Violating stereotypes:
Eye movements and comprehension processes
when text conflicts with world knowledge

SUSAN A. DUFFY and JESSICA A. KEIR
University of Massachusetts, Amherst, Massachusetts

We investigated the effect of discourse context on the access of word meaning during reading. Target words were role names (e.g., *electrician*) for which there was a gender stereotype (e.g., electricians are stereotypically male). Target sentences contained a reflexive pronoun that referred to the role name (e.g., *The electrician taught herself* . . .). Participants read these target sentences with or without paragraph context while their eye movements were monitored. In the absence of discourse context and in neutral discourse contexts, fixation times on the reflexive pronoun and immediately following the pronoun were inflated when the pronoun specified a gender that mismatched the stereotype, indicating that the gender stereotype was activated upon encountering the role name. When prior discourse context indicated the gender of the role-named character, this mismatch effect was eliminated. The mismatch effect indicates that gender stereotypes are automatically activated in the absence of disambiguating information. The lack of an effect when gender has previously been specified is consistent with the lexical reinterpretation model proposed by Hess, Foss, and Carroll (1995).

One question central to the comprehension of words in discourse context is whether and how the processing of individual words is modified by the discourse context. For example, consider a sentence that begins with the phrase *The electrician taught herself* . . . Because electricians are stereotypically male, the reader may encounter comprehension difficulty when processing the reflexive pronoun *herself*. Such difficulty may reflect activation of the gender stereotype when processing *electrician* initiates lexical access, or it may occur when the reader attempts to integrate *herself* with *electrician*. In particular, reading times on or after the reflexive pronoun should be inflated. But now consider the processing required to comprehend *The electrician* in the last sentence of the following paragraph:

> Jeff's power had been unreliable ever since the tornado. The electrician was a cautious woman who carefully secured her ladder to the side of the house before checking the roof. Jeff suspected that high winds had loosened the connection to

the power lines. *The electrician* taught herself a lot while fixing the problem.

When the reader encounters *electrician* in the last line of the paragraph, it is actually the second encounter of this concept within the discourse. Thus, there are now two sources of information available to comprehension processes trying to make sense of this word. First, semantic memory provides information about the meaning of the word *electrician*, including the male stereotype. Second, the discourse context provides information indicating that this particular electrician happens to be female. If the comprehension process proceeds as it does in the absence of extensive discourse context, then the gender stereotype may affect comprehension. If, however, the process is modified by information present in the discourse context, then the gender stereotype may not influence comprehension as reflected in reading times.

Both alternatives have support in current models in the literature. Under the assumptions of modular models of access, lexical access is always initially independent of discourse context. Thus, for example, in the classic models of processing for ambiguous words, both meanings are initially accessed despite disambiguating context (Gernsbacher, Varner, & Faust, 1990; Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; Swinney, 1979; Till, Mross, & Kintsch, 1988). Under modified models like the reordered access model (Duffy, Morris, & Rayner, 1988; Rayner, Pacht, & Duffy, 1994), although context hastens activation of the intended meaning, the unintended meaning is still activated. Thus, context does not initially prevent activation of the alternative meanings of an ambiguous word. This concept of initial context-independent

lexical access is incorporated into the construction phase of Kintsch's (1988) construction–integration model. Under this type of model, if the gender stereotype for a word like *electrician* is normally activated in the absence of informative discourse context, then the gender stereotype should also be activated in a discourse context like that given above.

Support for this view is also provided by the discourse prominence theory (Gordon & Hendrick, 1998), a theory of discourse processing that focuses in part on specifying antecedent retrieval processes for pronouns and definite noun phrases. Under this model, the critical difference between a pronoun with an antecedent and a definite noun phrase with an antecedent is that the pronoun immediately initiates a search for an antecedent, but a definite noun phrase does not. The definite noun phrase initiates a process that represents the noun as a new entity in the discourse model and presumably involves accessing its meaning in semantic memory and creating an entity with that meaning in the discourse model. Only after this is completed is there a search for a matching entity already in the discourse. Thus, in these models, the gender stereotype for *electrician* would be accessed before the electrician entity in the discourse model is retrieved. Consequently, the time to read *herself* should show the influence of the gender stereotype.

In contrast, the lexical reinterpretation model of Hess, Foss, and Carroll (1995; Foss & Speer, 1991) holds that the discourse representation is continually updated to maintain coherence. Then, as each word is processed, it is interpreted in terms of the global discourse model, a process that facilitates integration and overrides local semantic relations. Hess et al. tested the predictions of this type of model with a cross-modal priming procedure in which participants heard a discourse context and then named a target word that appeared on the video screen as the last word of the context. They used target sentences in which the subject of the sentence was a role name and the object was a word that was semantically related to the role name (e.g., The *English major* wrote the *poem*). When they presented their sentences in isolation, with no prior discourse context, they found that naming times for the target words (e.g., *poem*) were facilitated compared with a neutral control condition. However, when they embedded the sentences in discourse contexts that rendered the semantic features of the priming noun irrelevant (e.g., a paragraph about an English major who was having trouble in a computer science course and working on a big assignment in the course), the priming effect was eliminated. They explained this elimination by saying that within the discourse context, the concept of *English major* had been reinterpreted to reflect the context (e.g., as a student struggling in a computer science course). Thus, within the discourse model, it is a student in a technical course who *wrote the poem*. That revised model offers no basis for facilitation of *poem*.

