



HHS Public Access

Author manuscript

Brain Lang. Author manuscript; available in PMC 2015 March 18.

Published in final edited form as:

Brain Lang. 2000 February 1; 71(2): 310–336. doi:10.1006/brln.1999.2268.

Right and Left Hemisphere Cooperation for Drawing Predictive and Coherence Inferences during Normal Story Comprehension

Mark Jung Beeman,

Section of Cognitive Neuroscience, Department of Neurological Sciences, Rush Medical College

Edward M. Bowden, and

Section of Cognitive Neuroscience, Department of Neurological Sciences, Rush Medical College

Morton Ann Gernsbacher

Department of Psychology, University of Wisconsin-Madison

Abstract

In three experiments, healthy young participants listened to stories promoting inferences and named inference-related test words presented to the right visual field-Left Hemisphere (rvf-LH) or to the left visual field-Right Hemisphere (lvf-RH). Participants showed priming for *predictive* inferences only for target words presented to the lvf-RH; in contrast, they showed priming for *coherence* inferences only for target words presented to the rvf-LH. These results, plus the fact that patients with RH brain damage have difficulty drawing coherence inferences and do not show inference-related priming, suggest that information capable of supporting predictive inferences is more likely to be initially activated in the RH than the LH, but following coherence breaks these concepts (now coherence inferences) are completed in the LH. These results are consistent with the theory that the RH engages in relatively coarse semantic coding, which aids full comprehension of discourse.

When people comprehend stories, they often make inferences — they assume that some events occurred in the stories, even though those events were not explicitly described. Consider what happens when people read or hear the *premise* that “The shuttle sat on the ground in the distance, waiting for the signal to be given.” If they later read or hear a sentence with a *coherence break*, that is a sentence that interrupts the logical chain of a story, such as “After a huge roar and a bright flash, the shuttle disappeared into space leaving clouds of smoke in its wake,” most people infer that the shuttle was launched. As established decades ago, once comprehenders make inferences, they include these inferences when recalling stories (Glenn, 1978; Paris & Lindauer, 1976; Baggett, 1975), and they have a difficult time distinguishing inferred information from information explicitly stated (Johnson, Bransford, & Solomon, 1973). This suggests that normal comprehenders incorporate some inferences into their representation of the discourse.

© 2000 Academic Press

Correspondence and reprint requests should be addressed to Mark Jung Beeman, Rush Cognitive Neuroscience Section, 1645 West Jackson Boulevard, Suite 450, Chicago, IL 60612, mbeeman@-rush.edu.

Some RH-damaged (RHD) patients have difficulty drawing inferences, suggesting that the right hemisphere (RH) may play an important role when comprehenders draw such inferences (Beeman, 1993; Brownell, Potter, Bihrlé, & Gardner, 1986). The following experiments explore the idea that concepts necessary for supporting inferences are more likely to be initially activated in the RH than in the LH and that at the coherence break inferences are more likely to be selected—activated strongly enough to reach awareness, to be processed further, or to be output, in the sense proposed by Allport (1987)—in the LH for incorporation into the discourse representation (Beeman, 1993).

PREDICTIVE AND COHERENCE INFERENCES

In principle, comprehenders could draw several types of inferences. Comprehenders could make inferences to embellish or elaborate on a story. They could draw inferences to predict upcoming consequences. For example, as soon as people hear that, *The shuttle sat on the ground, waiting for the signal*, they might predict that the shuttle will be launched. If so, they would be generating a *predictive* (or “forward”) inference. On the other hand, comprehenders might draw inferences only when necessary to resolve a coherence break: When faced with the contradiction that the shuttle, which was once on the ground, has now disappeared into space, comprehenders might infer that it was launched. This would be a *coherence* (or “backward” or “bridging”) inference. Coherence inferences fill in missing information to resolve a contradiction between a premise and a changed state. In this way, coherence inferences tie together the events of a story and improve the story’s coherence.

It is generally agreed that coherence inferences are reliably drawn by comprehenders (Duffy, 1986; Graesser, Singer, & Trabasso, 1994; Kintsch, 1998; O’Brien, Shank, Myers, & Rayner, 1988; Potts, Keenan, & Golding, 1988; c.f., McKoon & Ratcliff, 1986, 1992). For example, when people judge whether a statement is true, based on information from a story they just heard, they are just as fast to verify information assumed through coherence inferences as they are to verify explicitly stated information (Singer, 1979, 1980; Singer & Ferreira, 1983), suggesting that people make coherence inferences during comprehension. Similarly, after reading stories that promote coherence inferences, people rapidly recognize words related to those inferences, such as the word LAUNCH for the above excerpt. People recognize test words related to coherence inferences as rapidly as they recognize test words explicitly stated in the story (e.g., Potts et al., 1988), whether recognition is measured by lexical decision or naming latency (Keenan, Golding, Potts, Jennings, & Aman, 1989; Potts et al., 1988). In addition, people recognize words related to coherence inferences faster after comprehending sentences implying these inferences than they do after sentences that do not imply inferences (Potts et al., 1988).

In contrast, there has been some debate about whether predictive inferences are drawn (see, e.g., Fincher-Kiefer, 1995; Garrod, O’Brien, Morris, & Rayner, 1990; Graesser et al., 1994; Keefe & McDaniel, 1993; McKoon & Ratcliff, 1992; Murray, Klin, & Myers, 1993; St. George, Mannes, & Hoffman, 1997; Singer, Graesser, & Trabasso, 1994; Whitney, Ritchie, & Clark, 1991). Intuitively, it seems likely that people attempt to predict the outcomes of stories in which they are very interested, and sometimes a person can finish another person’s sentence, which must require a predictive inference.

However, the evidence that people draw predictive inferences has been equivocal, at least until recently. Many studies have found that predictive inferences are not drawn, are only partially drawn, or are drawn only under restricted circumstances (Calvo & Castillo, 1996; McKoon & Ratcliff, 1992; Noordman, Vonk, & Kempff, 1992; Potts et al., 1988; Singer, 1979, 1980; Singer & Ferreira, 1983). Others suggest that various factors — story variables (e.g., Keefe & McDaniel, 1993; O'Brien et al., 1988) or individual differences in comprehenders — can mediate the likelihood that comprehenders draw predictive inferences. During long stories in which the inferable event remains in focus, comprehenders name inference-related test words faster than unrelated test words — that is, they show inference-related priming (Murray et al., 1993). Individual differences among comprehenders, such as in working memory capacity, also affect the likelihood of drawing predictive inferences (Whitney et al., 1991). For instance, in an Evoked Response Potential (ERP) paradigm, comprehenders with high working memory capacity show signs of having made predictive inferences, whereas comprehenders with low working memory capacity do not (St. George et al., 1997).

It is possible that comprehenders draw inferences with varying probability: Because, by definition, coherence inferences are necessary for comprehension, listeners are very motivated to make these inferences. In fact, even with coherence inferences, comprehenders are more likely to show evidence of drawing inferences when the text is more easily understood, uses more familiar terms, and uses familiar referential synonyms to connect the concepts linked by inferences (e.g., as in narrative rather than in expository texts (Singer, Harkness, & Stewart, 1997). Comprehenders may make predictive inferences only when very interested and when they have enough processing resources available to do so. It is also likely that several steps are involved in drawing an inference, and comprehenders may sometimes complete only some of the steps, without completing all of them. It has been proposed that comprehenders may need to both *activate* an inference and *select* it (Beeman, 1993), or activate it and *incorporate* it into their discourse representation in order to completely draw the inference (Bramucci, Budd, & Whitney, 1995; Calvo & Castillo, 1996; Keefe & McDaniel, 1993). While reading or listening to a story, comprehenders may activate — that is to say, think about at some level — many potential inferences that are not explicitly stated in the text being comprehended, most of which will be predictive or elaborative. This may mean merely activating the idea that A SHUTTLE CAN BE LAUNCHED, rather than the idea that this particular shuttle will be launched into space. Such initial activation could be considered a partial predictive inference (McKoon & Ratcliff, 1986, 1989, 1992). In contrast, comprehenders may incorporate into their representation only those inferences that are most highly activated and/or most strongly linked to multiple parts of the text. Usually, these are coherence inferences. Presumably, for inferences which turn out to bear no relevance to the subsequent story, let alone become necessary for coherence, the activation eventually decays (Keefe & McDaniel, 1993).

Therefore, whether one believes that comprehenders draw predictive inferences may depend on the criteria used to define “drawing an inference.” By the activation criterion, comprehenders may make predictive as well as coherence inferences, but by the incorporation criterion, comprehenders may make mostly (or only) coherence inferences.

The experiments reported below used a cross-modal priming paradigm to assess activation of inference-related information during comprehension, rather than directly addressing which inferences are incorporated.

