The Effects of Contextual Constraints on Meaning Selection in the Mental Lexicon

Carol Joy Madden
FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

THE EFFECTS OF CONTEXTUAL CONSTRAINTS
ON MEANING SELECTION IN THE MENTAL LEXICON

By

CAROL JOY MADDEN

A Dissertation submitted to the
Department of Psychology
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Spring Semester, 2004
The members of the Committee approve the dissertation of Carol Joy Madden

__________________________________
Rolf A. Zwaan
Professor Directing Thesis

__________________________________
Leonard L. LaPointe
Outside Committee Member

__________________________________
Frank Johnson
Committee Member

__________________________________
Donald J. Foss
Committee Member

__________________________________
Joseph K. Torgeson
Committee Member

The Office of Graduate Studies has verified and approved the above named committee
members.
## TABLE OF CONTENTS

List of Tables ....................................................................................................................iv
Abstract .............................................................................................................................v

INTRODUCTION ............................................................................................................1
EXPERIMENT ONE ........................................................................................................5
EXPERIMENT TWO .......................................................................................................14
EXPERIMENT THREE ...................................................................................................20
GENERAL DISCUSSION ...............................................................................................30
CONCLUSION.................................................................................................................32
APPENDICES ..................................................................................................................34
ENDNOTES .....................................................................................................................40
REFERENCES .................................................................................................................41
BIOGRAPHICAL SKETCH ............................................................................................44
# LIST OF TABLES

1. Means (Standard Deviations) for Response Times in Experiment.........................10
2. Means (Standard Deviations) for Response Times in Experiment Two .................16
3. Means (Standard Deviations) for Response Times in Experiment Three Using Passage Comprehension ...........................................................................................................23
4. Means (Standard Deviations) for Response Times in Experiment Three Using Reading Span.............................................................................................................................25
5. Means (Standard Deviations) for Response Times With Middle Scores Removed in Experiment Three Using Reading Span....................................................................................27
6. Means (Standard Deviations) for Response Times With Middle Scores Removed in Experiment Three Using Passage Comprehension.....................................................28
7. Mean % (Standard Deviation %) for Accuracy Data in Experiment One ...............35
8. Mean % (Standard Deviation %) for Accuracy Data in Experiment Three Using Reading Span ..........................................................................................................................37
9. Mean % (Standard Deviation %) for Accuracy Data in Experiment Three Using Passage Comprehension ..............................................................................................37
In this study, three experiments were conducted to investigate how context affects meaning selection during the comprehension of sentences for more skilled and less skilled comprehenders. In all three experiments, a picture was presented that either matched or mismatched an object’s shape as implied by the preceding sentence context. The pattern of results across the three experiments showed that after a short delay (750ms) both more skilled and less skilled comprehenders had constrained their representations to the contextually appropriate meaning of the target word from the preceding sentence. However only more skilled comprehenders had constrained their representations to the contextually appropriate meaning of the target word immediately upon hearing it in the context of the sentence, whereas less skilled comprehenders required more processing time before the appropriate meaning was activated. These results are comparable to results obtained using homographs, and suggest that the mental lexicon may be comprised of stored instances of each experience with a word or concept, rather than a single stored meaning.
INTRODUCTION

One of the most amazing capacities of the human mind is the ability to perceive squiggles on a page or rapid patterns of sound waves and transform this visual or auditory code into meaningful concepts and ideas. These resulting representations and the process by which they are constructed have been the focus of many debates in cognitive psychology over the past century. However, it is generally agreed that successful language comprehension of a text requires the construction of a situation model, a mental model of the described state of affairs (Zwaan & Radvansky, 1998; Van Dijk & Kintsch, 1983). Traditionally, such mental models were thought to be propositional in nature, consisting of amodal representations. Amodal representations are descriptive and not grounded in the perceptual systems that function during interaction with the environment. However, a growing body of recent research has suggested that the situation model is actually a perceptual simulation of the situation described by a text (Glenberg & Kaschak, 2002; Kan, Barsalou, Solomon, Minor, & Thompson-Schill, 2003; Klatzky, Pellegrino, McCloskey, & Doherty, 1989; Morrow & Clark, 1988; Pecher, Zeelenberg, & Barsalou, 2003; Stanfield & Zwaan, 2001; Zwaan, Stanfield, & Yaxley, 2002).

According to this new framework, simulations are presumably constructed by reactivating the perceptual memory traces that were originally laid down in memory during previous experiences with the described situation (Barsalou, 1999; Barsalou, Simmons, Barbey, & Wilson, 2003). It is postulated that these experiential traces are the building blocks for all levels of language comprehension, from the word unit to larger discourse units (Zwaan, in press). Therefore, the stored meaning of a given word is comprised of as many stored instances of that word and its referent as we have experienced in our lifetime. However, in order to perceptually simulate a word in a sentence, the comprehender must specify which of the word’s experiential traces will become activated in the simulation. For instance, most people have a multitude of stored
experiences with viewing, petting, smelling, and hearing dogs. Upon hearing the sentence, “The dog chased after the tennis ball,” one must use context to choose the appropriate perceptual representation of “dog” from these stored experiences to import into the specific simulation of the described event. Here, the appropriate perceptual representation should not be a dog curled up on the kitchen floor, but rather a dog in mid stride, presumably following a tennis ball.

We propose that when a word is read or heard, all associated memory instances are activated. The activated set of instances determines the perceived meaning of the current word. Stored instances with contexts that better overlap with the current context are most highly activated and provide the strongest contribution to the meaning of a word. The current context includes the situation that is being described by language (the sentence or larger discourse in which the word occurs) as well as the actual situation (internal and external) being experienced by the listener/reader. This process of activation occurs in similar fashion to the memory model proposed by Hintzman (1986, 1988), in which activation continuously and automatically flows to any memory traces that are sufficiently similar to the retrieval environment. When context matching occurs between the memory traces and the currently attended situation, memory retrieval is facilitated (Hintzman, 2002; Hintzman, Block, & Summers, 1973). Thus, the more contextual overlap between a stored instance and the current context, the more likely that instance is to become activated and participate in meaning selection.

Of course, the perceptual input of hearing a given word will resonate with many instances of having heard that word before. This type of broad activation of overlapping perceptual traces may precede the activation of instances that share more contextual or conceptual features. Thus the time course of meaning activation can vary based on many factors such as word frequency, and sense enumeration. Many studies have investigated the topic of contextual constraints on meaning selection, and a large majority of these studies have used homographs. A homograph is a word that has one spelling (usually one pronunciation), but multiple unrelated meanings. The word “nail,” for instance, has the hammer-nail meaning and the finger-nail meaning. These studies typically use a homograph in a sentence, biasing the context toward one or the other meaning, such as, “She carefully applied the polish to her nail.” With careful probing of each meaning
following the presentation of a homograph in a biased context, one can determine the contextual influence on the subsequent activation levels of each stored meaning of the homograph.