In the current studies, we extended Hess et al.'s findings to discourse contexts in which the semantic features of interest were actually in conflict with the discourse context

rather than simply irrelevant. We used role names with strong gender stereotypes, and in our critical discourse contexts the stereotypes were violated (as in the *electrician* paragraph given earlier). Rather than making use of a natural priming relationship between semantically related words, we employed an interfering relationship between role names and a reflexive pronoun that appeared later in the target sentences (e.g., *The electrician taught herself...*). On the basis of Hess et al.'s strategy, in Experiment 1 we established that the gender stereotypes for role names interfered with the comprehension of a later reflexive pronoun that violated the stereotype. In Experiment 2, we investigated whether this interference effect was reduced in discourse context.

A second difference from the Hess et al. procedure was that instead of using a probe word, we recorded eye movements as participants read the passages. While the probe technique captures the accessibility of the probed concept at a particular point in time, eye movements provide very good information about temporal processing (Rayner, 1998). Indeed, several recent eye movement experiments in which the reader's general knowledge conflicted with information in the text have demonstrated that the effect of the contradiction was either apparent immediately or became apparent downstream from the point at which it might first become apparent (Cook & Myers, 2004; Garrod & Terras, 2000). Our procedure enabled us to study the time course of processing, and to determine the location of effects of gender stereotype on reading times if such effects occurred at all in the presence of an earlier context that established the gender.

In summary, in Experiment 1, we used eyetracking methodology to determine whether violations of gender stereotypes caused disruption of the reading process, and if so, when that disruption took place. Experiment 2 was designed to determine whether a prior context that specified the gender would modify any effects of a second encounter with the role name.

EXPERIMENT 1

In this experiment, the goal was to establish that gender stereotypes are accessed during comprehension and do interfere with processing when the stereotype is violated. A number of other studies have found evidence for the activation of such stereotypes. In particular, Banaji and Hardin (1996) found that when a pronoun was presented immediately after a gender-stereotyped prime word (e.g., *nurse*, *mechanic*), judgments on the pronoun were faster when the pronoun matched the gender stereotype for the prime; this held for judgments that were gender-related (is the pronoun male or female?) and nongender-related (is the target word a pronoun or not?). These findings provide evidence for the activation of the gender stereotype associated with the role names. Carreiras, Garnham, Oakhill, and Cain (1996) found in a self-paced reading task that readers were slower to read a sentence that referred by pronoun to a preceding role name when the gender of the

pronoun mismatched the stereotype for the role name. Osterhout, Bersick, and McLaughlin (1997) monitored event-related brain potentials (ERPs) while participants read sentences containing reflexive pronouns whose antecedent nouns were either by definition male or female (e.g., *uncle*, *cowgirl*) or stereotypically male or female (e.g., *bartender*, *babysitter*). Violations of gender definition and of gender stereotype both produced larger positive-going waves (P600) than did nonviolations. Finally, Ehrlich (1983) did not find a large effect of stereotype violation in a study of eye movements as participants read sentences containing pronouns (*he* or *she*) that referred to a character introduced by role (e.g., *producer*). The lack of reliable effect in this study may simply reflect Ehrlich's finding that it was difficult to identify the precise fixation(s) on which participants processed the pronoun, since participants tend not to fixate frequent words that are only 2–3 letters long (see also Ehrlich & Rayner, 1983). On the other hand, in a somewhat different context, Sturt (2003) found clear evidence that when the gender of a reflexive pronoun mismatched the stereotype for the role name, fixation times on the pronoun increased.

In the present study, we monitored eye movements as participants read sentences that were structured similarly to those used by Osterhout et al. (1997). Our target pronouns were reflexive pronouns (*himself*, *herself*); thus, they were long enough to require direct fixation (as per Sturt, 2003). Example sentences are given in Table 1, each of which introduced a character by role name in a definite noun phrase (e.g., *the babysitter*). The role name was either stereotypically male or female and was the subject of the sentence. The noun phrase was followed by a verb and a reflexive pronoun that specified the gender of the subject noun. The gender of the reflexive pronoun either matched or mismatched the gender stereotype. We expected that the effect of stereotype mismatch would emerge in fixation times on the reflexive pronoun itself and on the region following it.

