. The constructionist theory attempts to account for the knowledge-based inferences that are constructed when readers comprehend narrative text. These inferences are generated when knowledge structures in long-term memory are activated by pattern recognition processes and part of that stored information is encoded in the meaning representation of the text. According to these constructionists, most of these knowledge structures contain contextually rich information that is grounded in experience, for example, a restaurant script. Graesser et al., however, do allow for generic or abstract knowledge structures, such as the schemas specified in the present logic model. Accordingly, this constructionist theory is in principle compatible with the mental-logic schemas in Braine et al.'s (1984) model. It is unclear, however, whether the two theories would produce the same predictions regarding the inferences in the present study. The constructionist position, as articulated by Graesser et al., emphasizes the role played by (a need for) coherence and downplays the contribution of forward, elaborative processes: "The [constructionist] theory predicts that readers do not normally construct inferences that forecast future episodes in the plot" (p. 372). As noted earlier, the logical inferences investigated in this article were deductive elaborations about what will happen next in the story. Therefore, the compatibility of the present results with the constructionist position appears to depend on whether or not these forward logical inferences constitute future episodes according to Graesser et al.

The construction-integration model proposed by Kintsch (1988) provides perhaps the most promising comprehension framework in which to place model inferences. According to this model, discourse comprehension proceeds in a two-stage fashion. In the first stage, a text base is constructed from the linguistic input as well as from the comprehender's knowledge base; many elements are activated in this phase. An integration process is then used to strengthen the contextually appropriate elements and inhibit the inappropriate ones. Although the construction-integration model provides a detailed account of how a proposition from the knowledge base is contacted, it does not address how an abstract schema is activated. However, Kintsch did consider how sets of objects can be represented by a propositional schema. For example, he discussed a part-whole schema in the process of describing how an arithmetic word problem is solved. It follows that propositional-logic schemas such as those in the logic model could, in principle, exist quite comfortably within a discourse-comprehension framework such as Kintsch's construction-integration model. Indeed, to be fully specified for use in discourse, the logic model might well benefit from many of the comprehension procedures afforded by a model such as Kintsch's.

To summarize, although the results of the current investigations are in line with the minimalists' rather broad easily available criterion, the logic model generally is more compatible with Graesser et al.'s (1994) constructionist position and with Kintsch's (1988) construction-integration model. Despite these similarities, however, none of these comprehension models alone is adequate to account for the data presented here.
It is worth noting that recently there has been a call for converging evidence of a specific sort for inference making. Graesser, Trabasso, and their colleagues (e.g., Magliano & Graesser, 1991; Suh & Trabasso, 1993) have advocated a three-pronged approach for the assessment of inferences in text. According to this strategy, three procedural criteria are met: (a) a discourse model provides a set of specific a priori inferences that readers could make during comprehension; (b) verbal protocol data are collected to evaluate whether people make the inferences predicted by the model in the places where the model predicts they will be made; and (c) less obtrusive on-line measures are used to provide convergent validation of Procedures a and b. The data presented here, together with that of Lea et al. (1990), meet these three criteria. The model makes specific predictions regarding when model inferences will be made in text; thus, the first criterion is satisfied. The second criterion is met by the protocols Lea et al. elicited from participants in their Experiment 3, in which participants were given premises and asked to write down whatever followed from them. Finally, the third criterion is satisfied by the nonobtrusive on-line measures obtained in the semantic-priming experiments presented here.

To conclude, there is converging evidence that logical inferences of the sort specified in the mental-logic model of Braine and his colleagues are part of the ordinary discourse processing that people readily engage in every day. Previous work has shown that people are very accurate at making these inferences, that readers make them easily enough that they often do not realize that they are making inferences, and the present research provides critical data supporting the model’s prediction that these inferences are made on-line independently of any demand for coherence. Clearly, any complete theory of discourse processing must include some account of how propositional-logic inferences contribute to comprehension. The inference-schema model discussed here provides a precise hypothesis about how such inferences are made in text.

References


Appendix A

Four Example Stories From Experiments 1 and 2 and Their Lexical Decision Probes

The naming probes used in Experiment 1b are in square brackets. Story versions (inference vs. no inference) differed only in the text's fourth sentence (either Inf or No Inf).

Dressing for the Weather

Paula was getting dressed for work and asked her roommate, Donna, if she knew what the weather would be.

"I heard it on the news last night and it didn't sound good," said Donna, "they said it would either be rainy or they said it would be cold—I can't remember which."

"Hmm, those are two very different dressing conditions," Paula thought to herself, "I wish she could remember which it was."

Inf: "Ok I just called the weather," shouted Donna, "and it's not going to rain."

No Inf: "Well, whatever happens, I hope I don't get caught in the rain."

Choosing Cars

Ben was at the amusement park watching the bumper cars and waiting his turn.

When Ben’s turn came, the attendant said “You can take either the fast car over there or the fancy red car over here.”

Ben had been eyeing the look and performance of each of the cars, but now he had to make a quick decision.

Inf: "Forget the red car!” he exclaimed.

No Inf: "That red car looks sleek out there,” he thought to himself.