This type of research has uncovered two main factors that influence the effects of context on lexical access, namely time and comprehension skill. The time course of contextual constraint on lexical access is often delayed, such that multiple meanings are initially activated, followed by the deactivation of the inappropriate meanings. Thus, only the contextually appropriate meaning is activated after a delay. This idea is consistent with the perceptual/contextual resonance-based activation of instances described above, and is supported by empirical studies (Conrad, 1974; Lucas, 1987; Onifer & Swinney, 1981; Swinney, 1979). However, some studies also show relatively immediate contextual constraint (Glucksberg, Kreuz, & Rho, 1986; Hess, Foss, & Carroll, 1995; Schvaneveldt, Meyer, & Becker, 1976; Simpson, 1981). This is also consistent with the described perceptual/contextual resonance-based activation of instances, as in some cases in which the context is sufficiently constraining, or the perceived word has a very narrow meaning. Thus, the initial, broad activation of perceptual traces of the sound or the sight of the word would largely activate the set of instances that are also contextually appropriate.

In addition, studies using homographs have shown that more skilled comprehenders are better able to use sentence context to constrain lexical access than less skilled comprehenders (Gernsbacher & Faust, 1991; Gernsbacher, Varner, & Faust, 1990; Madden, Kline, Zwaan, & Patrick, submitted; Van Petten, Weckerly, McIsaac, & Kutas, 1997). For instance, Madden, Kline, Zwaan, and Patrick examined the n400 component of ERP recordings during comprehension of homographs in sentences. The amplitude of this component is a measure of difficulty of integration into the overall context. In this study subjects read sentences in which the final word was contextually congruent, contextually incongruent, or lexically congruent (contextually incongruent, but associated with an inappropriate meaning of a nearby homograph).

As predicted, more skilled comprehenders produced large n400s in the lexically congruent and contextually incongruent conditions, but not in the contextually congruent condition, indicating that integration is facilitated by sentence-level context rather than by
irrelevant local associations. Less skilled comprehenders showed reduced n400s in the lexically congruent condition, indicating a deficit in the use of sentence-level context. This finding suggests that less skilled comprehenders relied too heavily on the initial, broad activation of perceptual traces of the sight of the target word, rather than the more constrained set of instances that also overlap with the perceptual sight of the word, but more importantly with the current context in which the word is used. This issue was further investigated in the present study.

The research described above identifies two main factors that have been shown to influence the effect of context on meaning selection, namely time and comprehension skill. In the current study, these two factors will be investigated using single-meaning words rather than homographs. According to the perceptual instance framework, every word is like a homograph in that the comprehender must disambiguate the appropriate meaning of the word from multiple possible representations (instances). If this is the case, then the two main factors outlined above, namely processing time and comprehension skill, should influence lexical access in normal words in the same manner as has been shown in homograph studies.
EXPERIMENT ONE

The current experiment employs single-meaning words rather than homographs to study the nature of lexical access. This experiment investigates the effects of comprehension skill and processing time on the ability to select contextually appropriate word meanings during sentence comprehension. The paradigm is modeled after that of Zwaan, Stanfield, and Yaxley (2002). In their study, a sentence describing the location of a target object was presented visually. The sentence was followed by a picture of the target object, which either matched or mismatched the contextual constraints of the preceding sentence. They found that participants were faster to respond to pictures that matched rather than mismatched the contextual constraints of the sentence.

In the current study, more skilled and less skilled comprehenders will hear sentences such as those used by Zwaan, Stanfield, and Yaxley, describing the location of a target object. Immediate and delayed picture presentations will be used to probe for context-appropriate and context-inappropriate perceptual features of the target words. Given the prior results of Zwaan et al. (2002), pictures matching the contextual constraints of the preceding sentence should be responded to more quickly than non-matching pictures. Also, consistent with previous research in the time course of meaning activation (Conrad, 1974; Lucas, 1987; Onifer & Swinney, 1981; Swinney, 1979), the contextually appropriate meaning of a word should be more available than the contextually inappropriate meaning at a delay rather than immediately.

This is because both the contextually appropriate and inappropriate meanings of a word may be available before the comprehender constrains his mental simulation of the described situation. However, by the time the picture is presented in the delayed condition, the mental simulation should be constrained and only the contextually appropriate meaning should be activated. And finally, as studies of comprehension skill
have shown (Gernsbacher & Faust, 1991; Gernsbacher, Varner, & Faust, 1990; Madden, Kline, Zwaan, & Patrick, submitted; Van Petten, Weckerly, McIsaac, & Kutas, 1997), more skilled comprehenders should be able to constrain their representations to the contextually appropriate meaning more quickly than less skilled comprehenders. Thus, pictures with contextually appropriate features will be more likely to match participants’ representations at a delay rather than immediately, and more skilled comprehenders will be able to activate these context-appropriate perceptual features more quickly than less skilled comprehenders.

Method

Participants

160 undergraduate students enrolled at Florida State University participated in the experiment as part of a course requirement\(^1\). All participants were native English speakers.

Materials

Twenty-eight experimental sentence pairs were adapted from Zwaan et al. (2002), each describing a target object in a location (See Appendix A for sample stimuli). The target object was always the final word of the sentence, for example, “In the sky/nest there was an eagle.” The sentence pairs were constructed such that each target object was described in two locations that implied different object shapes (e.g., eagle with wings outstretched or folded in). Two images, depicting the object in the two implied shapes, were also constructed to correspond to each experimental sentence pair. This yielded two sentences and two pictures for each target object. Each experimental sentence could be paired with a picture that matched or mismatched the implied shape of the target object, yielding four possible sentence-picture combinations. Participants were to see only one of these four possible combinations for each target object, so four experimental lists were created and counterbalanced with respect to implied shape and match/mismatch condition of the 28 target objects. For each of these four lists, the inter stimulus interval (ISI)
between the offset of the sentence and the presentation of the picture was varied. Pictures could appear immediately upon the offset of the final word of the sentence (0ms ISI), or 750 milliseconds after the offset of the final word of the sentence (750ms ISI), now yielding eight lists. ISI was varied between lists rather than within list, so that participants would not notice differences in picture onset delay from trial to trial.

In addition, 56 similar filler sentences and pictures were adapted from the earlier study. Each of the 112 sentences (56 filler sentences and 28 pairs of experimental sentences) was recorded to a wav file and the pictures were converted to black and white and scaled to occupy about three square inches on the center of the screen. The picture was mentioned in the sentence on half of the trials (all 28 experimental trials and 14 of the filler trials). Participants were told to respond as to whether or not the pictured object was mentioned in the preceding sentence, using keys labeled “Y” and “N.” On 24 of the filler trials, a question would appear after the picture-comparison response had been made. These questions required inferences about the sentences, and were included to ensure that participants would make an effort to process the sentences at a relatively deep level. Participants answered these questions using keys that were labeled with “Y” and “N” stickers. Because the participants didn’t know which sentences would be followed by a question, they had to comprehend each sentence to ensure a sufficient level of understanding. Both the experimental task and the Reading Span task described below were run on PCs with 19" flat-screen displays using the E-Prime stimulus presentation software (Schneider, Eschman, & Zuccolotto, 2002).

All participants completed a computer version of the Reading Span task (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002). On a given trial of the Reading Span task, a participant would read aloud a sentence, answer aloud “yes” or “no” as to whether it made sense, then read aloud the capitalized letter at the end of the sentence, remembering the letter for a later test. Participants would see 2, 3, 4, or 5 of these trials in a set before having to write the final letters they could recall for that set on a formatted sheet. Participants completed 3 practice sets and 12 experimental sets. The experimenter sat with the participants and controlled the progression from trial to trial to assure that participants had no time to rehearse the letters before they had to write them down at the end of the set.
This Reading Span task assesses a participant’s ability to maintain linguistic information in working memory while simultaneously processing sentences. This is a crucial component of the language comprehension process, as readers/listeners constantly hold words or clauses in working memory while processing other words or clauses until both can be integrated. Also, this measure correlates well with other measures of reading comprehension, such as verbal SAT (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). Based on a median split of scores on this test, participants were classified as more skilled or less skilled comprehenders.