Method

Participants

The participants were 40 students from the University of Massachusetts who were paid or received course credit for participation.

Materials

Gender stereotypes were established in a norming study in which the participants were asked to rate the gender stereotype for 100 roles or occupations. The participants used a rating scale in which 1 = *extremely male* and 7 = *extremely female*. For this experiment, 24 role names were selected, 12 with mean ratings of less than 2 and 12 with mean ratings of greater than 6.

The stimuli consisted of 24 sentences, one for each of the 24 role names. Each sentence introduced a character by role name in subject position and then referred to this character with a reflexive pronoun that followed the main verb. We decided on reflexive pronouns because they unambiguously refer to the subject of the sentence and because they are long enough to be directly fixated. For a given participant, half of the reflexive pronouns mismatched the gender stereotype for the role character, and the other half matched it. Within each of these conditions, for each participant, half of the sentences contained role names that were stereotypically female, and half contained role names that were stereotypically male. Each participant saw only one version of each sentence, but each sentence appeared in both conditions across the full set of participants. An additional 80 filler sentences were included, but they did not have reflexive pronouns or any stereotype mismatches. The presentation order for the full set of sentences was randomized for each participant.

Apparatus

Target sentences were presented on an NEC MultiSync computer monitor in standard upper- and lower-case letters. The participants were seated 62 cm from the monitor; three characters equaled one degree of visual angle. Sentences were displayed on two lines up to 80 characters long and were arranged so that a given critical region appeared on one line. The participants' eye movements were monitored by a Fourward Technologies Dual Purkinje Eyetracker (Generation V). Viewing was binocular, but eye movements were recorded from the right eye. The eyetracker signal was sampled every millisecond by the computer.

Procedure

To eliminate head movements, a bite bar was prepared for each participant and attached to the eyetracker platform. Once the participant was in place on the bite bar, the eyetracker was calibrated. This initial calibration process took about 5 min. The participants were informed that they would be reading a series of unrelated sentences and that they should be sure to read for comprehension. They were also told that occasionally a sentence would be followed by a comprehension question. Each trial began with the display of a series of boxes positioned across the screen. The participants were instructed to fixate the left-hand box, which marked the beginning of the sentence. Once the participants fixated this box, the experimenter displayed the sentence. The participants read the sentence at their own pace and then pressed a response button which erased the sentence from the screen. After 25% of the sentences, the participants were given a yes–no comprehension question. They were quite accurate in answering these questions (92% correct).

Results

We examined fixations on the reflexive pronoun itself and on the post-pronoun region, which consisted of the next one to four words (mean number of words = 2.6) immediately after the pronoun. The latter region was included because other studies have indicated that processing on a target word often continues after the eyes have moved into the next region (e.g., Duffy & Rayner, 1990; Rayner & Duffy, 1986). The primary measures were first

---

Table 1
Example Sentences Used in Experiment 1

Female Stereotype Role Names
    The babysitter found herself/himself humming while walking up to the door.
    The secretary treated herself/himself to a large sundae after finishing work.
Male Stereotype Role Names
    The senator perjured himself/herself on the stand in an attempt to avoid prison.
    The firefighter burned himself/herself while rescuing victims from the building.

Table 2
Mean First Pass and Go-Past Times (in Milliseconds) for the Pronoun and Post-Pronoun Regions in Experiment 1

Measure	Mismatch	Match	Difference
	Pronoun Region		
First pass	284	272	12
Go-past	340	316	24
	Post-Pronoun Region		
First pass	512	481	31
Go-past	624	546	78

pass time and go-past time for each region. First pass time is the sum of all fixation durations from the first fixation to enter the region to the last fixation before leaving the region in any direction. For a single word, this measure is also called gaze duration (Rayner, 1998). Go-past time is the sum of all fixations beginning with the first fixation in the region and ending with the last fixation before going forward to fixate a later position in the sentence (Duffy et al., 1988; Liversedge, Paterson, & Pickering, 1998; Rayner, 1998; Rayner & Duffy, 1986). This measure includes the first pass fixations in the region plus any regressions made from the region, and any refixations of the region before going past it. Go-past time is a measure of how much processing the reader needed to comprehend the region well enough to move beyond it (e.g., Duffy et al., 1988; Duffy & Rayner, 1990). Statistical tests based on both participants (F_1) and items (F_2) variability are reported; 5% of the data were lost due to track losses.