RH CONTRIBUTIONS TO DRAWING INFERENCES

As noted earlier, some RHD patients (generally, studies have used stroke patients with large lesions in the RH and compared these with age-matched controls; in Beeman, 1993, all patients had lesions that included the RH homolog to Wernicke's area) clearly have difficulty drawing inferences, as well as other subtle difficulties with complex discourse, despite passing most simple linguistic tests (see Beeman, 1998, for review). RHD patients have difficulty answering true/false questions about inferable information, despite answering questions about explicitly stated information as well as do age-matched controls (Beeman, 1993; Brownell, et al., 1986). These RHD patients are willing to accept inferences as true, but are not adept at choosing the most appropriate inference, which would tie together the changed state with the premise (Beeman, 1993). Importantly, RHD patients' difficulties are not from responding too literally (Beeman, 1993; Weylman, Brownell, Roman, & Gardner, 1989), as previously suggested in the literature (e.g., Winner & Gardner, 1977; Van Lancker & Kempler, 1987). Rather, RHD patients appear to be impaired at generating inferences that best integrate distinct parts of the discourse. Failing to draw inferences does have some advantage. When normal comprehenders hear short scenarios that promote inferences, they often mistakenly judge that they heard or read inferred sentences (e.g., Johnson et al., 1973). LH-damaged (LHD) patients also make such mistakes, but RHD patients are less prone to do so, most likely because they do not generate inferences in the first place (Grafman, Salazar, Vance, Weingartner, & Amin 1987; cf. McDonald & Wales, 1986).

Why do RHD patients have difficulty drawing inferences? Unlike age-matched control participants, while listening to inference-promoting sentences RHD patients show no evidence of activating semantic information related to the intended inferences; i.e., they show no priming for inference-related test words (Beeman, 1993). The lack of inference-related priming eliminates the possibility that RHD patients draw inferences, but forget them or do not connect them to their representations of explicit input. Rather, RHD patients fail to activate some semantic information necessary to draw coherence inferences.

RH COARSE SEMANTIC CODING

By presenting words to each hemifield, we can assess the activation of the intended inferences in each hemisphere somewhat independently. Even when the corpus callosum is intact, requiring an immediate response to test words presented to one hemifield can tap into the current processing state (e.g., activation) of the contralateral hemisphere. (Presentations must be rapid, faster than 200 ms, to preclude eye movements that would allow participants to foveate test words.) Qualitatively distinct patterns of priming, including crossover interactions (e.g., Beeman et al., 1994), demonstrate that left visual field-Right Hemisphere (lvf-RH) test words are not merely relayed across the callosum to be fully processed in the LH. Moreover, these priming asymmetries can occur for foveally (for review, Chiarello, 1998) or binaurally presented primes, even as long as 15 s after prime onset (Bowden &

Beeman, 1998), indicating that some parallel processing occurs even in people with connected hemispheres.

Comprehenders display qualitatively distinct patterns of semantic priming when processing words in the lvf-RH, relative to priming seen for test words presented to the right visual field-LH (rvf-LH) (for review, Beeman, 1998; Chiarello, 1998; Faust, 1998). The RH appears to maintain activation of distant semantic relations of words (Beeman et al., 1994; Chiarello, Burgess, Richards, & Pollock, 1990; Nakagawa, 1991) or multiple interpretations of ambiguous words (Burgess & Simpson, 1988), whereas the LH appears to activate close associates and a single interpretation.

One interpretation of these results is that when comprehenders encounter words in discourse, the LH engages in relatively *fine* semantic coding, strongly activating small semantic fields (i.e., only concepts closely related to the input, given the local context), whereas the RH engages in relatively coarse semantic coding, weakly activating large semantic fields (including concepts distantly related to the input). In this view, the meaning of a word consists of all of its relations, and probably does not vary across the hemispheres (which are viewed as having roughly equivalent lexicons). The interpretation in a given context comprises the semantic field, that is, the specific relations that are activated and remain activated, and this is what varies by hemisphere. Strongly focused activation allows the LH to quickly select the most relevant meanings of words, and connections between words, which can then be incorporated into a mental representation of the discourse. Thus, relatively fine semantic coding makes the LH extremely adept at most language tasks. In contrast, diffuse activation in the RH is not efficient for selecting a single meaning, as several distinct meanings or shades of meaning remain active, rather than one or a few representations being much more active than the others. Thus, the RH provides only a coarse interpretation of each word, and is not efficient for most language tasks. However, the larger (albeit weakly activated) semantic fields are more likely to overlap — that is, multiple words of the discourse will each weakly activate the same semantic representations — and this overlap conveys some advantages.

“Summation priming” data are consistent with the view that the RH is especially sensitive to weak but overlapping activation from distantly related words (Beeman et al., 1994; Beeman, Bowden, Hassenfeld, & Shivde, 1996). Summation primes are three-word primes (*shuttle*, *ground*, *space*) where each word is weakly related to the test word (*launch*). Participants benefit more from summation primes (presented simultaneously above, at, and below the fixation point) when responding to lvf-RH test words than when responding to rvf-LH test words, at least when controlled semantic processing is possible. In contrast, they show more direct priming — from one prime word (*scissors*) strongly related to the test word (*cut*) — than summation priming for rvf-LH test words, but at least as much summation as direct priming for lvf-RH test words (Beeman et al., 1994). This RH advantage in summation priming is strongest when each of the three prime words is very distantly related to the test word, and all three converge on the same meaning (Beeman et al., 1996).

Some neuroimaging data are also consistent with this description of RH coarse semantic coding and LH fine semantic coding. For instance, the RH shows increased PET signal

when participants comprehend metaphors, which require distant semantic relations to be noted (Bottini *et al.*, 1994). Also, increased signal occurs in the RH when people generate unusual uses of nouns compared to when they generate the first use that comes to mind, as measured both by ERP (Abdullaev & Posner, 1997) and by fMRI (Seger, Desmond, Glover, & Gabrieli, in press). Generating unusual uses of nouns presumably relies more heavily on information less closely related to the noun than on generating usual uses. Finally, fMRI signal increases primarily in the RH when subjects comprehend untitled rather than titled stories, which putatively places greater demand on recognizing distant semantic relations (St. George, Kutas, Martinez, & Sereno, 1998). Similarly, patients with RH damage have greater difficulty recalling story themes than do patients with LH damage (Hough, 1990) and depend more on presentation of a theme sentence for arranging sentences into a paragraph (Schneiderman, Murasugi, & Saddy, 1992).

When working in concert with fine semantic coding in the LH, coarse semantic coding in the RH might be very helpful for comprehending discourse that requires people to notice distant semantic relations, as is necessary to draw inferences (for review, Beeman, 1998). According to Keefe and McDaniel,

Predictive inferences are presumably only weakly associated to individual concepts in a sentence but become strongly activated when the activation from the individual concepts in a sentence combine during the integration process, (p. 461)

According to our view, concepts related to the inference are more likely to receive converging input from multiple words in the RH than in the LH, but following a coherence break the inference concept is more likely to be selected, via strong focused activation, in the LH. This is not to say that the hemispheres take turns in some serial form of processing. The crux of our theory is that both hemispheres process all information in parallel, with some sharing of information at critical times. We believe that a person's action or conscious thought is the product of simultaneous processing and activation of both hemispheres, supported by population coding — that is, the summed distributed activity of many thousands of neurons, without the need of an executive, homunculus, or grandmother cell. This population is divided over both hemispheres and is the product of both processing styles. Because we posit that the LH strongly activates a narrow field of information, most of which will also be weakly activated in the RH, the information activated in the LH tends to dominate the population code underlying selection of information for consciousness and responses.

Experiments 1 through 3 examined whether people comprehending stories had semantic activation capable of supporting inferences in a given hemisphere by assessing inference-related priming in the rvf-LH or in the lvf-RH at various points in the stories. If comprehenders activate inferences, then they should name (pronounce or read aloud) visual test words related to coherence inferences more rapidly (or accurately) than unrelated test words — that is, they will show priming. Across the three experiments, we examined inference-related priming at five test points (1 through 5), as noted in Table 1. Each of the following experiments tested inference-related priming at two test points: Experiment 1 tested priming at the Predictive point (1), the first clause boundary at which participants could reasonably make a predictive inference, and at the Coherence point (4), the first clause

boundary following the coherence break (at which point the inference would be a coherence inference). Experiment 2 also tested priming at the Coherence point (4) and at the Resolution point (5) at the following clause boundary, when we expected the inference to have been completed. Finally, to strengthen results from Experiment 1, Experiment 3 tested priming at a second Predictive test point (2), between the predictive and coherence sentences, and at the Transition test point (3), a few words into the next sentence, just as the coherence break began to develop. Thus, we will assess inference-related activation in each hemisphere, at points when the inferences would be either predictive or coherence inferences. The primary prediction is that inference-related priming will occur for lvf-RH target words at an earlier point than for rvf-LH target words.

EXPERIMENTS 1 AND 2

Experiments 1 and 2 were run simultaneously, randomly assigning participants to each Experiment. Priming was assessed at three test points, two in each experiment, with the middle point—the Coherence test point (4)—being tested in both experiments, to allow the possibility of replication. We predicted that at the Predictive test point (1) participants would show priming for inference-related test words presented to the lvf-RH, because the RH is more sensitive to converging distant semantic relations, as demonstrated with Summation priming (Beeman et al., 1994). We predicted less or no priming for test words viewed in the rvf-LH at the Predictive test point (1), given that certain RHD patients (with intact LH) have difficulty drawing inferences (Beeman, 1993; Brownell et al., 1986). By the time comprehenders reached the Coherence test point (4), we predicted that participants would show inference-related priming in the rvf-LH because at this point the inferences must be drawn to maintain coherence. In a pilot study using a similar test point, participants did show inference-related priming in the rvf-LH, but not in the lvf-RH. At the Resolution test point (5), we predicted priming in both hemi-fields because the inference should be incorporated into the discourse representation.