SLOW [FAST]

The Model

Chris, the director of the arts program at State University, needed to come up with a male model for the drawing class.

"I could hire a man to model for the class,” he said to his assistant, "or I could use a statue instead."

"Don’t the students need to draw the body in several poses?” asked the assistant.

Inf: "You're right about that,” the director replied, “I can’t use a statue for the class."

No Inf: “Not necessarily,” the director replied, “I think the class should vote on this matter of the statue.”

WOMAN [MAN]

Choosing a Hair Color

Linda finally decided to be daring and to get a dye job done on her hair.

According to her hairdresser, Linda’s complexion would look good with either a dark shade or a red shade.

She let her hairdresser, Yvette, make the final decision.

Inf: “Well,” said Yvette, “I’m sure you want to look different from everybody else this year, so we’re not going to dye your hair red.”

No Inf: Linda flipped through a few hairstyle magazines and saw a lot of people with red hair.

LIGHT [DARK]

Appendix B

Four Example Stories From Experiments 3 and 4 and Their Lexical Decision Probes

Story versions (inference vs. no inference) differed only in the text’s fourth sentence (either Inf or No Inf).

The Incomplete Shirt

George was distressed to find that his favorite shirt was missing a button.

“If I can’t find another shirt to wear on my date tonight,” George thought, “then I had better find a needle.”

George frantically looked through his closet to find another shirt to wear.

Inf: But George couldn’t find another shirt to wear because all of his other shirts were dirty.

No Inf: George realized that he had better figure out which shirt to wear soon because he was late.

THREAD

Environmental Protector

Fred, the forest ranger, was trying to decide which National Park to work for.

He knew that if he decided on the one in Alaska, then he would be working to protect the eagles.

Fred discussed his options carefully with a friend who was a veteran of the forest service.

Inf: After much consideration, Fred decided to work for the park in Alaska.

No Inf: Fred wondered whether the park in Alaska would be too cold for his liking.

BIRD

Sniffing Out the Goods

Special Investigator Evans was trying to crack a Cuban cigar smuggling ring.

As he entered the dark warehouse he whispered into his walkie-talkie: “If the dog starts to bark, then we know there’s tobacco around.”

“Ok, I’ll be listening,” replied his partner, Newstead, who was monitoring the situation from their unmarked patrol car.

Inf: There was quiet for a few minutes, then the silence was broken by the sound of Evans’s dog barking.

No Inf: There was quiet for a few minutes, and Newstead wondered whether he was close enough to hear the sound of the dog barking.

SMOKE

(Appendix continues on next page)
The Reconciliation Gift

Russell had had a fight with his girlfriend, and was eager to make up with her.
"If she will let me see her today, then I'll bring her a bouquet of blossoms from my garden," he thought.

Russell called and called her, but for hours there was no answer.

Finally she picked up the phone, "Ok, come on over," she said.

He sat staring out the window wondering whether he’d get to see her.

Call for Nominations

The Publications and Communications Board has opened nominations for the editorships of the Journal of Experimental Psychology: Animal Behavior Processes, the "Personality Processes and Individual Differences" section of the Journal of Personality and Social Psychology, the Journal of Family Psychology, Psychological Assessment, and Psychology and Aging for the years 1998-2003. Stewart H. Hulse, PhD; Russell G. Geen, PhD; Ronald F. Levant, EdD; James N. Butcher, PhD; and Timothy A. Salthouse, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 1997 to prepare for issues published in 1998. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees.

To nominate candidates, prepare a statement of one page or less in support of each candidate and send to the attention of the chair of the appropriate search committee. Search committee chairs are

- Joe L. Martinez, PhD, for JEP: Animal Behavior Processes. Members of the search committee are Russell M. Church, PhD; Michael Domjan, PhD; Michael S. Fanselow, PhD; and William D. Timberlake, PhD.

- David L. Rosenhan, PhD, for the "Personality Processes and Individual Differences" section of the Journal of Personality and Social Psychology. Members of the search committee are Nancy E. Cantor, PhD; Susan Fiske, PhD; Oliver John, PhD; and Mark Snyder, PhD.

- Carl E. Thoresen, PhD, for the Journal of Family Psychology. Members of the search committee are Arthur M. Bodin, PhD; James H. Bray, PhD; Lucia Gilbert, PhD; John M. Gottman, PhD; and Howard A. Liddle, EdD.

- Hans H. Strupp, PhD, for Psychological Assessment. Members of the search committee are Lee Anna Clark, PhD; Ken Pope, PhD; M. Tracie Shea, PhD; and Auke Tellegen, PhD.

- Lyle E. Bourne, PhD, for Psychology and Aging. Members of the search committee are Carol Barnes, PhD; Alfred Kaszniaè, PhD; M. Powell Lawton, PhD; and Michael A. Smyer, PhD.

Address all nominations to the appropriate search committee at the following address:

Lee Cron
P&C Board Search Liaison
Room 2004
American Psychological Association
750 First Street, NE
Washington, DC 20002-4242

First review of nominations will begin December 11, 1995.