Mean first pass times and go-past times for the reflexive pronoun and for the post-pronoun region are given in Table 2. In both measures for both regions, readers' times were longer in the stereotype mismatch condition, compared with the match condition. For the reflexive pronouns, this effect was marginal in the analysis of first pass times [$F_1(1,39) = 3.36, p = .07; F_2(1,23) = 1.72, p = .20$] and in the participants' analysis of go-past times [$F_1(1,39) = 2.61, p = .11; F_2(1,23) = 5.13, p < .05$]. For the post-pronoun region, this mismatch effect was marginal in the participants' analysis of first pass times [$F_1(1,39) = 2.80, p = .10; F_2(1,23) = 4.40, p < .05$], but fully reliable in the analysis of go-past times [$F_1(1,39) = 10.51, p < .01; F_2(1,23) = 18.09, p < .001$]. Further analyses of the data indicated that this reliable go-past effect in the post-pronoun region in part reflected a tendency for readers to regress from this region more frequently in the mismatch condition (14% regressions) than in the match condition (7% regressions) [$F_1(1,39) = 10.33, p < .01; F_2(1,23) = 5.71, p < .05$].

Given the reliable regression effect, we also analyzed second pass times for earlier regions in the sentence. Second pass time is the time spent refixating a region after the eyes have left the region. Table 3 presents second pass times for the role name, with the verb following the role name and the reflexive pronoun. For all regions, readers spent longer refixating the region in the stereotype mismatch condition compared with the match condition. This

effect was significant for the role name [$F_1(1,39) = 8.84, p < .01; F_2(1,23) = 8.73, p < .01$] and for the reflexive pronoun [$F_1(1,39) = 18.22, p < .001; F_2(1,23) = 10.03, p < .001$], and marginal in the participants analysis, but significant in the items analysis for the verb [$F_1(1,39) = 3.58, p = .066; F_2(1,23) = 8.96, p < .01$].

Discussion

Gender stereotypes for the role names did interfere with processing of reflexive pronouns that specified a gender mismatching the stereotype. This interference emerged immediately on the reflexive pronoun itself, and was significant in the region that followed the reflexive pronoun. Readers did not go past this region without registering the mismatch, and their response to the mismatch was a tendency to reread earlier parts of the sentence. The question we ask in Experiment 2 is whether the gender stereotypes for the role names continue to affect reading times when the target sentences are placed in a larger discourse context in which the role-name character has already been introduced and the gender of this character has already been specified.

EXPERIMENT 2

In this study, we created a series of paragraphs focused on a particular character who was identified with a role or occupation that has a gender stereotype (e.g., *electrician*). Examples are given in Table 4. In the prior disambiguating versions of the paragraphs, the gender of the character was explicitly stated in the second sentence, where the character was introduced. In the neutral versions of the paragraphs, the gender of the character was not stated. The fourth sentence in each paragraph was the target sentence. The target sentences had the same structure as those in Experiment 1. Eye movements were monitored while participants read the paragraphs. For the neutral contexts, we expected again to observe an effect of stereotype mismatch on the reflexive pronoun and post-pronoun region of the target sentences. The critical question is whether this effect is also observed for the prior disambiguating contexts. Under the lexical reinterpretation model (Hess et al., 1995), the *electrician* in the discourse model should have been instantiated as a woman following the first encounter. Therefore, the reflexive pronoun *herself* should be readily integrated into the discourse model, causing the mismatch effect to disappear. Under models in which semantic information is activated independent of discourse context, then, an effect of the contradiction in genders

Table 3
Mean Second Pass Times (in Milliseconds) for Role Name, Verb, and Reflexive Pronoun Regions in Experiment 1

Region	Mismatch	Match	Difference
Role name	60	29	31
Verb	91	64	27
Pronoun	73	45	28

Table 4
Sample Stimulus Paragraphs for Experiment 2

Neutral Context

Jeff's/Lucy's power had been unreliable ever since the tornado.
The electrician was cautious and carefully
secured the ladder to the side of the house before
checking the roof. Jeff/Lucy suspected that high winds
had loosened the connection to the power lines.
The [electrician] [taught] [herself/himself] [a lot while fixing] the problem.

Prior Disambiguating Context

Jeff's/Lucy's power had been unreliable ever since the tornado.
The electrician was a cautious [woman/man] [who carefully]
secured her/his ladder to the side of the house before
checking the roof. Jeff/Lucy suspected that high winds
had loosened the connection to the power lines.
The [electrician] [taught] [herself/himself] [a lot while fixing] the problem.

Note—Critical regions are in brackets.

should be apparent either at the reflexive pronoun or in the subsequent post-pronoun region.

Carreiras et al. (1996) asked a similar question involving gender stereotype mismatches, employing English and Spanish stimuli. They first established that in English when a pronoun specifies a gender that mismatches the stereotypical gender for its referent, reading times for the pronoun sentence are increased. They then carried out a set of similar studies in Spanish, in which the gender of the role names is disambiguated by the preceding article (e.g., *la carpintera* indicates a female carpenter, whereas *el carpintero* indicates a male carpenter). Thus, in these studies, the first sentence introduced a character by role name, but this reference included a specification of gender. The second sentence then referred to this character by pronoun. This time Carreiras et al. found lengthened reading times on the first sentence when the article indicating the gender stereotype was violated, but no effect of stereotype mismatch on the second sentence containing the pronoun. Thus, having established a female carpenter, the reader had no trouble processing a pronoun that refers to the carpenter as *she*. Carreiras et al. concluded that the pronoun accessed an established entity in the mental model, a view similar to the lexical reinterpretation explanation offered by Hess et al. for their results.