Method

Participants—Participants were 107 undergraduates who participated as partial fulfillment of a course requirement, randomly assigned to Experiment 1 or 2. All participants were right-handed and native speakers of American English. Nine participants were replaced due to poor performance (more than 2.5 SDs from the mean in question or naming accuracy or naming speed): 6 answered fewer than 75% of the true/false comprehension questions accurately; 1 correctly named only 52.5% of the test words; 2 responded slowly on naming trials. Data from two participants were lost and replaced due to experimenter error. The remaining 96 participants (48 in each experiment) correctly answered 90.1% ($SD = 5.75\%$) of the true/false questions and correctly named 92.4% ($SD = 9.93\%$) of the test words. Response times missed by the voice key or those exceeding 2.5 SDs from each subject's overall mean latency, were excluded from analyses.

Materials—Materials consisted of 20 stories, each promoting four coherence inferences in four paragraphs, for a total of 80 promoted inferences. One paragraph promoting an inference is given in Table 1. Each inference presented a premise in one sentence (which

suggested, sometimes only weakly, a predictive inference) and a changed state in the next sentence. The coherence break sentences immediately followed the premise sentences. We followed Kurtzman's (1990) recommendation and wrote the stories to avoid gender stereotypes (e.g., police officers and doctors were often women, and a man rather than a woman was rescued by a lifeguard). These changes were not so unusual that they disrupted comprehension (based on feedback from pilot participants). Because the stories often related tragic events (e.g., personal injuries, bombs exploding, and objects breaking), we wrote positive endings to many stories (e.g., no one was hurt in the fire). Half the stories were recorded by a female reader and the other half by a male reader. Stories varied in duration from 63 to 86 s with a mean duration of 75 s. The stories were digitized using a Macintosh Quadra 950 and SoundEdit Pro. Participants listened to the stories, played by the Superlab experiment program, through headphones connected to a Macintosh LC computer.

Eight words — four test words and four filler words — were visually presented during each story. Of the test words, two were related to the inference promoted by the paragraphs in which they occurred and two were taken from corresponding paragraphs of another story and therefore were unrelated to the paragraphs in which they occurred (balanced across material sets, each word was its own control). The inference-related test words were not closely related to individual words in the stories, according to association norms (Jenkins, 1970; Keppel & Strand, 1970; Marshall & Cofer, 1970) and two pilot studies. In one pilot study, participants (at least 12 for each test word) produced three associated to each test word (e.g., launch) and each potentially related word from the story (e.g., shuttle, space). If more than 20% of participants' responses indicated a relation between a test word and a word from the related story, the word in the story was changed. Several rounds of this type of study were performed. In another pilot study, eight participants read the stories and circled words in the stories that they thought may have been related to the test word. These words were eliminated. This is not to say that all the story words were completely unrelated to the test words — in fact, our theory suggests that distant semantic relations help activate the inferable concepts — merely that the story words were not closely related to the test words, by the same criteria used in most similar studies. The fillers were unrelated to any of the stories. All words were three to eight letters in length (mean = 4.675 letters; maximum of 97 pixels for experimental test words and 106 pixels for fillers).

The onset of the visual test words occurred at three different clause breaks within the premise and coherence break sentences, as illustrated in Table 1. The Predictive test point (1) was the first clause break after comprehenders could reasonably make the intended predictive inference. The Coherence test point (4) was the first clause break after the coherence break. The Resolution test point was the next clause break after that and the last before the end of the sentence. The Coherence test point (4) in this experiment was roughly equivalent to the test point used for RHD patients and age-matched controls in prior work (Beeman, 1993). The filler words occurred at unpredictable points in the story (i.e., not at clause boundaries). Computer presentation of the digitized stories and test words permitted precise temporal placement of the visual test words within the auditory stories.

All test words were presented in black on a white background on a 15-inch Apple RGB monitor so that the proximal end of each word (left end for words in rvf-LH, right end for

words in the lvf-LH) was 24 pixels from the center of the screen. Participants were seated 45 cm from the monitor so that the inner edge of the words appeared approximately 1.0° of visual angle from fixation, and the outer edge ranged from 2.5° to 3.5° of visual angle away from fixation. A pattern mask, consisting of letter fragments, followed the presentation of test words, to ensure that participants could not read any image persisting on the screen. The mask extended 3.5° of visual angle to both sides of fixation and about 4° vertically.

There were eight material sets to completely counterbalance hemifield of presentation of test word (rvf-LH or lvf-RH), type of test word (Related or Unrelated control), and test point (Predictive versus Coherence in Experiment 1, or Coherence versus Resolution in Experiment 2). Each participant saw 10 test words per condition, and each test word occurred in each condition across the eight material sets.

Procedure—Participants were tested individually and told that they had two tasks: (i) While listening to stories, they were to focus on the central plus sign, and at random intervals words would appear randomly to the right or the left side of that plus sign. They were to read these words aloud (name them), as quickly and accurately as they could. Naming (i.e., pronunciation) latencies are thought (e.g., Potts et al., 1988) to be less susceptible to postlexical mechanisms such as compatibility checking which may affect lexical decision latencies (basing a yes/no response on the relatedness of the target and the prime — in this case, the story). Responses were scored by the experimenter, and latencies were collected with a voice activated relay. (ii) To encourage comprehension, they were also told that following each story, they would be asked four true/false questions regarding the story. These questions related to facts and themes of the stories, but were not directly related to the intended inferences, so that participants would not adopt special comprehension strategies just to answer the questions (McKoon & Ratcliff, 1992). Participants listened to one practice story while pronouncing test words and subsequently answered questions to familiarize them with the procedure.

Participants listened to 20 stories each promoting four coherence inferences. For each story, a fixation point (+) was displayed in the center of the computer monitor for 1000 ms, and then the story began. While the story was being presented the fixation point remained on the screen except when a test word, filler word, or the mask was presented. Test words and filler words were visually presented laterally for 180 ms (it takes approximately 200 ms to make a saccade and refixate, so if subjects were fixated at the onset of the trial, the rapid presentation ensured that participants could not foveate the target words). It appears that participants did maintain fixation because, in Experiments 1–3, participants were able to read over 90% of 180-ms test words and showed the same rvf-LH advantage in accuracy and latency (especially for unrelated test words) typically obtained in visual hemifield paradigms. Hemifield, relatedness, and test point (Points (1) and (4) in Experiment 1, and Points (4) and (5) in Experiment 2) were factorially manipulated.

Experiment 1 Results

Data the mean correct naming latencies from 48 participants from 10 observations for each Hemifield × Test Point × Relatedness condition (F_1 analyses), are listed in Table 2 and

graphed as priming data (unrelated-related latencies) together with Experiments 2 and 3 results in Fig. 1. Alpha was set to $p < .05$. There were some uninteresting interactions with the material set, so the material set was included as a between-participants variable, but will not be further discussed. (Neither the pattern nor the reliability of results changed when analyses were collapsed across material sets.) An $8 \times 2 \times 2 \times 2$ between-within ANOVA was performed with material set as a between participant variable and hemisphere, test point, and relatedness as within participant variables. Another analysis examined the mean naming latencies for 80 test words (F_2 analyses) from 48 participants (6 for each material set rotating Test Point \times Relatedness \times Hemifield). Parallel analyses were performed on the accuracy data (reported in Table 2) but because participants were closer to ceiling in overall accuracy for rvf-LH than for lvf-RH test words, these analyses could be misleading and will not be discussed.

We begin with the planned analyses of the priming effects conducted for each hemisphere at each test point.¹ At the Predictive test point (1) participants named inference-related test words an average of 24 ms more quickly than unrelated test words when viewed in the lvf-RH, $F_1(1, 40) = 7.71$, $MSe = 13,278$, $F_2(1, 79) = 2.88$, $MSe = 20,612$, $p < .10$. In contrast, they named inference-related and unrelated test words equally quickly (mean difference of 5 ms) when viewed in the rvf-LH, $F_s < 1.0$.

At the Coherence test point (4), participants named inference-related and unrelated test words equally quickly when presented to the rvf-LH (mean difference = 10 ms), $F_1(1, 40) = 1.63$, $MSe = 2,460$, $p < .21$, $F_2 < 1.0$, and to the lvf-RH (mean difference = 6 ms) $F_s < 1.0$. Thus, as predicted, Experiment 1 found priming for predictive inferences with lvf-RH and not rvf-LH test words. However, the priming for coherence inferences was not reliable in either hemifield, contrary to the results of prior experiments (Beeman, 1990, 1993).

All three main effects were reliable. As expected, participants named rvf-LH test words (645 ms, $SE = 7.0$) more quickly than they named lvf-RH test words (673 ms, $SE = 7.3$) across test point and relatedness, $F_1(1, 40) = 41.27$, $MSe = 75,040$, $F_2(1, 79) = 30.96$, $MSe = 119,001$, thus producing the normal rvf-LH advantage for reading. Also as expected, participants named inference-related test words (653 ms, $SE = 7.2$) more quickly than they named unrelated test words (664 ms, $SE = 7.2$), $F_1(1, 40) = 6.13$, $MSe = 12,399$, $F_2(1, 79) = 6.91$, $MSe = 27,053$, thus producing the expected inference-related priming. Finally, participants responded more quickly at the Predictive test point (1) (mean = 654, $SE = 7.2$) than at the Coherence test point (4) (mean = 664, $SE = 7.2$), $F_1(1, 40) = 6.22$, $MSe = 9,068$, $F_2(1, 79) = 3.45$, $MSe = 14,621$. The three-way interaction was not statistically reliable, $F(1, 40) = 2.95$, $MSe = 2,860$, $p < .10$, $F_2 < 1.0$.