Procedure

Participants were met one at a time in the laboratory and asked to sign a consent form. Then the participant was shown to another room where they completed the Reading Span task. The experimenter sat next to the participant as he completed this task, advancing the participant from trial to trial so that rehearsal or other memory strategies would be avoided. After completing the Reading Span task, participants began the picture-response experiment. On a given trial, participants pressed the spacebar to hear a sentence over headphones (e.g., “In the sky, there was an eagle.”). The sentence was followed by a picture presentation either immediately after the offset of the final word (0ms between sentence offset and picture presentation) or after a delay (750ms between sentence offset and picture presentation). The participant’s task was to make a response as to whether the pictured object was mentioned in the preceding sentence. If the pictured object was mentioned in the sentence, as was the case for all 28 of the experimental trials and 14 of the filler trials, the participant was to press the j-key, which was covered by a “Y” sticker. If the pictured object was not mentioned in the preceding sentence, as was the case for the 42 remaining filler trials, the participant was to press the f-key, which was covered by an “N” sticker. Once the experiment was finished, participants were debriefed and assigned partial course credit for their participation and dismissed.

Design and Predictions

Experiment 1 incorporated a 2 (match vs. mismatch) X 2 (ISI: 0 ms vs. 750 ms) X 2 (Comprehension Skill: high vs. low on Reading Span) mixed design, with match as a within subjects variable, and ISI and comprehension skill as between subjects variables.
On all 28 experimental trials, the presented image depicted the final word of the sentence (the target object), and thus required a yes response. However, on these experimental trials, the target object could be pictured in a shape that either matched or did not match the contextual constraints of the sentence. For example, if the participant heard, “In the sky there was an eagle”, a matching picture would be an eagle with wings outstretched and a mismatching picture would be an eagle with wings folded in. The participant should respond “yes” in either case, because the pictured object was mentioned in the sentence, but we expected the responses to be faster for matching rather than mismatching pictures.

To summarize, we expect that when the perceptual properties of the pictured object match the contextual constraints of the preceding sentence, responses will be facilitated relative to when the picture mismatches the contextual constraints of the preceding sentence. Thus responses will be faster for matching than for mismatching pictures. In addition, we predict that the contextually appropriate representation of the described situation is more likely to be constructed when the participant is given a longer time to comprehend the sentence before the test picture appears, so the 750ms ISI condition should show a stronger match effect than the immediate ISI condition. Furthermore, more skilled comprehenders should be able to constrain their representations of the target object to the contextually appropriate meaning more quickly than less skilled comprehenders, yielding faster response times for the matching rather than mismatching picture, even at the early ISI. Thus, we predict a three-way interaction among contextual match, ISI, and comprehension skill, whereby more skilled but not less skilled comprehenders show the match effect at the 0ms ISI, and both more skilled and less skilled comprehenders show the match effect at the 750ms ISI.

Results and Discussion

The dependent measure of interest was the participant’s time to respond to the presented picture. Analyses using response accuracy were also conducted for
Experiments 1 and 3 and are reported in Appendix B. Although List was included as a factor in all analyses for all three experiments (Pollatsek & Well, 1995; Raaijmakers, Schrijnemakers, & Gremmen, 1999), effects for the list variable will not be reported given their lack of theoretical relevance. All analyses were conducted both with variability due to subjects and variability due to items in the error term. These analyses are indicated by the subscripts 1 and 2 respectively for all three experiments. Incorrect responses were not included in the reported analyses, and any response time above or below 2.5 standard deviations from a participant’s mean for a given condition was removed prior to running the analyses. This constituted removal of less than 3% of the data.

Table 1
Means (Standard Deviations) for Response Times in Experiment One

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>809 (173)</td>
<td>911 (287)</td>
</tr>
<tr>
<td>750ms</td>
<td>816 (198)</td>
<td>893 (218)</td>
</tr>
</tbody>
</table>

The means and standard deviations for the response times from Experiment 1 are displayed in Table 1. Participants were classified as more skilled or less skilled based on a median split of scores on the Reading Span task [M = 31.6, SD = 4.2; More skilled m = 34.9, sd = 2.6; Less skilled m = 28.3, sd = 2.6]. The overall mixed analysis of variance (ANOVA) with list as a between subjects factor showed a clear effect of match, indicating that participants were faster to respond to the picture when it matched rather than mismatched the contextual constraints of the preceding sentence [F(1,144) = 32.92, p < .001, MSe = 11644; F(1,52) = 27.40, p < .001, MSe = 29577]. This effect was
qualified by a three-way interaction among match, ISI, and comprehension skill, indicating that the effect of match varied for more skilled and less skilled comprehenders at the two ISIs \(F_1(1,144) = 4.18, p < .05; F_2(1,52) = 27.40, p = .067, \text{MSe = 10321}\).

To understand the nature of this three-way interaction, separate analyses were conducted for the 0ms ISI and the 750ms ISI. The overall effect of match was observed in both the 0ms ISI analysis \(F_1(1,72) = 11.41, p < .01, \text{MSe = 11244}; F_2(1,52) = 19.22, p < .001, \text{MSe = 14162}\) as well as the 750ms analysis \(F_1(1,72) = 22.23, p < .001, \text{MSe = 12045}; F_2(1,52) = 20.83, p < .001, \text{MSe = 27098}\). However, the 0ms ISI analysis also revealed an interaction between match and comprehension skill \(F_1(1,72) = 7.21, p < .01; F_2(1,52) = 5.03, p < .05, \text{MSe = 12659}\), whereas no such interaction was present in the 750ms ISI analysis \(\text{both } F_2 < 1\). This suggests that more skilled comprehenders yielded a larger difference between match and mismatch responses than less skilled comprehenders only in the 0ms ISI condition.

Within-group comparisons for more skilled and less skilled comprehenders at each ISI were conducted to confirm this notion. Indeed, at the 0ms ISI the more skilled comprehenders showed the predicted match effect \(F_1(1,36) = 12.60, p < .01, \text{MSe = 16466}; F_2(1,52) = 16.47, p < .001, \text{MSe = 18198}\) whereas less skilled comprehenders did not show a significant difference between responses to the matching and mismatching pictures at the early ISI by subjects \(F_1 < 1\), although this comparison did reach significance by items \(F_2(1,52) = 4.20, p < .05, \text{MSe = 8623}\). At the 750ms ISI both more skilled and less skilled comprehenders showed the predicted match effect [more skilled: \(F_1(1,36) = 13.16, p < .01, \text{MSe = 9173}; F_2(1,52) = 11.14, p < .01, \text{MSe = 24270}; \text{less skilled: } F_1(1,36) = 9.90, p < .01, \text{MSe = 14916}; F_2(1,52) = 18.05, p < .001, \text{MSe = 16316}\].