In the present experiment, we established the protagonist's gender in English-language passages by explicitly stating the gender associated with the role (e.g., *The electrician was a woman who. . .*). Later in the passage, in contrast to Carreiras et al.'s stimuli, the target sentences contained a second explicit mention of the role name, rather than a pronoun that referred to the character introduced earlier. This allowed us to focus on how the reading process goes forward after the second encounter of the role name. Experiment 1 demonstrated that a mismatch of stereotype and stated gender caused readers difficulty. In Experiment 2, we ask whether such difficulties also will appear on the second encounter. Or does comprehension proceed in a different way, reflecting the presence of the entity in the discourse model?

Within the target sentences (e.g., *The electrician taught herself*), there are actually two general areas in which an

effect of stereotype mismatch might occur. Difficulties due to the mismatch between the gender stereotype and the reflexive pronoun might emerge on the reflexive pronoun and the post-pronoun region as in Experiment 1. For the prior disambiguating contexts, however, there is another possibility. The gender stereotype may be activated when *electrician* is processed; if so, it may then be eliminated once the antecedent for *electrician* has been retrieved from the discourse model and it becomes apparent that this entity is female. If the stereotype is eliminated from processing before the reader encounters the reflexive pronoun, then no mismatch effect should emerge on the pronoun, but rather an effect should emerge on or immediately after the second encounter of *electrician*. This effect would reflect the simultaneous activation of the male gender stereotype on the second encounter and the female *electrician* from earlier in the discourse.

Method

Participants

Thirty-six students from the University of Massachusetts participated in this study. Students were paid or received course credit for participation.

Materials

A total of 36 role names were selected from the earlier norming study. Half of the names received mean ratings of less than 2 (*strongly female*), and the other half received mean ratings of greater than 6 (*strongly male*). A series of 36 paragraph sets was written, one set for each of the roles chosen in the norming task. Example paragraphs are presented in Table 4. In each paragraph, the character in the stereotyped role was introduced in the second sentence by role name (*The electrician*). In the prior disambiguating context condition, this character's gender was explicitly stated in the second sentence as well (by means of an explicit noun that directly identified gender [*was a cautious **woman***] and a pronoun that specified gender [*who carefully secured **her** ladder . . .*]) . For the neutral context condition, these references to the character's gender were removed (*was cautious and carefully secured the ladder . . .*). The first and third sentences of the paragraph referred to a second character by proper name rather than role; this second character was always a different gender from the role-named character. This proper-named character was introduced in the paragraph because we found that if there was only one character in the paragraph, the role-named character, it became awkward to refer to this character by role rather than

by pronouns. In the target sentence—sentence 4—we wanted to refer to the role-named character using the role name in a definite noun phrase. Shifting the focus of the paragraph to the proper-named character in sentence 3 allowed us to then naturally refer to the role-named character by role name rather than pronoun in sentence 4.

As noted, sentence 4 was the target sentence. It always referred to the role-named character by role name in a definite noun phrase followed by a verb and a reflexive pronoun that specified the gender of the character. The specified gender either matched or mismatched the gender stereotype. For the prior disambiguating contexts, the gender specified earlier always matched the gender specified in the target sentence (that is, the gender specified in sentence 2 never mismatched the gender specified in sentence 4). The reflexive pronoun always appeared toward the center of the line, and the next one to four words formed the post-pronoun region (mean number of words in this region = 2.7).

There were four versions of each of the 36 paragraphs, created by the crossing of two factors: prior context (disambiguating vs. neutral) and stereotype consistency (match vs. mismatch). Each participant read nine paragraphs in each of the four conditions. Across the full set of participants, each paragraph appeared in all four conditions. In addition, there were 24 filler paragraphs that did not feature any mismatches of gender stereotype. Thus, the participants read a total of 60 paragraphs.

Apparatus. The apparatus was the same as that used in Experiment 1.

Procedure. The procedure was the same as that used in Experiment 1. The only change was that the stimuli were paragraphs which were displayed on five to seven lines of up to 80 characters. Again, the words in the paragraphs were arranged so that a given critical region appeared on only one line. Participants were given a yes–no question to answer after 25 of the 60 paragraphs (95% correct).

Results

Four percent of the data were eliminated due to track losses. In addition, fixation times that were more than three standard deviations from the mean were eliminated. Table 5 contains first pass times and go-past times for the following regions in the target sentence: the role name, the verb following the role name, the reflexive pronoun, and the post-pronoun region. Within the analyses of the role name and verb regions, no effects were significant.