¹We also conducted two-way ANOVAs on hemisphere and relatedness at each test point. At the Predictive test point (1) there were main effects of hemisphere, $F_1(1, 40) = 13.79$, $MSe = 28,373$, $F_2(1, 79) = 6.63$, $MSe = 45,339$, and of relatedness, $F_1(1, 40) = 7.15$, $MSe = 10,005$, $F_2(1, 79) = 4.26$, $MSe = 20,432$. There was also a trend toward an interaction, $F_1(1, 40) = 2.81$, $MSe = 3,960$, $p = .10$, $F_2 < 1.0$. At the Coherence test point (4) there was a main effect of hemisphere, $F_1(1, 40) = 36.04$, $MSe = 47,944$, $F_2(1, 79) = 8.14$, $MSe = 75,584$, but not of relatedness, $F_1(1, 40) = 1.41$, $MSe = 3,300$, $p = .24$, $F_2(1, 79) = 2.37$, $MSe = 8040$, $p = .13$. There was no interaction, $F_s < 1.0$.

Experiment 2 Results

Experiment 2 examined participants' naming times for inference-related and unrelated test words presented at the Coherence (4) and Resolution (5) test points, which are listed in Table 2 and graphed as priming data (unrelated-related latencies) together with results from Experiments 1 and 3 in Fig. 1. Planned analyses of the priming effects were conducted for each hemisphere at each test point.² At the Coherence test point (4), participants named related test words an average of 23 ms more quickly than unrelated test words for rvf-LH presentation, $F_1(1,40) = 13.84$, $MSe = 12,285$, $F_2(1, 79) = 10.42$, $MSe = 25,883$. In contrast, they named inference-related and unrelated test words equally quickly (mean difference of 7 ms) for lvf-RH presentation, $F_s < 1.0$.

At the Resolution test point (5), participants named related test words an average of 23 ms more quickly than unrelated test words for rvf-LH presentation, $F_1(1, 40) = 7.82$, $MSe = 12,627$, $F_2(1, 79) = 5.37$, $MSe = 18,512$. They also named related test words 27 ms more quickly than unrelated test words for lvf-RH presentation, $F_1(1, 40) = 12.26$, $MSe = 17,794$, $F_2(1, 79) = 6.76$, $MSe = 27,641$.

All three main effects were reliable. As expected, participants named rvf-LH test words (649 ms, $SE = 6.4$) more quickly than they named lvf-RH test words (668 ms, $SE = 6.7$), $F_1(1, 40) = 18.66$, $MSe = 33,975$, $F_2(1, 79) = 17.24$, $MSe = 56,663$, thus producing the normal rvf-LH advantage for reading. Also as expected, participants named inference-related test words (648 ms, $SE = 6.4$) more quickly than they named unrelated test words (668 ms, $SE = 6.7$), $F_1(1, 40) = 21.03$, $MSe = 38,280$, $F_2(1, 79) = 15.18$, $MSe = 65,044$, thus producing the expected inference-related priming. Finally, participants responded more quickly at the Coherence test point (4) (mean = 650, $SE = 6.4$) than at the Resolution test point (5) (mean = 667, $SE = 6.8$), $F_1(1,40) = 25.19$, $MSe = 27,304$, $F_2(1,79) = 11.66$, $MSe = 45,731$. The three-way interaction was not statistically reliable, $F_1(1, 40) = 1.92$, $MSe = 2,360$, $F_2(1, 79) = 1.37$, $MSe = 5198$.

Combined Results of Coherence Test Point (4) from Experiments 1 and 2

Because the Coherence test point (4) was identical for Experiments 1 and 2, and the participants were selected simultaneously from the same pool, we combined the data and performed an additional 8 (material set) \times 2 (hemisphere) \times 2 (relatedness) ANOVA. (Initial analyses indicated that the variable Experiment (1 versus 2) had no effects and did not interact with other variables, and so data was collapsed across Experiment).

In the planned comparisons, participants named related test words 17 ms more quickly than unrelated test words when viewed in the rvf-LH, $F_1(1, 88) = 10.48$, $MSe = 12,871$, $F_2(1, 79) = 9.09$, $MSe = 21,809$, but only 7 ms more quickly when viewed in the lvf-RH, $F_1(1, 88) = 1.41$, $MSe = 2201$, $p < .24$, $F_2 < 1.0$. Thus combined data from Experiments 1 and 2

²We also conducted two-way ANOVAs on hemisphere and relatedness at each test point. At the Coherence test point (4) there was a main effect of hemisphere, $F_1(1, 40) = 13.62$, $p < .001$, $F_2(1, 79) = 11.81$, $MSe = 32,765$, a main effect of relatedness, $F_1(1, 40) = 9.62$, $p < .005$, $F_2(1, 79) = 5.06$, $MSe = 21,583$, but the interaction was not quite reliable, $F_1(1,40) = 2.34$, $MSe = 2,899$, $p = .13$, $F_2(1, 79) = 2.20$, $MSe = 6,498$, $p = .14$. At the Resolution test point (5) there were main effects of hemisphere, $F_1(1, 40) = 9.69$, $MSe = 15052$, $F_2(1, 79) = 6.78$, $MSe = 24,221$, and relatedness, $F_1(1, 40) = 17.75$, $MSe = 30,200$, $F_2(1, 79) = 12.52$, $MSe = 45,697$. There was no interaction, $F_s < 1.0$.

demonstrate inference-related priming for rvf-LH test words but not for lvf-RH test words at the Coherence test point (4). However, the interaction between relatedness and hemifield was not reliable, $F_1(1, 88) = 1.94$, $MS = 2,214$, $p < .17$, $F_2 = 2.62$, $MSe = 7,772$, $p = .11$.

The two main effects were also reliable: Participants named rvf-LH test words more quickly than lvf-RH test words, $F_1(1, 88) = 40.23$, $MSe = 63,706$, $F_2(1, 79) = 8.14$, $MSe = 32,140$, and named related test words more quickly than unrelated test words, $F_1(1, 88) = 7.80$, $MSe = 12,858$, $F_2(1, 79) = 4.36$, $MSe = 14,567$.

EXPERIMENT 3

Experiments 1 and 2 together demonstrated a RH advantage in inference-related priming at the Predictive test point (1), and a slight LH advantage in inference-related priming at Coherence test point (4). Although the LH priming at the Coherence test point (4) was not reliable in Experiment 1, it was reliable when the data from both experiments were combined and was reliable in a prior experiment using different materials when subjects made lexical decisions (Beeman, 1990). The goals of Experiment 3 were to replicate the RH advantage in inference-related priming at predictive test points and to get more information about the time course of inference-related activation in both hemispheres.

Method

Participants—Fifty-five undergraduates participated to partially fulfill a course requirement. All were right-handed and native speakers of American English. Seven participants were excluded due to equipment failure or poor performance (more than 2.5 SDs from the mean, in question or naming accuracy or latency; some participants failed multiple criteria): 3 correctly named fewer than 74% of the test words; 2 answered fewer than 78% of the true/false comprehension questions accurately; 2 showed reverse hemifield advantages for both latency and accuracy (lvf-RH responses were faster and more accurate); 2 had long naming latencies; and for one participant the microphone missed 30/80 naming responses. The remaining 48 participants correctly answered 89.2% ($SD = 4.5\%$) of the true/false questions and correctly named 89.4% ($SD = 6.14\%$) of the test words. Response times missed by the built-in voice key (4.37% of total responses), or exceeding 2.5 SDs from each subject's overall mean latency, were excluded from analyses.

Materials and procedure—Experiment 3 was the same as Experiments 1 and 2 except that latencies were collected with the built-in Macintosh microphone (via Superlab software); and, more importantly, test words appeared at two new test points, between those previously used. In Experiment 3, test words appeared at the Predictive point (2), the end of the premise sentence, or at the Transition test point (3), a few words into the coherence break sentence. At the end of the premise sentence, the promoted inference is still predictive, because no coherence break has occurred. By a few words into the next sentence, the coherence break has begun to occur (in some stories more than others), so at this point the inference could be called predictive, coherence, or somewhere in between depending on the story. Having test words appear at the end of a sentence is similar to having them occur at the end of a clause. However, having test words appear in the middle of a clause (Transition point 3) could differ from the other test points, if, for instance, ongoing text interferes with

responses to the test words. Therefore, we tried to locate these test points at commas or other natural breaks and when necessary inserted some silence into the story so that there was a pause at least 500 ms long as the test word appeared.

Results

Experiment 3 examined participants' naming times for inference-related and unrelated test words presented at the second Predictive (2) point and at the Transition test point (3), just as the coherence break unfolded. Mean latencies and accuracies are listed in Table 2, and graphed as priming data (unrelated-related latencies) together with Experiments 1 and 2 results in Fig. 1. Planned analyses of the priming effects were conducted for each hemisphere at each test point.³ At the Predictive test point (2) participants named inference-related test words an average of 43 ms more quickly than unrelated test words when viewed in the lvf-RH, $F_1(1, 40) = 13.29$, $MSe = 44,333$, $F_2(1,79) = 10.58$, $MSe = 55,652$. In contrast, they named inference-related test words only 7 ms more quickly than unrelated test words when viewed in the rvf-LH, $F_s < 1.0$. The difference in priming effects for lvf-RH and rvf-LH test words was reflected in a hemifield by relatedness interaction, reliable with participants as the random variable $F_1(1, 40) = 5.62$, $MSe = 15,878$, $F_2(1, 79) = 2.20$, $MSe = 9,757$.