The pattern of results for Experiment 1 demonstrates that the more skilled comprehenders were more quickly able to constrain their representation of the target word to the appropriate contextual meaning. The interactions and within-group comparisons show that the match effect at the immediate ISI was largely due to the more skilled comprehenders, who showed a stronger effect of match than the less skilled comprehenders. More skilled comprehenders were faster to verify pictures when the preceding sentence context matched rather than mismatched the shape of the depicted
object, whereas less skilled comprehenders showed only a weak advantage for matching vs. non-matching pictures, only significant in the contrast by items. At the 750ms ISI, both more skilled and less skilled comprehenders were faster to verify matching pictures than mismatching pictures. Just as in lexical ambiguity studies using homographs, the current experiment showed that when context is sufficiently constraining, more skilled comprehenders were able use sentence context to quickly activate the correct meaning of the target word, while less skilled comprehenders could only do so after the delay.

The observed pattern of responses supported the idea that the comprehension of any word requires that the contextually appropriate meaning be selected from a set of alternatives. At the immediate ISI, the less skilled comprehenders had not yet selected the appropriate meaning of the target word. The relatively fast responses of the less skilled comprehenders to both matching and mismatching pictures at the immediate ISI suggested that multiple meanings were activated at this point. We speculated that less skilled comprehenders had not yet constructed a useful simulation of the described situation at the immediate ISI because too many context-irrelevant instances of the target word were activated. At this stage, the less skilled comprehenders were performing like the less skilled comprehenders in the ERP study described in the introduction. They had broadly activated the stored instances that matched the perceptual input of the sound of the word, but had not yet constrained the set of activated instances to those that overlapped with the current context in which the word occurred. However, by the time the picture was presented in the 750ms ISI, they had had enough time to deactivate the irrelevant instances and construct a contextually appropriate simulation of the described situation. At this point, the matching picture was more similar than the mismatching picture to the activated perceptual simulation, and they showed the predicted match effect just as the more skilled comprehenders did.

A potential criticism of the present experiment is that the decision task taps processes that occur relatively late in the course of lexical processing, and thus are susceptible to post-access influences. The nature of the comparison-response task employed in Experiment 1 allowed for the alternative explanation that participants first accessed a singular, context-independent, lexical representation, and then manipulated that representation afterwards, according to the contextual constraints of the sentence.
This alternative explanation weakened the intended claim that context-specific representations were stored in the lexicon and were accessed when the current context was sufficiently similar. To address this criticism, another experiment was conducted in which a naming task was used. The naming task is thought to assess the lexical access process in its early stages, decreasing its susceptibility to post-access influences.
Experiment 2 was conceived to address the issue of higher-level cognitive processes that may have been implicated in the comparison-response task used in Experiment 1. Therefore, Experiment 2 was identical to Experiment 1 except that instead of making a response as to whether the pictured object was mentioned in the sentence, participants were instructed to name the pictured object as quickly as possible. The naming task was a more direct measure of lexical access, as it eliminated some of the cognitive processes that are required for the task used in Experiment 1. The decision task in Experiment 1 required recognition of the picture, lexical access of the name of the object, comparison of that name to the words in the sentence that preceded the picture, and an affirmative or negative response based on that comparison. In contrast, the naming task in Experiment 2 only required recognition of the pictured object, lexical access of the name of that object, and a vocal response, reporting the accessed name.

Method

Participants

160 undergraduate students enrolled at Florida State University participated in the experiment as part of a course requirement. All participants were native English speakers, and none had participated in the previous experiment.

Materials and Procedure

The materials and procedure for Experiment 2 were identical to those of Experiment 1, except that during the sentence-picture experiment, participants were told to simply name the pictured object as fast as possible, regardless of whether the picture
was related to the sentence. A microphone attached to the headphone set relayed the voice input to a response box where E-Prime software logged the latency of voice onset for each trial. An experimenter sat with the participant during the experiment to record any misnamed trials, or trials in which the microphone did not record the response correctly. It should be noted that it is possible for participants to perform this naming task without actually attending to the sentences. Thus, the same yes/no questions that required inferences about the sentences on 24 of the filler trials from Experiment 1 were again used in Experiment 2. Participants answered these questions using keys that were labeled with “y” and “n” stickers. Because the participants didn’t know which sentences would be followed by a question, they had to comprehend each sentence to ensure a sufficient level of understanding.

Design and Predictions

The design for Experiment 2 was identical to that for Experiment 1. It was a 2 (match vs. mismatch) X 2 (ISI: 0 ms vs. 750 ms) X 2 (Comprehension Skill: high vs. low on Reading Span) mixed design, with match as a within subjects variable, and ISI and comprehension skill as between subjects variables. Just as in Experiment 1, we expected that when the picture of the object matched the contextual constraints of the preceding sentence, lexical access would be facilitated, relative to when the picture mismatched the contextual constraints of the preceding sentence. In addition, we again predicted that the contextually appropriate representation was more likely to be constructed when participants were given a longer time to comprehend the sentences before the test picture appeared, so the 750 ms ISI condition should show a stronger match effect than the immediate ISI condition. Finally, we expected more skilled comprehenders to be faster to constrain their representations to the contextually appropriate meaning than less skilled comprehenders. Thus, we predicted another 3-way interaction among contextual match, ISI, and comprehension skill. However, given that the task in Experiment 2 did not require explicit comparison of the picture and the sentence as in Experiment 1, it is probable that these predicted effects would be weaker in this experiment.
Results and Discussion

The dependent measure of interest in Experiment 2 was the participant’s time to name the presented picture. Misnamed trials and trials in which the microphone did not record the response correctly were not included in the reported analyses. Because it was impossible to distinguish between misnamed trials and equipment error, accuracy data were not analyzed for Experiment 2. Naming latencies over 3 seconds as well as naming latencies over or under 2 standard deviations from a participant’s mean for a given condition were removed prior to running the analyses. This constituted removal of less than 6% of the data. The data from 2 of the 160 participants were excluded for having too many bad trials (misnames or equipment error) and two new participants were run to replace them. Also, the data from 5 of the 160 participants were excluded for having at least one of their condition means above 1000 ms. Seven new participants were run to replace these data.

Table 2
Means (Standard Deviations) for Response Times in Experiment Two

<table>
<thead>
<tr>
<th></th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISI</td>
<td>Match Mismatch</td>
<td>Match Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>671 (93) 698 (113)</td>
<td>686 (93) 687 (92)</td>
</tr>
<tr>
<td>750ms</td>
<td>661 (104) 686 (115)</td>
<td>711 (103) 731 (109)</td>
</tr>
</tbody>
</table>

The means and standard deviations for the naming times from Experiment 2 are displayed in Table 2. Participants were classified as more skilled or less skilled based on a median split of scores on the Reading Span task [M = 31.5, SD = 4.7; More skilled m = 35.3, sd = 2.6; Less skilled m = 27.6, sd = 2.7]. As in Experiment 2, the overall mixed
analysis of variance (ANOVA) with list as a between subjects factor showed a clear effect of match, indicating that participants were faster to name the pictured object when it matched rather than mismatched the contextual constraints of the preceding sentence \( F_1(1,144) = 11.16, p < .01, \text{MSe} = 2321; F_2(1,52) = 22.14, p < .001, \text{MSe} = 2587 \). In the overall analysis, neither the interaction between match and comprehension skill nor the three-way interaction among match, comprehension skill, and ISI reached significance. Nonetheless, given the observed pattern of results from Experiment 1 as well as the current hypotheses concerning the group differences, separate analyses for the 0ms ISI and the 750ms ISI were conducted.