For the reflexive pronoun, as in Experiment 1, readers spent longer in the first pass on the pronoun when there was a mismatch in the neutral contexts [$F_1(1,35) = 7.88$, $p < .01$; $F_2(1,35) = 4.26$, $p < .05$]. In contrast, with the prior disambiguating contexts, there was no difference between the mismatch and match conditions ($Fs < 1$). This resulted in a significant interaction of context and consistency [$F_1(1,35) = 6.88$, $p < .05$; $F_2(1,35) = 7.02$, $p < .05$]. Exactly the same pattern of results was observed for the go-past measure, with readers having longer go-past times for mismatch than for match in the neutral context [$F_1(1,35) = 7.99$, $p < .01$; $F_2(1,35) = 4.77$, $p < .05$], no difference with the prior disambiguating context ($Fs < 1$), and an interaction [$F_1(1,35) = 7.26$, $p < .05$; $F_2(1,35) = 7.48$, $p < .05$].

In the post-pronoun region, stereotype mismatch produced longer first pass times in the neutral contexts but not in the disambiguating contexts; this interaction of context and consistency was significant by participants and marginally so when tested against items variability [$F_1(1,35) = 5.51$, $p < .05$; $F_2(1,35) = 3.00$, $p = .092$]. Simple effects analyses indicated that the mismatch did lengthen first pass times in the neutral contexts [$F_1(1,35) = 9.52$, $p < .01$; $F_2(1,35) = 7.06$, $p < .05$], but not in the prior disambiguating contexts (both $Fs < 1$).

For the go-past measure within the post-pronoun region, there were significant main effects for both context and consistency, with readers producing longer go-past times for the gender mismatch conditions [$F_1(1,35) = 14.93$, $p < .001$; $F_2(1,35) = 27.91$, $p < .001$] and longer go-past times for the neutral contexts [$F_1(1,35) = 6.71$, $p < .05$; $F_2(1,35) = 5.51$, $p < .05$]. More importantly, however, these effects were qualified by an interaction of context and consistency [$F_1(1,35) = 10.56$, $p < .01$; $F_2(1,35) = 6.08$, $p < .05$]. Simple effects analyses indicated that readers produced longer go-past times when the reflexive pronoun mismatched the gender stereotype in the neutral contexts [$F_1(1,35) = 24.49$, $p < .001$;

Table 5
Mean First Pass and Go-Past Times (in Milliseconds)
for the Critical Regions in Experiment 2

Measure	Neutral Contexts			Prior Disambiguating Contexts		
	Mismatch	Match	Difference	Mismatch	Match	Difference
Role Name Region						
First pass	328	331	−3	329	335	−6
Go-past	370	366	4	351	367	−16
Verb Region						
First pass	316	311	5	316	295	21
Go-past	361	347	14	350	350	0
Reflexive Pronoun Region						
First pass	297	277	20	276	283	−7
Go-past	351	323	28	342	336	6
Post-Pronoun Region						
First pass	504	444	60	458	463	−5
Go-past	684	541	143	569	548	21

$F_2(1,35) = 23.95, p < .001]$, but there was no effect of gender mismatch in the disambiguating contexts (both $Fs < 1$).

For the disambiguating contexts, we also examined the region in the second sentence that explicitly stated the gender of the role-named character. We expected to find an effect of stereotype mismatch when the reader encountered the word *man* or *woman*. In Table 6, first pass and go-past times are given for the man/woman region in the disambiguating contexts. First pass times in the region were marginally longer when the gender was violated $[F_1(1,35) = 2.91, p = .097; F_2(1,35) = 4.10, p = .053]$. However, go-past times in this region were longer when the gender stereotype was violated $[F_1(1,35) = 8.67, p < .01; F_2(1,35) = 4.99, p < .05]$. Additional analyses indicated that this effect reflected an increase in the number of regressions out of this region to earlier regions in the text (15% when the stereotype was violated; 6% when it was not) $[F_1(1,35) = 6.94, p < .05; F_2(1,35) = 6.51, p < .05]$.

Discussion

The results demonstrate that a mismatch of gender stereotype with text information interferes with the reading process. This interference was observed on the man/woman region for the prior disambiguating contexts and on the reflexive pronouns in the neutral contexts when the discourse specified a gender that mismatched the stereotype. Thus, it is clear that gender stereotypes do play a role in discourse comprehension. However, the mismatch effect that appeared on the reflexive pronoun in the neutral contexts disappeared when the gender of the role character was stated earlier in the paragraph. This pattern of effects is consistent with the lexical reinterpretation model (Foss & Speer, 1991; Hess et al., 1995). Applying that model to the present paradigm, there is no difficulty integrating the reflexive pronoun into an ongoing discourse representation in which the referent's gender has earlier been established. The critical point within this framework is that although the gender may mismatch the stereotype, it does not mismatch the gender instantiated in the discourse model. Therefore, the pronoun is readily integrated into that discourse model.