At the Transition test point (3), participants named inference-related test words an average of 20 ms more quickly than unrelated test words when viewed in the lvf-RH, $F_1(1, 40) = 5.85$, $MSe = 9,923$, $F_2 < 1.0$, and named inference-related test words an average of 21 ms more quickly than unrelated test words when viewed in the rvf-LH, $F_1(1, 40) = 7.91$, $MSe = 10,732$, $F_2(1, 79) = 2.93$, $MSe = 16,974$. The hemifield by relatedness interaction was not reliable, $F_s < 1.0$.

All three main effects were reliable. As expected, participants named rvf-LH test words (668 ms, $SE = 9.14$ ms) faster than they named lvf-RH test words (701 ms, $SE = 9.74$ ms), $F_1(1, 40) = 41.08$, $MSe = 106,967$, $F_2(1, 79) = 30.91$, $MSe = 142,683$, thus producing the normal rvf-LH advantage for reading. Also as expected, participants named inference-related test words (673 ms, $SE = 8.74$) more quickly than they named unrelated test words (696, $SE = 10.18$), $F_1(1, 40) = 17.75$, $MSe = 49,754$, $F_2(1, 79) = 11.91$, $MSe = 65,853$. Finally, participants responded more quickly at the Predictive test point (2) (mean = 673, $SE = 8.57$) than at the Transition test point (3) (697 ms, $SE = 10.31$), $F_1(1, 40) = 14.76$, $MSe = 53,889$, $F_2(1, 79) = 10.46$, $MSe = 55,577$. These main effects were qualified by a reliable three-way interaction, $F_1(1, 40) = 5.00$, $MSe = 8,298$, $F_2(1, 79) = 3.97$, $MSe = 12,006$.

Discussion

These data show an intriguing pattern of inference-related activation in the cerebral hemispheres. In Experiments 1 and 3, we observed inference-related priming at Predictive

³We also conducted two-way ANOVAs on hemisphere and relatedness at each test point. At the Predictive test point (2) there were main effects of hemisphere, $F_1(1, 40) = 42.31$, $MSe = 60,066.75$, $F_2(1, 79) = 25.10$, $MSe = 103,572.03$, and of relatedness, $F_1(1, 40) = 13.68$, $MSe = 29,502$, $F_2(1, 79) = 9.26$, $MSe = 55,151$. There was also an interaction, $F_1(1, 40) = 5.62$, $MSe = 15,878$, $F_2(1, 79) = 2.20$, $MSe = 9,757$. At the Coherence test point (3) there was a main effect of hemisphere, $F_1(1, 40) = 19.63$, $MSe = 47,282$, $F_2(1, 79) = 7.37$, $MSe = 45,101$, and relatedness, $F_1(1, 40) = 9.90$, $MSe = 20,646$, $F_2(1, 79) = 1.97$, $MSe = 16,402$, but there was no interaction, $F_1(1, 40) = 1.88$, $MSe = 1817$, $F_2 < 1.0$.

test points (1) and (2) only in the lvf-RH. This confirms our prediction that the RH activates concepts that could support predictive inferences and is more likely to do so than the LH.

In the combined data from Experiments 1 and 2, immediately following the coherence break, at the Coherence test point (4) we observed priming for coherence inferences only in the rvf-LH, replicating a similar observation from a pilot study using different stories and test words, in which participants made lexical decisions. In Experiment 2, at the following clause break, by which point we might expect some resolution of the inference, we observed activation of the inference being sustained in both hemispheres at the Resolution test point (5). Note also that simple lexical priming effects — priming due to individual words in the stories, rather than due to inferences — would not be expected to vary the way our priming effects did across the hemifield and time point conditions in these experiments.

Exactly when or why the priming/activation shifts from RH advantage at test points (1) and (2) to a LH advantage at test point (4) remains somewhat unclear. This may partly be due to the materials, which do not contain a clause break at a point which clearly separates predictive from coherence inferences, because they were not designed to look for such a crossover point. However, it is also likely that individual differences, such as working memory capacity (St. George et al., 1997; Whitney et al., 1991), mediate when inferences are drawn, so that there will often be a point in a story at which some participants have drawn inferences but other subjects have not yet done so.

Sex Differences

In some studies of lateralized processing of language, men appear more lateralized than women. For instance, when making rhyme judgments, PET activation increases bilaterally in women, but only in the LH of men (Shaywitz et al., 1995; c.f., Frost et al., 1999). In our studies on laterality, we have rarely found any statistically reliable sex differences. When a sex difference emerges, it usually follows the pattern that both men and women show the observed laterality on the strongly lateralized conditions, but only men show it (or men show it more) in the weakly lateralized conditions. For instance, in one summation priming study, participants benefited more from Summation Inference primes when naming lvf-RH test words than when naming rvf-LH test words and displayed the opposite pattern for Direct primes. However, for another type of summation prime, men showed more priming when naming lvf-RH test words than when naming rvf-LH test words, but in women this asymmetry was not significant (Beeman et al., 1996). Similarly, in our problem-solving studies, after either 7 or 15 s of solving effort both men and women showed a RH advantage in priming when naming solution words, and in raw reaction times when judging whether test words were solutions. In contrast, after 2 s of solving effort, men showed a reliable RH advantage in priming and a marginal RH advantage in recognizing solutions, whereas women showed equal priming and equal latencies for recognition times across hemifields, (Beeman & Bowden, in press; Bowden & Beeman, 1998).

However, in Experiment 3, with equal numbers of men and women, we found stronger asymmetries for women than for men (no sex effects were observed in the other experiments, in which there were generally more women than men). The crossover in priming advantage from lvf-RH (at point 2, when the inference was still a predictive

inference) to rvf-LH (at point 3, as the coherence break unfolded) was reliable for women, $F_1(1, 40) = 4.27$, $MSe = 7,892.51$ but not for men, $F_1(1, 40) = 1.17$, $MSe = 1598.52$. The sex \times hemifield \times relatedness interaction was not reliable, $F_s < 1.0$, but the sex effect should be noted because it is opposite of the that often reported in the literature (i.e., stronger laterality for men).

GENERAL DISCUSSION

These results indicate that multiple processes and both hemispheres are involved in drawing inferences. Existing theories posit that (i) comprehend-ers activate many potential inferences but incorporate few into the representation of the text, usually after coherence breaks (Bramucci et al., 1995; Calvo & Castillo, 1996; Keefe & McDaniel, 1993); and (ii) when people comprehend text, the RH weakly activates large semantic fields, whereas the LH strongly activates small semantic fields (Beeman, 1993, 1998; Beeman et al., 1994; also, Burgess & Simpson, 1988; Chiarello et al., 1990; Kiefer, Weisbrod, Kern, Maier, & Spitzer, 1998; Nakagawa, 1991; Rodell, Cook, Regard, & Landis, 1992; Titone, 1998). We integrate these theories to propose a five-step model for drawing inferences based on parallel but interactive processing in the RH and LH. Step 3 is somewhat speculative, but there is strong empirical support for each of the others:

1. When people comprehend discourse, semantic processing proceeds somewhat in parallel: For each input word the RH activates large semantic fields, whereas the LH independently activates (or maintains activation of) smaller context-relevant semantic fields.
2. Because large semantic fields tend to overlap, activation capable of supporting predictive inferences is more likely to occur in the RH than in the LH. These predictive inferences are not routinely incorporated into the discourse representation, but can be under certain circumstances (high focus, high predictability, high-working-memory comprehenders, clear referential connectivity, etc.).
3. When a coherence break is encountered, a search ensues for potentially connecting information that has been previously activated by explicit mention, by close relation to input words, or by overlap from distantly related input words. We speculate that if the activation is found initially only in one hemisphere (the RH when relying on overlapping activation from distantly related words), it is output to the other hemisphere, with some momentary cost for further output of that concept in the source hemisphere.
4. When activation of a potentially connective concept is found, the concept is selected by enhancing its activation. This is more likely to occur in the LH, due to its tendency to strongly activate small semantic fields.
5. If the selected concept resolves the coherence break, it is incorporated into the discourse representation. Incorporation of inferences involves the same integrative processing undergone by concepts or propositions activated by explicit mention in

the discourse (Kintsch, 1988; Kintsch & Welsch, 1991), so that it is later difficult to distinguish inferred concepts from those explicitly encountered.

Evidence that the RH and LH produce complementary semantic activation during comprehension (Step 1) comes from many studies of semantic priming from single words, multiple words, or sentences (for review: Beeman, 1998; Chiarello, 1998; Faust, 1998).

Strong evidence that the RH is more likely than the LH to activate concepts that can support predictive inferences (Step 2) comes from Experiments 1 and 3, when participants named test words related to predictive inferences faster than they named unrelated test words, but only in the lvf-RH. It is possible that we did not observe inference-related activation in the rvf-LH at predictive test points (1) and (2) because the LH activated the inferences more slowly, or because the LH activation decayed extremely rapidly. However, sustained activation only in the RH is consistent with the fact that some RHD patients have difficulty activating inference-related information (Beeman, 1993) and answering questions about inferable information (Brownell et al., 1986; Beeman, 1993). The LH, relying only on its own resources, could probably move its “spotlight” of activation around to activate the information necessary to make the appropriate inferences. However, during highly demanding discourse processing, this mechanism is less efficient than capitalizing on diffuse activation in the RH, so comprehenders limited to the moving spotlight mechanism will miss some possible inferences, particularly inferences that are not necessary for maintaining coherence of the discourse as a whole.