In the immediate ISI analysis there was an effect of match (marginally significant by subjects), indicating that participants were already able to name the pictured objects more quickly when they matched rather than mismatched the contextual constraints of the preceding sentence \( F_1(1,72) = 2.76, p = .10, \text{MSe} = 2633; F_2(1,52) = 7.75, p < .01, \text{MSe} = 2836 \). This main effect was qualified by an interaction between match and comprehension skill, although this interaction did not reach conventional levels of significance \( F_1(1,72) = 2.51, p = .12, \text{MSe} = 2633; F_2(1,52) = 2.63, p = .11, \text{MSe} = 2190 \). This suggested that just like Experiment 1 the match effect at the immediate ISI was largely due to the more skilled comprehenders. Contrast tests for more skilled and less skilled comprehenders confirmed this notion, as the more skilled comprehenders showed a significant match effect \( F_1(1,36) = 4.82, p < .05, \text{MSe} = 2870; F_2(1,52) = 8.43, p < .01, \text{MSe} = 2980 \), whereas less skilled comprehenders showed no difference in naming the matching and mismatching pictures at the immediate ISI [both Fs <1].

The 750ms ISI analysis again yielded a main effect of match, indicating that overall, participants were faster to name the pictures when they matched rather than mismatched the contextual constraints of the preceding sentence \( F_1(1,72) = 10.10, p < .01, \text{MSe} = 2010; F_2(1,52) = 13.83, p < .001, \text{MSe} = 2616 \). However, the 750ms ISI analysis did not reveal an interaction between match and comprehension skill [both Fs <1], suggesting that both more skilled and less skilled comprehenders were contributing to the match effect in this condition. The contrast tests for the more skilled comprehenders showed an effect of match \( F_1(1,36) = 7.16, p < .05, \text{MSe} = 1700; F_2(1,52) = 12.58, p < .01, \text{MSe} = 2027 \). The less skilled comprehenders also showed an effect of match,
although only marginally significant [$F_1(1,36) = 3.58, p = .07, MSe = 2319; F_2(1,52) = 3.36, p = .07, MSe = 3562$].

The pattern of results for Experiment 2 was much the same as the pattern observed for Experiment 1. The more skilled comprehenders were able to constrain their representations to the appropriate meaning of the target word immediately upon hearing the sentence. More skilled comprehenders showed a significant match effect at both the immediate ISI and the 750 ms ISI. Less skilled comprehenders, however, were equally fast to name both the contextually appropriate and inappropriate pictures at the immediate ISI. Only after 750ms did they exhibit the predicted match effect, albeit only marginally significant.

Given the task differences between the two experiments, the effects in Experiment 2 were expected to be weaker. In Experiment 1 the participant was forced to compare the response-eliciting picture to the preceding sentence, so it is more likely that participant’s representation would affect the response to the picture. The task in Experiment 2 did not require the comprehenders to refer to the preceding sentence. Speeded naming of the pictured object could have been executed without associating it in any way with the preceding sentence, whereas in Experiment 1, the pictured object needed to be compared to the preceding sentence before a response could be made. Even though the sentences had to be kept active in short term memory in order to answer the comprehension questions that sometimes followed, the comprehender might have been able to perform the naming task independently of the sentence comprehension task. In fact, it is possible that comprehenders only constructed a coherent simulation of the sentence after the fact, when a question was encountered. Thus, it is likely that the match effect was constrained by the somewhat shallow level of processing that the sentence received as participants settled into the task demands of Experiment 2. Given that the task de-emphasized the sentence context, and the naming task assessed the lexical access process at an earlier stage than in Experiment 1, Experiment 2 provided a stronger test of the hypothesis than Experiment 1.

Although the statistical tests in Experiment 2 were not as convincing as they were in Experiment 1, the condition means looked very similar in the two experiments. The fact that the overall pattern of results was the same, even when a more direct measure like
naming time was used, provided an important manipulation check with regards to the comparison-response task used in Experiment 1. The consistent pattern of results from Experiment 2 addressed the possible criticism that the nature of the comparison-response task employed in Experiment 1 was susceptible to post-access influences. Experiment 2 provided additional support for the ideas asserted in Experiment 1 that context-specific representations are stored in the lexicon and accessed when the current context is sufficiently similar.

The results from the preceding two experiments tell a coherent story so far, but the claim could be made that too much time had elapsed between the presentation of disambiguating context (“sky” or “nest”) and the picture probe. Although the target word to be disambiguated was the final word of the sentence, and it occurred immediately before the picture probe in the 0ms ISI condition, the information that disambiguated the target word occurred much earlier in the sentence. It could be argued that participants were using the early information to construct a background context, which helped to disambiguate the target word before it was even presented. Indeed, research suggests that comprehenders do begin to construct a functional model of a sentence upon the first words, integrating each new word as it is encountered (Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2001; Van Petten & Kutas, 1991). These studies support the idea that meaning convergence becomes more constrained as greater context is provided, lowering the integration effort as the sentence unfolds. In order to address this issue of serial position of disambiguating contextual information in the sentence, a third experiment will be run in which the target occurs earlier in the sentence, and the disambiguating word occurs immediately prior to the picture probe.
EXPERIMENT THREE

Experiment 3 was intended to replicate the pattern of results from the previous two experiments and extend these findings to account for the immediate integration of disambiguating information in the sentence final position. Thus, the sentence structure for all experimental and filler sentences was altered to yield constructions such as, “There was an eagle in the sky/nest.” Whereas the target object used to occur in the sentence final position in Experiments 1 and 2, it occurred as the fourth word of the sentence in Experiment 3. In this experiment, the sentence final position was instead occupied by the disambiguating contextual information. Thus, Experiment 3 was intended to test the ability of participants to activate the contextually appropriate instances of the target word (e.g., “eagle”) immediately upon hearing the disambiguating contextual information. The instances that are contextually inappropriate with respect to object shape can only be deactivated upon hearing the final word, which, in specifying the target object’s location, constrains the shape that the object may take. If this deactivation of inappropriate instances is not immediate, the match effect will not be observed at the 0ms ISI.

In addition to the Reading Span task, participants in Experiment 3 also completed the Passage Comprehension subtest of the Woodcock Reading Mastery Test (Woodcock, 1998). This subtest is designed to measure participants’ ability to comprehend a short passage and use the context to identify a key word missing from the passage. This task will be further described in the Method section.

Method

Participants
178 undergraduate students enrolled at Florida State University participated in the experiment as part of a course requirement. All participants were native English speakers, and none had participated in the previous experiments.

Materials and Procedure

The materials and procedure for Experiment 3 were identical to those for Experiment 1, except that each sentence was rewritten to conform to the construction, “There was an X in the Y.” This is the same sentence format that was used in a previous study (Zwaan et al., 2002), which showed faster responses for pictures that matched rather than mismatched the contextual constraints of the preceding sentence. However, that study presented sentences visually and did not manipulate time between the sentence and the picture (always 250 ms) or include a measure of individual differences in comprehension skill. The comparison-response task from Experiment 1 was preferable to the naming task from Experiment 2, as the comparison-response task has been shown to yield larger effects than the naming task. This is most likely because the comparison-response task necessitates reference to the preceding sentence, while the naming task does not.