As in Experiment 1, the effect of stereotype mismatch was generally quite immediate. Indeed, the size of the mismatch effect was consistent with results reported by Sturt (2003). In the prior disambiguating context, it was the noun man or woman that revealed the mismatch, and

the mismatch effect emerged immediately, before the participants' eyes went past that noun. Likewise, in the neutral context, it was the reflexive pronoun that revealed the mismatch, and the mismatch effect appeared immediately and continued after the participants' eyes moved on to the region following the pronoun. With respect to this finding, it is interesting that within the pronoun literature, there are a number of studies in which retrieval of the antecedent for a pronoun is not always immediate (e.g., Ehrlich & Rayner, 1983; Garrod, Freudenthal, & Boyle, 1994; Greene, McKoon, & Ratcliff, 1992; Sanford & Garrod, 1989).

GENERAL DISCUSSION

The results of Experiment 1 and the neutral condition of Experiment 2 demonstrate that when text and gender-stereotyped information conflict, the normal process of reading text suffers interference at the point where the conflict can first be realized. This may reflect immediate activation of the gender stereotype when the role name was read, resulting in a conflict when the reflexive pronoun was subsequently encountered. Alternatively, it is possible that the gender information was not activated until integration with the pronoun was required. The present data do not permit us to decide between these two possibilities.

Although gender mismatches resulted in comprehension difficulty in Experiment 1 and in the neutral condition of Experiment 2, there was no interference when the prior context specified the *man* or *woman*, a result consistent with the lexical reinterpretation model proposed by Hess et al. (1995). At first glance, this pattern of results might seem consistent with the contrast between context-independent and context-dependent properties of concepts proposed by Barsalou (1982). Context-independent properties are always activated when a particular concept is encountered (i.e., in neutral as well as related contexts), whereas context-dependent properties are only activated in a context that elicits these properties. Under this set of assumptions, the gender stereotype for a role name would be a context-independent property since it is activated in a neutral context. However, the hallmark of context-independent properties is that they are always activated. This is not the case here. Rather, a property that is reliably activated in neutral contexts (gender stereotype) is not activated in a context that renders that property untrue. Thus, this pattern of results provides evidence against the general claim that access of context-independent properties cannot be influenced by context.

It might be noted that the properties of male and female gender are mutually exclusive. These two properties cannot both be true of the same entity, and thus they differ from the properties that Barsalou (1982) identified as context-independent and context-dependent. The properties investigated by Hess et al. (1995), however, are similar to the context-independent properties proposed by Barsalou. For example, the fact that *English major* primes *poem* is presumably based on the activation of a central property of *English major* that might be considered

Table 6
Mean Go-Past Times (in Milliseconds)
and Percentage Regressions Out of the Man/Woman Region
in the Prior Disambiguating Contexts in Experiment 2

	Mismatch	Match	Difference
First pass	279	261	18
Go-past	366	305	61
% Regress out	15	6	9

context independent. Hess et al., however, eliminated this priming effect by embedding the target sentence in a context in which the literary aspects of *English major* were irrelevant. Thus, these studies also provide evidence against the claim that context-independent properties cannot be influenced by discourse context.

In contrast to the present results and to those of Hess et al. (1995), Rayner et al. (1994) found a different pattern of results for second encounters of ambiguous words. They presented biased ambiguous words preceded by information that disambiguated them toward their less-frequent meaning. In such contexts, fixation times on the target words are longer, compared with unambiguous control words, an effect that has been labeled the *subordinate bias effect* (Duffy, Kambe, & Rayner, 2001; Rayner et al., 1994). In Rayner et al.'s paragraphs, these ambiguous target words appeared twice, each time disambiguated by immediately preceding context, and each time disambiguated toward the same (less-frequent) meaning. The expectation was that the subordinate bias effect would disappear on second encounter because only the intended meaning would be accessed. Surprisingly, they observed a comparable subordinate bias effect for the second encounter of the target word, as well as for the first. This result was interpreted as indicating that, on the second encounter, lexical access proceeded just as it had on the first encounter, with no selective access of the intended meaning despite the biasing discourse context.