Because so much information is diffusely active in the RH, the predictive inference-related activation may not be much stronger than activation of other concepts, including other potential predictions. The diffuse activation and further coarse semantic coding by the RH make it difficult for the RH alone to select this information for further processing, such as incorporation into the mental representation of the discourse. Indeed, there is ample evidence that predictive inferences are not completed as readily as coherence inferences (cf., Murray et al., 1993). For instance, comprehenders are slower to answer questions about predictive inferences than questions about coherence inferences or explicitly mentioned information (Singer & Ferreira, 1983).

The activation that could support predictive inferences may generally be passive. Various characteristics of the story (length, episodic structure, ease of referential connectiveness), the inference (degree of predictability, causal importance, focus, number of connections to the story), and the comprehender (working memory capacity, motivation) can affect which inferences are activated and/or incorporated (Graesser & Bertus, 1998; McKoon & Ratcliff, 1992; Murray et al., 1993; O'Brien et al., 1988; St. George et al., 1997; Singer et al., 1997; Whitney et al., 1991; Whitney, Ritchie, & Crane, 1992). Several of these characteristics are confounded with the distinction between predictive and coherence inferences. Coherence inferences are more likely than predictive inferences to be incorporated for several reasons: They are reactivated or reinforced by multiple parts of the story; they are probably more in focus and more important to the causal chain of the story; and comprehenders are more motivated to make these inferences that are, by definition, necessary to maintain coherence.

Priming for predictive inferences has been observed previously (e.g., Calvo & Castillo, 1996; Keefe & McDaniel, 1993; Murray et al., 1993). The priming obtained at predictive points (1) and (2) appear large, numerically, compared to other effects observed in the literature. This brings up the possibility that lateralized presentation of test words to the lvf-RH uncovers RH-specific activation that may be masked when target words are presented foveally in most standard paradigms. However, it should also be noted that we used a large number (80) of inferences (although the hemifield manipulation decreased the number of observations per cell) and a large number of participants (48). Also, although our effects were numerically large, latencies were slower and more variable than in some prior studies, due to the dual task, cross-modal paradigm and hemifield presentations. In particular, it is more difficult to name words briefly presented laterally, and the longer latencies and higher variability allow more room for priming. The fact that participants activated information that could support predictive inferences accords with the consensus view on deep text processing (although automatic processing during shallow reading may not include activation of predictive inferences, McKoon & Rateliff, 1992).

Searching for activation of an appropriate concept to resolve a coherence break (Step 3) is presumably necessary in selecting the necessary concept, although the search could be passive rather than active (and attention-demanding). Experiments 1 through 3 suggest that some cost is associated with outputting activation of the inference concept from one hemisphere to the other: Immediately following the coherence break, participants did not show inference-related priming for lvf-RH test words, but do for rvf-LH test words. Given that inference related priming was observed for lvf-RH test words at all other points, we speculate that this lack of priming at the Coherence point (4) reflects momentary inhibition of further output of the particular concept within the source hemisphere (in this case, the RH). This posited cost is specific to the output concept, as a general slowing should also have affected the unrelated test words. We argue that the cost is inhibition of output, rather than true suppression, because there was priming in accuracy for lvf-RH test words at test point (4) in Experiments 1 and 2.⁴ This is the most speculative of our conclusions, and it is, of course, possible that the lack of priming for lvf-RH test words at test point (4) is a type II error. However, we did not find lvf-RH priming at this test point either in Experiment 1 or in Experiment 2 or in a pilot study (with 64 participants) using lexical decision responses at an analogous test point in a different set of 32 inferences. It is also possible that the RH initially activates predictive inferences, but this activation for this concept decays in the RH just before the LH activates it, and then suddenly the RH reactivates the information. However, this does not seem to fit the entire set of data, including the fact that RHD patients do not show inference-related priming at a similar test point (Beeman, 1993), suggesting that the LH capitalizes on inference-related activation that existed in the RH prior to the coherence break.

Evidence that the inference concept is more likely to be selected in the LH than the RH (Step 4) is manifest when, immediately after coherence breaks, participants showed inference-related priming only for rvf-LH test words. This result was not reliable in

⁴Combined accuracy results from Experiments 1 and 2 at the Coherence point (4) for lvf-RH test words, $F_1(1, 40) = 14.42$, $MSe = 1.220$, $F_2(1, 79) = 18.23$, $MSe = 1240$, and for rvf-LH test words, $F_1(1, 40) = 6.85$, $MSe = 402$, $F_2(1, 79) = 4.54$, $MSe = 265$.

Experiment 1, but was in the analysis combining Experiments 1 and 2, as well as in a prior experiment (Beeman, 1990). We use the term selection in the sense proposed by Allport (1987), that is, that enhancement of activation for one particular concept (or motor program) allows that concept to be coupled with other processing, such as further processing or production. We propose that the LH plays a larger role than the RH in selecting a coherence inference and incorporating it into the representation of the text. The LH is well suited for this role because of its tendency to strongly activate a limited subset of semantic information, whereas the RH is poorly suited because of its tendency to maintain weak activation for many concepts. Selection could occur a number of ways: any information activated above an absolute threshold is selected; any information activated more strongly than competing information is selected when the difference in activation exceeds some relative threshold; or, simply, the strongest activation is selected. Once selected, the information is available for processing by other modules (such as speech output), or for inclusion in long-term memory, according to this model.

Evidence that comprehenders incorporate inferences (Step 5) comes from the aforementioned pilot study, as well as previous results (Beeman, 1990, 1993) in which participants answered Inference questions as accurately as Explicit questions. This is consistent with previous results showing that comprehenders answer questions about coherence inferences as rapidly as they answer questions about explicit information, and faster than questions about predictive inferences (Singer & Ferreira, 1983). Indeed, they have a difficult time distinguishing inferred information from information explicitly stated in the discourse (Johnson et al., 1973).

Parallel Processing and Sharing across the Hemispheres

These data contribute to a picture of interhemispheric cooperation in which the RH and LH differently process semantic information, in parallel, but share their outputs when necessary, and a person's overall behavior and thoughts results from the combined activation across hemispheres —population coding with the population of neurons divided across the hemispheres. That connected hemispheres of normal comprehenders process semantic information in parallel is suggested by asymmetric semantic priming observed at several points in the ongoing stories. Similarly, when people attempt to solve short insight-like word problems, the RH shows more activation for the solution of unsolved problems than the LH, and this seems to increase over time, up to at least 15 (Beeman & Bowden, in press; Bowden & Beeman, 1998).

It is also clear that, at some point, the hemispheres must share information. For instance, we have speculated that the LH capitalizes on inference-related activation of predictive inferences in the RH in order to select and incorporate coherence inferences. The issue of *how* information is shared between the hemispheres has come under increasing scrutiny recently (Banich & Karol, 1992; for review, Banich & Nicholas, 1998; Clarke, McCann, & Zaidel, 1998), and the current results merely highlight the need for delineating the processes of interhemispheric cooperation.

One characterization of hemispheric interaction consistent with our results, but admittedly speculative, is that when the RH outputs activation of potential inferences to the LH, it self-

inhibits the same information to prevent repetition of the output (similar to proposals that output-induced inhibition underlies repetition blindness). We suggest inhibition of output rather than suppression because the participants did respond more accurately to inference-related test words, just not more quickly.

It is also possible that when the LH receives activation from the RH, a feedback mechanism inhibits the source activation. A third potential mechanism for inhibition, that the hemispheres automatically inhibit each other homotopically (Cook, 1984; Cook & Beech, 1990; Rodell et al., 1992), is not compatible with the fact that both hemispheres sustained activation of inferences at the Resolution test point (5). Further experiments are necessary to demonstrate the true nature of hemispheric interaction during comprehension.

CONCLUSIONS

These results are consistent with the interpretation that both hemispheres are involved in drawing inferences. For an individual to fully comprehend discourse, interhemispheric cooperation must capitalize on the strengths of both hemispheres. We claim that these strengths are coarse semantic coding in the RH and the relatively fine semantic coding in the LH. RH coarse semantic coding makes the RH more likely than the LH to activate concepts that could support predictive inferences: When there is semantic overlap of distantly related information from multiple words of the story, weak activation from the related words passively summates (but is not selected) as a predictive inference in the RH. When a break in coherence occurs, LH fine semantic coding makes the LH more likely than the RH to select coherence inferences to be incorporated into the representation of the discourse. Additional research is needed to further specify the nature and mechanisms of interhemispheric cooperation in drawing inferences from discourse.

Acknowledgments

This work was supported in part by Grant R29 DC02160 from the National Institute on Deafness and other Communication Disorders, NIH, to M.J.B. and by Grant ROI NS 29926 from NIH and Army Research Institute awards DASW0194-K-0004, DASW0196-K-0013, and DAAG55-97-1-0224 to M.A.G. Thanks to Jim Tanaka, Marie Banich, Christine Chiarello, Art Graesser, and Edward O'Brien for useful comments on an earlier versions of the manuscript. Thanks to the following people for help with stimulus preparation and data collection: Caroline Bollinger, Julie Foertsch, Kim Hassenfeld, Donald Houghton, Dana Janes, Ying Lu, Rachel Roberston, Geeta Shivde, Beth Travis, and Linda Whitcombe.