In addition, a second measure of comprehension skill was incorporated into the design for Experiment 3. While the Reading Span task is correlated with other measures of reading comprehension, such as verbal SAT (Daneman & Carpenter, 1980; Daneman & Merikle, 1996), this is an indirect measure of comprehension skill. Therefore, the Passage Comprehension subtest of the Woodcock Reading Mastery Test (Woodcock, 1998) was also administered as an additional and perhaps more direct measure of comprehension skill. This subtest is designed to measure participants’ ability to comprehend a short passage and use the passage context to identify a key word missing from the passage. There were a total of 44 items on the Passage Comprehension subtest, and raw scores for number correct were used in the analyses. Thus, this additional measure provides a different assessment of comprehension skill and allows for comparison of the results when participants are assigned to skill groups based on the two different measures.
Design and Predictions

As in the previous two experiments, Experiment 3 used a 2 (match vs. mismatch) \( \times 2 \) (ISI: 0 ms vs. 750 ms) \( \times 2 \) (Comprehension Skill: high vs. low on Reading Span or Passage Comprehension) mixed design, with match as a within subjects variable, and ISI and comprehension skill as between subjects variables. Because we used two different measures of comprehension skill in Experiment 3, all analyses were conducted twice. Again, we expected that when the picture of the object matched the contextual constraints of the preceding sentence, responses would be facilitated, relative to when the picture mismatched the contextual constraints of the preceding sentence. In addition, we predicted that the 750ms ISI condition would show a stronger match effect than the immediate ISI condition, as the activation of meaning selection may take time to settle. Finally, we expected more skilled comprehenders to be faster to constrain their representations to the contextually appropriate meaning than less skilled comprehenders. Thus, we predicted another 3-way interaction among contextual match, ISI, and comprehension skill. Given that the disambiguating information occurred so late in the sentence, it was possible that it would be more difficult and thus take longer to constrain representations once the final word was presented than it did in the previous experiments. However, we still predicted that overlapping context would automatically activate the contextually appropriate instances and more skilled comprehenders would show the effect at both the immediate and the late ISIs.

Results and Discussion

The dependent measure of interest was the participant’s time to respond to the presented picture. Just as in Experiment 1, incorrect responses were not included in the reported analyses, and any response time above or below 2.5 standard deviations from a participant’s mean for a given condition was removed prior to running the analyses. This constituted removal of less than 3% of the data. All analyses were run twice, using both Passage Comprehension and Reading Span as the skill-grouping measure. The results for
analyses using Passage Comprehension as the skill-grouping measure are reported first, followed by the results for analyses using Reading Span.

Table 3
Means (Standard Deviations) for Response Times in Experiment Three Using Passage Comprehension

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>783 (206)</td>
<td>836 (207)</td>
</tr>
<tr>
<td>750ms</td>
<td>825 (256)</td>
<td>932 (312)</td>
</tr>
</tbody>
</table>

Passage Comprehension

Participants were classified as more skilled or less skilled based on a median split of scores on the Passage Comprehension subtest of the WRMT [M = 32.8, SD = 4.3; More skilled m = 36.3, sd = 2.2; Less skilled m = 29.3, sd = 2.8]. The means and standard deviations for the response times from Experiment 3 using Passage Comprehension as the skill-grouping measure are displayed in Table 3. The overall mixed analysis of variance (ANOVA) with list as a between subjects factor once again showed a clear effect of match, indicating that participants were faster to respond to the picture when it matched rather than mismatched the contextual constraints of the preceding sentence \[F_1(1,144) = 19.02, p < .001, MSe = 12313; F_2(1,52) = 25.32, p < .001, MSe = 19735\]. This effect was qualified by an interaction between match and comprehension skill, indicating that the strength of the match effect differed between the more skilled and less skilled comprehenders \[F_1(1,144) = 3.72, p = .056; F_2(1,52) = 8.35, p < .01, MSe = 10106\].
To understand the nature of this interaction, and to report analyses consistent with those conducted in the previous two experiments, separate analyses were conducted on each of the ISI conditions. The analysis for the 0ms ISI yielded a main effect of match \(F_1(1,72) = 3.82, p = .05, \text{MSe} = 15829; F_2(1,52) = 9.88, p < .01, \text{MSe} = 14947\], but no interaction between match and comprehension skill was observed (although this interaction did approach significance by items \(F_2(1,52) = 2.12, p = .15, \text{MSe} = 8597\]).

Regardless, investigation of the within-group comparisons was warranted by the current hypothesis as well as the observed patterns of results in the previous two experiments. The within-group comparisons showed an effect of match for more skilled comprehenders at the immediate ISI \(F_1(1,36) = 5.56, p < .05, \text{MSe} = 10141; F_2(1,52) = 14.00, p < .001, \text{MSe} = 9633\], but not for less skilled comprehenders (although the match effect approached significance by items \(F_1 < 1; F_2(1,52) = 2.23, p = .14, \text{MSe} = 13911\]). Thus, it appeared that more skilled comprehenders were able to respond to the matching pictures more quickly than mismatching pictures immediately upon hearing the disambiguating context information, whereas less skilled comprehenders were not able to discriminate between the pictures as well at the 0ms ISI.

The analysis for the 750ms ISI also yielded a main effect of match \(F_1(1,72) = 21.87, p < .001, \text{MSe} = 8797; F_2(1,52) = 29.24, p < .001, \text{MSe} = 12950\]. In addition, there was an interaction between match and comprehension skill \(F_1(1,72) = 5.14, p < .05; F_2(1,52) = 5.75, p < .05, \text{MSe} = 13219\], indicating that the match effect was stronger for more skilled than for less skilled comprehenders. This interpretation was confirmed as the within-group comparisons yielded a strong effect of match for the more skilled comprehenders \(F_1(1,36) = 27.50, p < .001, \text{MSe} = 7713; F_2(1,52) = 24.10, p < .001, \text{MSe} = 16473\], but a match effect that did not reach conventional levels of significance by subjects for the less skilled comprehenders \(F_1(1,36) = 2.58, p = .12, \text{MSe} = 9880; F_2(1,52) = 5.95, p < .05, \text{MSe} = 9695\].

These results are consistent with the previous two experiments as well as our current hypotheses. The more skilled comprehenders were able to constrain their representations to the appropriate meaning of the target word immediately upon hearing the disambiguating location information, but the less skilled comprehenders were less able to do so. In fact, the less skilled comprehenders showed only a weak effect of match
even after a delay. This might be due to the fact that the information useful in disambiguating the meaning of the target word came later in the sentence than in Experiments 1 and 2.