Although the Rayner et al. (1994) pattern of results would seem to conflict with those reported here in Experiment 2 and in Hess et al., there are a number of potentially critical differences between these studies. First, the target words in Rayner et al. were ambiguous. It is possible that processing differs for ambiguous words. Second, the Rayner et al. target words were not always coreferential with the earlier mention of the word (e.g., for one of the sample texts, the second mention of *speaker* was not coreferential with the first mention of *speaker*). Third, the intended meaning of the Rayner et al. target word was prestored in semantic memory, whereas the intended interpretations of the target role names in our Experiment 2 and in Hess et al. were created by the discourse context and not prestored in semantic memory. Rather, they were presumably stored in the current discourse model as part of the representation for the antecedent for the target role name. Finally, the contexts that Rayner et al. used might not have been as strong as the contexts used here (e.g., Rayner et al.'s readers had to infer what kind of *speaker* was being referred to; in our Experiment 2, the prior disambiguating text explicitly states that the electrician is a woman).

Finally, in the present experiments, we have used gender stereotype not as the direct focus of analysis but as a useful tool for studying comprehension processes in discourse context. Nevertheless, the research does converge with other studies on gender stereotypes that suggest that these stereotypes are automatically activated when a stereotyped role name is encountered and they play a part in processing, once activated (e.g., Banaji & Hardin, 1996). It is

somewhat disheartening that in a discourse context where no mention of gender has been made, readers simply assume that the electrician they are reading about is male and the nurse they are reading about is female. On a more positive note, our results also suggest that this automatic leap to conclusions about the gender of a particular character can be overridden, at least for a short time, by explicit specification of gender early in the paragraph. Further study might investigate the duration of this override.

REFERENCES

BANAJI, M. R., & HARDIN, C. D. (1996). Automatic stereotyping. *Psychological Science*, **7**, 136-141.

BARSALOU, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition*, **10**, 82-93.

CARREIRAS, M., GARNHAM, A., OAKHILL, J., & CAIN, K. (1996). The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *Quarterly Journal of Experimental Psychology*, **49A**, 639-663.

COOK, A., & MYERS, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory & Language*, **50**, 268-288.

DUFFY, S. A., KAMBE, G., & RAYNER, K. (2001). The effect of prior disambiguating context on the comprehension of ambiguous words: Evidence from eye movements. In D. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 27-44). Washington, DC: American Psychological Association.

DUFFY, S. A., MORRIS, R., & RAYNER, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory & Language*, **27**, 429-446.

DUFFY, S. A., & RAYNER, K. (1990). Eye movements and anaphor resolution: Effect of antecedent typicality and distance. *Language & Speech*, **33**, 1003-1019.

EHRLICH, K. (1983). Eye movements in pronoun assignment: A study of sentence integration. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 253-268). New York: Academic Press.

EHRLICH, K., & RAYNER, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning & Verbal Behavior*, **22**, 75-87.

FOSS, D. J., & SPEER, S. R. (1991). Global and local context effects in sentence processing. In R. R. Hoffman & D. S. Palermo (Eds.), *Cognition and the symbolic processes: Applied and ecological perspectives* (pp. 115-139). Hillsdale, NJ: Erlbaum.

GARROD, S., FREUDENTHAL, D., & BOYLE, E. (1994). The role of different types of anaphor on the on-line resolution of sentences in a discourse. *Journal of Memory & Language*, **33**, 39-68.

GARROD, S., & TERRAS, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory & Language*, **42**, 526-544.

GERNSBACHER, M. A., VARNER, K. R., & FAUST, M. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory & Cognition*, **16**, 430-445.

GORDON, P. C., & HENDRICK, R. (1998). Representation and processing of coreference in discourse. *Cognitive Science*, **22**, 389-424.

GREENE, S. V., MCKOON, G., & RATCLIFF, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 266-283.

HESS, D. J., FOSS, D. J., & CARROLL, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, **124**, 62-82.

KINTSCH, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, **95**, 163-182.

LIVERSEDGE, S. P., PATERSON, K. B., & PICKERING, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye movements in reading and scene perception* (pp. 201-222). Oxford: Elsevier.

OSTERHOUT, L., BERSICK, M., & MCLAUGHLIN, J. (1997). Brain poten-

tials reflect violations of gender stereotypes. *Memory & Cognition*, **25**, 273-285.

RAYNER, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372-422.

RAYNER, K., & DUFFY. S. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, **14**, 191-201.

RAYNER, K., PACHT, J., & DUFFY, S. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory & Language*, **33**, 527-544.

SANFORD, A. J., & GARROD, S. C. (1989). What, when, and how? Questions of immediacy in anaphoric reference resolution. *Language & Cognitive Processes*, **4**, SI 325-262.

SEIDENBERG, M. S., TANENHAUS, J. K., LEIMAN, J. M., & BIENKOWSKI, M.

(1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology*, **114**, 489-537.

STURT, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory & Language*, **48**, 542-562.

SWINNEY, D. A. (1979). Lexical access during sentence comprehension: (Re)consideration of context effects. *Journal of Verbal Learning & Verbal Behavior*, **18**, 645-659.

TILL, R. E., MROSS, E. F., & KINTSCH, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition*, **16**, 283-298.