References

- Abdullaev YG, Posner MI. Time course of activating brain areas in generating verbal associations. *Psychological Science*. 1997; 8:56–59.
- Allport, A. Selection for action: Some behavioral and neurophysiological considerations of attention and action. In: Heuer, H.; Sanders, AF., editors. *Perspectives on perception and action*. Hillsdale, NJ: Erlbaum; 1987. p. 395-415.
- Baggett P. Memory for explicit and implicit information in picture stories. *Journal of Verbal Learning and Verbal Behavior*. 1975; 14:538–548.
- Banich MT, Karol DL. The sum of the parts does not equal the whole: Evidence from bihemispheric processing. *Journal of Experimental Psychology: Human Perception and Performance*. 1992; 18:562–577. [PubMed: 1593235]

- Banich, MT.; Nicholas, CD. Integration of processing between the hemispheres in word recognition. In: Beeman, M.; Chiarello, C., editors. *Getting it right: The cognitive neuroscience of right hemisphere language comprehension*. Mahwah, NJ: Erlbaum; 1998. p. 51-78.
- Beeman, M. Unpublished doctoral dissertation. 1990. Coherence inferencing and structure building in the cerebral hemispheres.
- Beeman M. Semantic processing in the right hemisphere may contribute to drawing inferences during comprehension. *Brain and Language*. 1993; 44:80–120. [PubMed: 8467379]
- Beeman, M. Coarse semantic coding and discourse comprehension. In: Beeman, M.; Chiarello, C., editors. *Getting it right: The cognitive neuroscience of right hemisphere language comprehension*. Mahwah, NJ: Erlbaum; 1998. p. 225-284.
- Beeman MJ, Bowden EM. The right hemisphere maintains solution-related activation in yet-to-be-solved insight problems. *Memory & Cognition*. In press.
- Beeman, ML.; Bowden, E.; Hassenfeld, K.; Shivde, G. Right hemisphere advantage for some summation primes, not others; Poster presented at the 37th annual meeting of the Psychonomic Society; Chicago, IL; Nov. 1996
- Beeman M, Friedman RB, Grafman J, Perez E, Diamond S, Lindsay MB. Summation priming and coarse coding in the right hemisphere. *Journal of Cognitive Neuroscience*. 1994; 6:26–45. [PubMed: 23962328]
- Bottini G, Corcoran R, Sterzi R, Paulesu E, Schenone P, Scarpa P, Frackowiak RSJ, Frith CD. The role of the right hemisphere in the interpretation of figurative aspects of language: A positron emission tomography activation study. *Brain*. 1994; 117:1241–1253. [PubMed: 7820563]
- Bowden E, Beeman MJ. Getting the right idea: Semantic activation in the right hemisphere may help solve insight problems. *Psychological Science*. 1998; 6:435–440.
- Bramucci, RS.; Budd, D.; Whitney, P. Task effects on the activation and incorporation of predictive inferences; Poster presented at the 36th Annual Meeting of the Psychonomic Society; Los Angeles, CA. Nov. 1995
- Brownell HH, Potter HH, Bihrlle AM, Gardner H. Inference deficits in right brain-damaged patients. *Brain and Language*. 1986; 29:310–321. [PubMed: 3955344]
- Burgess C, Simpson G. Hemispheric processing of ambiguous words. *Brain and Language*. 1988; 33:86–104. [PubMed: 3342321]
- Calvo MG, Castillo MD. Predictive inferences occur online, but with delay: Convergence of naming and reading times. *Discourse Processes*. 1996; 22:57–78.
- Chiarello, C. On codes of meaning and the meaning of codes: Semantic access and retrieval within and between the hemispheres. In: Beeman, M.; Chiarello, C., editors. *Getting it right: The cognitive neuroscience of right hemisphere language comprehension*. Mahwah, NJ: Erlbaum; 1998. p. 141-160.
- Chiarello C, Burgess C, Richards L, Pollock A. Semantic and associative priming in the cerebral hemispheres: Some words do, some don't ... sometimes, some places. *Brain and Language*. 1990; 38:75–104. [PubMed: 2302547]
- Clarke, JM.; McCann, CM.; Zaidel, E. The corpus callosum and language: Anatomical-Behavioral relationships. In: Beeman, M.; Chiarello, C., editors. *Getting it right: The cognitive neuroscience of right hemisphere language comprehension*. Mahwah, NJ: Erlbaum; 1998. p. 27-50.
- Cook ND. Homotopic callosal inhibition. *Brain and Language*. 1984; 23:116–125. [PubMed: 6478187]
- Cook ND, Beech AR. The cerebral hemispheres and bilateral neural nets. *International Journal of Neuroscience*. 1990; 52:201–220. [PubMed: 2269607]
- Duffy SA. Role of expectations in sentence integration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1986; 12:208–219.
- Faust, M. Obtaining evidence of language comprehension from sentence priming. In: Beeman, M.; Chiarello, C., editors. *Getting it right: The cognitive neuroscience of right hemisphere language comprehension*. Mahwah, NJ: Erlbaum; 1998. p. 161-186.
- Faust ME, Gernsbacher MA. Cerebral mechanisms for suppression of inappropriate information during sentence comprehension. *Brain and Language*. 1996; 53:234–259. [PubMed: 8726535]
- Fincher-Kiefer R. Relative inhibition following the encoding of bridging and predictive inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1995; 21:981–995.

- Frost JA, Binder JR, Springer JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. Language processing is strongly left lateralized in both sexes: Evidence from functional MRI. *Brain*. 1999; 122:199–208. [PubMed: 10071049]
- Garrod S, O'Brien EJ, Morris RK, Rayner K. Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1990; 16:250–257.
- Gernsbacher, MA. *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum; 1990.
- Glenn CG. The role of episodic structure and of story length in children's recall of simple stories. *Journal of Verbal Learning and Verbal Behavior*. 1978; 17:229–247.
- Graesser AC, Bertus EL. The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading*. 1998; 2:247–269.
- Graesser AC, Singer M, Trabasso T. Constructing inferences during narrative text comprehension. *Psychological Review*. 1994; 101:371–395. [PubMed: 7938337]
- Grafman J, Salazar AM, Vance SC, Weingartner H, Amin D. Immediate memory for story discourse in Vietnam veterans with penetrating brain wounds. *Journal of Clinical and Experimental Neuropsychology*. 1987; 9:23–34.
- Hough MS. Narrative comprehension in adults with right and left hemisphere brain-damage: Theme organization. *Brain and Language*. 1990; 38:253–277. [PubMed: 1691038]
- Jenkins, JJ. The 1952 Minnesota word association norms. In: Postman, L.; Keppel, G., editors. *Norms of word association*. New York: Academic Press; 1970.
- Johnson MK, Bransford JD, Solomon SK. Memory of tacit implications of sentences. *Journal of Experimental Psychology*. 1973; 98:203–205.
- Kaplan, R. On process models for sentence analysis. In: Norman, D.; Rumelhart, D., editors. *Explorations in cognition*. San Francisco: Freeman; 1975. p. 117-135.
- Keefe DE, McDaniel MA. The time course and durability of predictive inferences. *Journal of Memory and Language*. 1993; 32:446–463.
- Keenan, JM.; Golding, JM.; Potts, GR.; Jennings, TM.; Aman, CJ. Methodological issues in evaluating the occurrence of inferences. In: Graesser, AC.; Bower, GH., editors. *The psychology of learning and motivation*. New York: Academic Press; 1989.
- Keppel, G.; Strand, B. Free association responses to the primary purposes and other responses selected from the Palermo-Jenkins norms. In: Postman, L.; Keppel, G., editors. *Norms of word association*. New York: Academic Press; 1970.
- Kiefer M, Weisbrod M, Kern IL, Maier S, Spitzer M. Right hemisphere activation during indirect semantic priming: Evidence from event-related potentials. *Brain and Language*. 1998; 64:377–408. [PubMed: 9743549]
- Kintsch W. The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*. 1988; 95:163–182. [PubMed: 3375398]
- Kintsch, W. *Comprehension: A paradigm for cognition*. New York: Cambridge University Press; 1998.
- Kintsch, W.; Welsch, DM. The construction-integration model: A framework for studying memory for text. In: Hockley, WW.; Lewendowsky, S., editors. *Relating theory and data: Essays in honor of Bennet B Murdoch*. Hillsdale, NJ: Erlbaum; 1991.
- Kurtzman HS. Sex bias in language stimuli. *American Psychologist*. 1990; 45:1273–1275.
- Marshall, GR.; Cofer, C. Single-word free association norms for 328 responses from the Connecticut cultural norms for verbal items in categories. In: Postman, L.; Keppel, G., editors. *Norms of word association*. New York: Academic Press;
- McDonald S, Wales R. An investigation of the ability to process inferences in language following right hemisphere brain damage. *Brain and Language*. 1986; 29:68–80. [PubMed: 3756462]
- McKoon G, Ratcliff R. Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1986; 12:82–91.
- McKoon G, Ratcliff R. Semantic associations and elaborative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1989; 15:326–338.
- McKoon G, Ratcliff R. Inference during reading. *Psychological Review*. 1992; 99:440–466. [PubMed: 1502273]