Table 4
Means (Standard Deviations) for Response Times in Experiment Three Using Reading Span

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>820 (286)</td>
<td>856 (251)</td>
</tr>
<tr>
<td>750ms</td>
<td>808 (251)</td>
<td>890 (297)</td>
</tr>
</tbody>
</table>

Reading Span
Participants were classified as more skilled or less skilled based on a median split of scores on the Reading Span task [M = 31.7, SD = 4.7; More skilled m = 35.4, sd = 2.6; Less skilled m = 28.1, sd = 3.2]. The means and standard deviations for the response times from Experiment 3 using Reading Span as the skill-grouping measure are displayed in Table 4. The overall mixed analysis of variance (ANOVA) with list as a between subjects factor once again showed a clear effect of match, indicating that participants were faster to respond to the picture when it matched rather than mismatched the contextual constraints of the preceding sentence [F₁(1,144) = 21.73, p < .001, MSe = 11098; F₂(1,52) = 25.83, p < .001, MSe = 189270]. This effect was qualified by a three-way interaction among match, ISI, and comprehension skill (did not reach conventional levels of significance by subjects), indicating that the match effect differed between the more skilled and less skilled comprehenders at each ISI [F₁(1,144) = 2.23, p = .14; F₂(1,52) = 5.97, p < .05, MSe = 10598].
To understand the nature of this interaction, separate analyses were conducted on each of the ISI conditions. The analysis for the 0ms ISI yielded a main effect of match [$F_1(1,72) = 6.20, p = .05, \text{MSe} = 15706; F_2(1,52) = 14.18, p < .001, \text{MSe} = 13965$], but no interaction between match and comprehension skill was observed (although this interaction did approach significance by items [$F_2(1,52) = 2.67, p = .10, \text{MSe} = 12336$]). Investigation of the within-group comparison for the more skilled comprehenders at the immediate ISI showed an effect of match that only approached significance in the item analysis [$F_1(1,36) = 1.78, p = .19, \text{MSe} = 14442; F_2(1,52) = 2.83, p = .098, \text{MSe} = 11985$], indicating that more skilled comprehenders were hardly able to constrain their representations to the contextually appropriate meaning of the target word at the immediate ISI. However, less skilled comprehenders showed a significant effect of match at the immediate ISI [$F_1(1,36) = 4.64, p < .05, \text{MSe} = 16971; F_2(1,52) = 13.84, p < .001, \text{MSe} = 14316$]. Thus, it appears that immediately upon hearing the disambiguating context information, less skilled but not more skilled comprehenders were able to respond to the matching pictures more quickly than mismatching pictures. This finding is inconsistent with the results reported for the previous two experiments and will be further examined after the analyses for the 750ms ISI condition are reported.

The analysis for the 750ms ISI also yielded a main effect of match [$F_1(1,72) = 22.56, p < .001, \text{MSe} = 6489; F_2(1,52) = 23.43, p < .001, \text{MSe} = 11856$]. In addition, there was a marginally significant interaction between match and comprehension skill [$F_1(1,72) = 2.90, p = .09; F_2(1,52) = 3.93, p = .05, \text{MSe} = 7487$], indicating that the match effect was stronger for more the skilled than for less skilled comprehenders. Just as in the Passage Comprehension analysis, this interpretation was confirmed as the within-group comparisons yielded a strong effect of match for the more skilled comprehenders [$F_1(1,36) = 19.17, p < .001, \text{MSe} = 7045; F_2(1,52) = 19.80, p < .001, \text{MSe} = 12322$], and a significant, yet weaker match effect for the less skilled comprehenders [$F_1(1,36) = 5.08, p < .05, \text{MSe} = 5933; F_2(1,52) = 9.00, p < .01, \text{MSe} = 7022$].

Overall, the results for Experiment 3 using Reading Span as the skill-grouping measure were inconsistent with the previously reported patterns of results. Here, the less skilled comprehenders showed a stronger effect of match than the more skilled comprehenders did at the immediate ISI, which was problematic for the current
hypothesis. This might have been due to participants scoring near the middle of the distribution on the comprehension skill measure, which may have obscured the pattern of results. To investigate the apparent inconsistency in the data for Experiment 3, the within-group comparisons were recalculated after excluding the participants who scored at or near the median of the distribution of Reading Span scores. This procedure addressed the potential problem that some more skilled comprehenders may have scored near the middle of the distribution and thus may have been analyzed as an less skilled comprehender and vice versa. Once these middle scores were excluded, the more skilled and less skilled groups were more polarized [Reading Span mid-removed: More skilled m = 36.6, sd = 2.1; Less skilled m = 27.0, sd = 2.9] and had much less likelihood of including participants who did not belong in each group. Although the pattern of results for the Passage Comprehension data was not inconsistent, the same procedure was applied to these data to fully investigate the pattern of results in Experiment 3 as well [Passage Comprehension mid-removed: More skilled m = 37.1, sd = 2.0; Less skilled m = 28.5, sd = 2.4]. The new resulting means and standard deviations for both Reading Span and Passage Comprehension by subjects and items are displayed in Table 5 and Table 6.

Table 5
Means (Standard Deviations) for Response Times With Middle Scores Removed in Experiment Three Using Reading Span

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>783 (251)</td>
<td>837 (241)</td>
</tr>
<tr>
<td>750ms</td>
<td>814 (214)</td>
<td>888 (284)</td>
</tr>
</tbody>
</table>
Table 6
Means (Standard Deviations) for Response Times With Middle Scores Removed in Experiment Three Using Passage Comprehension

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>759 (203)</td>
<td>822 (211)</td>
</tr>
<tr>
<td>750ms</td>
<td>800 (224)</td>
<td>885 (282)</td>
</tr>
</tbody>
</table>

The recalculated within-group comparisons using the data with excluded middle scores on the Reading Span and Passage Comprehension did in fact show a pattern of results that was more consistent with the findings from Experiments 1 and 2. At the immediate ISI, the more skilled comprehenders now showed a significant effect of match by subjects and by items for both Reading Span and Passage Comprehension [Reading Span: $F_1(1,34) = 6.17, p < .05, MSe = 9987; F_2(1,52) = 6.64, p < .05, MSe = 13542$; Passage Comprehension: $F_1(1,31) = 7.72, p < .01, MSe = 8585; F_2(1,52) = 7.33, p < .01, MSe = 15675$], indicating that they had already constrained their representation to the appropriate meaning of the target word. The less skilled comprehenders did not show a match effect at the immediate ISI for Passage Comprehension by subjects nor by items, nor for Reading Span by subjects, although the match effect by items for Reading Span was significant [$F_2(1,52) = 8.49, p < .01, MSe = 16043$]. This suggested that the less skilled comprehenders may had constrained their representations to the appropriate meaning of the target word, but clearly the match effect was weaker than that elicited by the more skilled comprehenders. At the 750ms ISI, the more skilled comprehenders again showed an effect of match by subjects and by items for both Reading Span and Passage Comprehension [Reading Span: $F_1(1,36) = 6.74, p < .05, MSe = 12338; F_2(1,52)$]
Finally, the less skilled comprehenders again showed a weak effect of match at the 750ms ISI, which did not reach conventional levels of significance in two of the four analyses [Reading Span: \( F_1(1,29) = 2.29, p = .14, \text{MSe} = 12338; F_2(1,52) = 4.15, p < .05, \text{MSe} = 18673 \); Passage Comprehension: \( F_1(1,35) = 2.48, p = .12, \text{MSe} = 10193; F_2(1,52) = 4.68, p < .05, \text{MSe} = 13550 \)].

This follow up investigation supports the idea that participants near the median of the distribution of Reading Span scores were inadvertently sorted into the wrong skill groups and thus caused the apparent inconsistency in the Reading Span analyses for Experiment 3. The original pattern of results, showing a match effect for less skilled but not the more skilled comprehenders at the immediate ISI, were most likely obscured by misclassified participants. Once the scores around the median were removed from the analysis the predicted effects emerged, showing a stronger effect of match for the more skilled comprehenders than the less skilled comprehenders at the 0ms ISI.