- Murray JD, Klin CM, Myers JL. Forward inferences in narrative text. *Journal of Memory and Language*. 1993; 32:464–473.
- Nakagawa A. Role of anterior and posterior attention networks in hemispheric asymmetries during lexical decisions. *Journal of Cognitive Neuroscience*. 1991; 3:315–321.
- Noordman LG, Vonk W, Kempff HJ. Causal inferences during the reading of expository texts. *Journal of Memory and Language*. 1992; 31:573–590.
- O'Brien EJ, Shank DM, Myers JL, Rayner K. Elaborative inferences during reading: Do they occur online? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1988; 14:410–420.
- Paris S, Lindauer BK. The role of inferences in children's comprehension and memory for sentences. *Cognitive Psychology*. 1976; 8:211–221.
- Potts GR, Keenan JM, Golding JM. Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language*. 1988; 27:399–415.
- Rodel M, Cook ND, Regard M, Landis T. Hemispheric dissociation in judging semantic relations: Complementarity for close and distant associates. *Brain and Language*. 1992; 43:448–459. [PubMed: 1446212]
- St George M, Mannes S, Hoffman JE. Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience*. 1997; 9:776–787. [PubMed: 23964599]
- St George M, Kutas M, Martinez A, Sereno MI. Semantic integration in reading: A role for the right hemisphere in language comprehension. 1998 Manuscript submitted for publication.
- Schneiderman EI, Murasugi KG, Saddy JD. Story arrangement abilities in right-brain damaged patients. *Brain and Language*. 1992; 43:107–120. [PubMed: 1643506]
- Seger CA, Desmond JE, Glover GH, Gabrieli JDE. fMRI evidence for right hemisphere involvement in processing unusual semantic relationships. *Neuropsychology*. In press.
- Shaywitz BA, Shaywitz SE, Pugh KR, Constable RT, Skudlarski P, Fulbright RK, Bronen RA, Fletcher JM, Shankweiler DP, Katz L, Gore JC. Sex differences in the functional organization of the brain for language. *Nature*. 1995; 373:607–609. [PubMed: 7854416]
- Singer M. Processes of inference in sentence encoding. *Memory and Cognition*. 1979; 7:192–200.
- Singer M. The role of case-filling inferences in the coherence of brief passages. *Discourse Processes*. 1980; 3:185–200.
- Singer M, Ferreira F. Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior*. 1983; 22:437–448.
- Singer M, Graesser AC, Trabasso T. Minimal or global inference during reading. *Journal of Memory and Language*. 1994; 33:421–441.
- Singer M, Harkness D, Stewart ST. Constructing inferences in expository text comprehension. *Discourse Processes*. 1997; 24:199–228.
- Titone D. Hemispheric differences in context sensitivity during lexical ambiguity resolution. *Brain and Language*. 1998; 65:361–394. [PubMed: 9843609]
- Van Lancker DR, Kempler D. Comprehension of familiar phrases by left- but not right-hemisphere damaged patients. *Brain and Language*. 1987; 32:265–277. [PubMed: 2446699]
- Weylman ST, Brownell HH, Roman M, Gardner H. Appreciation of indirect requests by left- and right-brain-damaged patients: The effects of verbal context and conventionality of wording. *Brain and Language*. 1989; 36:580–591. [PubMed: 2470462]
- Whitney P, Ritchie BG, Crane BS. The effects of foregrounding on readers' use of predictive inferences. *Memory and Cognition*. 1992; 20:424–432. [PubMed: 1495404]
- Whitney P, Ritchie BG, Clark MB. Working-memory capacity and the use of elaborative inferences in text comprehension. *Discourse Processes*. 1991; 14:133–145.
- Winner E, Gardner H. The comprehension of metaphor in brain-damaged patients. *Brain*. 1977; 100:719–727.

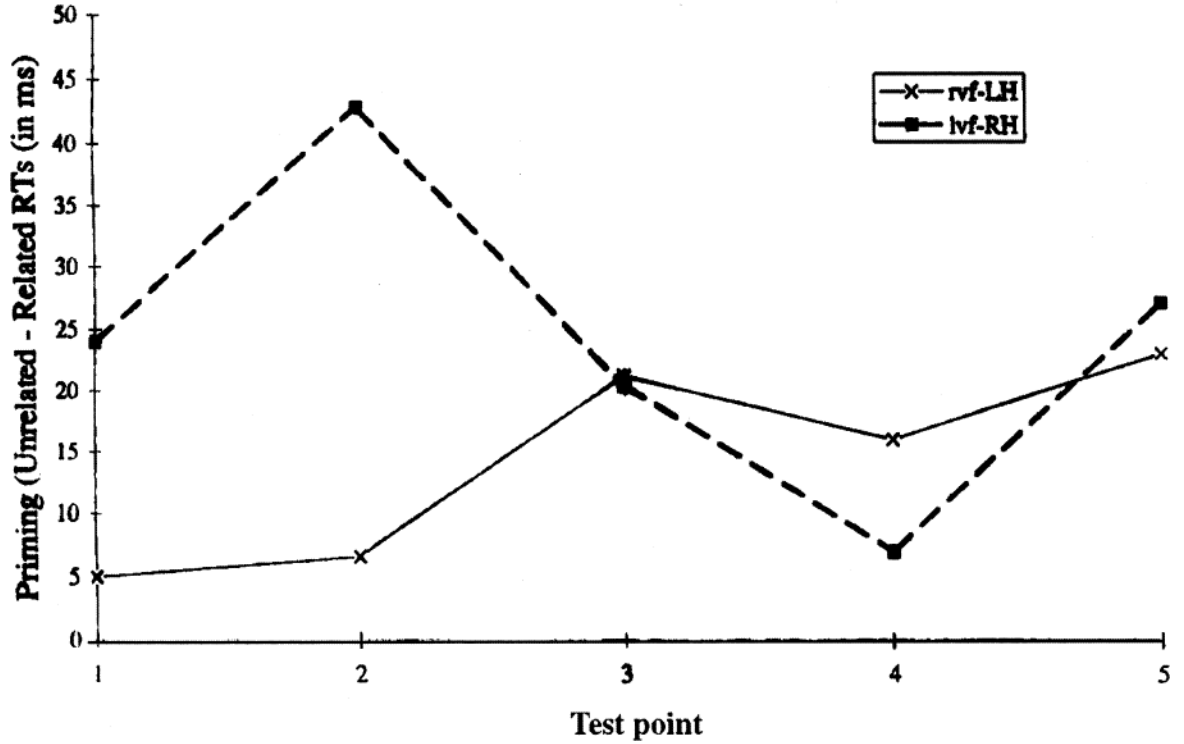


FIG. 1. Participants' mean priming (Unrelated-Inference-related latencies) for test words presented to the lvf-RH and rvf-LH at five test points: Test point (1) from Experiment 1; (2) and (3) from Experiment 3; (4) from combined data of Experiments 1 and 2; and (5) from Experiment 2. Test points (1) and (2) assessed inference-related activation at points when inferences would be considered predictive, test point (3) assessed activation during the transition from predictive to coherence inferences, and Coherence point (4) and Resolution point (5) assessed inference-related activation at points when inferences were necessary for coherence: "The shuttle sat on the ground in the distance (1), waiting for the signal to be given (2). After a huge roar (3) and a bright flash, the shuttle disappeared into space (4), leaving clouds of smoke in its wake (5), and the audience cheered."

TABLE 1

Example Paragraph Used in Experiments 1 through 3, with the Five Test Points Marked: Predictive (1) and (2), Transition (3), Coherence (4), and Resolution (5)

Bob took his daughter Karen out of school for the day so she could enjoy a very historic event that would take place that morning. The shuttle sat on the ground in the distance (1), waiting for the signal to be given (2). After a huge roar (3) and a bright flash, the shuttle disappeared into space (4), leaving clouds of smoke in its wake (5), and the audience cheered.

Note. Inference-related test word, *launch*, unrelated test word, *sand*.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Mean and Standard Error Pronunciation Latencies and Accuracies for Inference-Related and Unrelated Test Words by Hemisphere and Test Point

TABLE 2

Test points:	Predictive (1)		Predictive (2)		Transition (3)		Coherence (4) ^a		Resolution (5)		
	rvf-LH	lvf-RH	rvf-LH	lvf-RH	rvf-LH	lvf-RH	rvf-LH	lvf-RH	rvf-LH	lvf-RH	
Latencies (ms)											
Related	M	639	654	652	669	670	702	636	666	647	662
	SE	15.0	13.0	15.6	14.4	19.3	19.7	8.8	10.4	13.0	13.0
Unrelated	M	644	678	658	712	691	723	652	673	669	689
	SE	15.0	14.0	17.5	19.7	20.4	22.9	8.8	10.1	14.0	14.0
Priming		5	24**	6	43**	21**	20*	16**	7	22**	27**
Accuracy (%)											
Related	M	94.2	91.1	94.4	89.2	94.6	87.8	95.9	93.8	95.2	94.0
	SE	1.3	1.6	1.2	1.5	1.4	1.5	0.7	0.9	1.0	1.1
Unrelated	M	92.4	89.5	92.1	83.3	92.3	81.4	93.0	88.8	89.2	90.0
	SE	1.2	1.6	1.4	2.2	1.2	2.1	1.0	1.0	1.6	1.6
Priming		1.8	1.6	2.3	5.8**	2.3	6.5**	2.9**	5.0**	6.0**	4.0*

^aData from Experiments 2 and 3, 96 participants total (all other Ns = 48).

* $p < .05$.

** $p < .01$.