Also, once the middle scores were removed from the distribution of Reading Span scores, the pattern of results for the more skilled and less skilled comprehenders at both the immediate and delayed ISI are much the same as those for the Passage Comprehension data. Although the two measures are not highly correlated in the current sample (\( r = .236, p < .001 \)), they seem to be tapping a similar comprehension construct in their discrimination between more skilled and less skilled comprehenders.
GENERAL DISCUSSION

The current study was aimed at better understanding the process of lexical access and meaning selection as it occurs within the perceptual instance-based framework. The three experiments presented here employed sentence-picture tasks to assess the influence of comprehension skill and processing time on the use of contextual constraints during meaning selection. Experiment 1 used a comparison-response task in which the final word of the sentence was the word to be disambiguated. This experiment yielded a 3-way interaction among contextual match, ISI, and comprehension skill, whereby only the more skilled comprehenders showed the predicted match effect at the immediate ISI, but all comprehenders showed the match effect at the later ISI. Experiment 2 provided a manipulation check for Experiment 1, in that a more direct measure of lexical access was used. This experiment replicated the pattern of results reported in Experiment 1 using a naming task. However, the effects in Experiment 2 were weaker than those in Experiment 1, probably due to the decreased reference to the preceding sentence in the naming task. Experiment 3 used the original comparison-response task, but with sentences in which the final word of the sentence was the disambiguating information rather than the word to be disambiguated. Consistent with the previous findings, Experiment 3 showed an effect of match for more skilled comprehenders at the immediate and the delayed ISI, and a very weak effect of match for the less skilled comprehenders at the immediate ISI that gained strength at the later ISI. However, these results were only observed once the middle of the comprehension skill distribution was excluded.

One difference between the overall pattern of results for Experiment 3 and the previous two experiments is that in Experiment 3, the less skilled comprehenders seem to show a weak match effect at the 0ms ISI. In Experiments 1 and 2, the less skilled
comprehenders show no significant match effect at the immediate ISI. This difference may simply be due to error variability in the data for the experiments, but it also might be due to the difference between the sentence constructions used in Experiment 1 and Experiment 3.

In Experiment 3, the target word is presented early in the sentence and the disambiguating information is withheld until the last word of the sentence. All stored instances that match the sound of the target word will be activated when it is encountered early in the sentence, and the deactivation of the contextually inappropriate instances cannot begin until this last word is heard. In Experiment 1, the activation of potentially consistent instances AND the deactivation of the contextually inappropriate instances must occur at the same time, as the target word is presented last. Thus, if it is the deactivation of inappropriate instances that is problematic for less skilled comprehenders, they should show a null effect for match at the immediate ISI for both types of sentence constructions, as deactivation is immediately required in both cases.

However, if less skilled comprehenders show a match effect at the immediate ISI in Experiment 3 but no match effect at the immediate ISI in Experiment 1, then it seems there is some benefit of having extra time to activate the set of potentially useful instances in Experiment 3. In this case, the process of deactivation is not problematic for less skilled comprehenders, as they can execute this process immediately and exhibit the match effect for sentences like those in Experiment 3. It is only when the process of deactivation, which they can execute immediately, is paired with the process of activation that the less skilled comprehenders fail to show a match effect. In this case it would appear that the process of instance activation is problematic for the less skilled comprehenders. There is very weak empirical evidence to support this idea in the present study, and the above explanation is purely speculative. However, if sentence construction is manipulated within rather than between experiment as was done in the present study, this idea could be tested in future research.
CONCLUSION

The present set of results suggests that any given word must be disambiguated before its referent can be used in a perceptual simulation of the described event. The same process of lexical access that occurs for homographs seems to be operating on so-called single-meaning words. Accordingly, the effects of time and comprehension skill on lexical access that have been shown in studies using homographs are also now extended to single-meaning words.

In all three experiments the pattern of results for less skilled comprehenders is consistent with similar studies that have used homographs to show that the time course of contextual constraint on lexical access is often delayed, such that multiple meanings are initially activated, followed by the deactivation of the inappropriate meanings (Conrad, 1974; Lucas, 1987; Onifer & Swinney, 1981; Swinney, 1979). In contrast, the pattern of results for the more skilled comprehenders is consistent with studies that show immediate contextual constraint (Glucksberg et al., 1986; Hess et al., 1995; Schvaneveldt et al., 1976; Simpson, 1981). Also, these results are consistent with studies that have shown that more skilled comprehenders are better able to use sentence context to constrain lexical access than less skilled comprehenders (Gernsbacher & Faust, 1991; Gernsbacher et al., 1990; Madden et al., submitted; Van Petten et al., 1997).

These results support the idea that even single-meaning words have a variety of stored perceptual representations, similar to the traditional idea of multiple stored meanings of homographs. Furthermore, this is consistent with the idea that the lexicon is comprised of stored instances of each experience with a word or concept, rather than a single stored meaning. According to this new framework, each instance is a viable candidate for activation in a perceptual simulation of the current sentence, and overlapping context determines the subset of stored instances that are consistent with the
current use of the word. These instances become activated and are incorporated into the perceptual simulation of the described event.
APPENDIX A

SAMPLES OF SENTENCE-PICTURE PAIRS

In the skillet/refrigerator there was an egg.

In the nest/sky there was an eagle.

In the box/bowl there was spaghetti.
APPENDIX B

ACCURACY DATA FOR EXPERIMENTS ONE AND THREE

Experiment One

The accuracy data for Experiment 1 are displayed in Table 7. The overall mixed analysis of variance (ANOVA) with list as a between subjects factor showed an effect of match whereby responses were more accurate for matching than mismatching pictures \[ F_1(1,144) = 32.94, p < .001, \text{MSe} = 2.114 \times 10^{-3}; F_2(1,52) = 14.23, p < .001, \text{MSe} = 7.042 \times 10^{-3} \]. In addition, there was an interaction between match and ISI, indicating that the match effect was stronger at the earlier ISI \[ F_1(1,144) = 6.27, p < .05; F_2(1,52) = 5.16, p < .05, \text{MSe} = 3.642 \times 10^{-3} \]. There was an effect of match at both the 0ms ISI \[ F_1(1,72) = 25.27, p < .001, \text{MSe} = 2.843 \times 10^{-3}; F_2(1,52) = 15.92, p < .001, \text{MSe} = 6.463 \times 10^{-3} \] as well as the 750ms ISI \[ F_1(1,72) = 7.98, p < .01, \text{MSe} = 1.385 \times 10^{-3}; F_2(1,52) = 3.82, p = .056, \text{MSe} = 4.220 \times 10^{-3} \].

Table 7

Mean % (Standard Deviation %) for Accuracy Data in Experiment One

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>98 (04)</td>
<td>94 (09)</td>
</tr>
<tr>
<td>750ms</td>
<td>98 (04)</td>
<td>96 (05)</td>
</tr>
</tbody>
</table>
Experiment Three

The accuracy data for Experiment 3 are displayed in Table 8 (Reading Span) and Table 9 (Passage Comprehension). The overall mixed analysis of variance (ANOVA) for Reading Span with list as a between subjects factor showed an effect of match whereby responses were more accurate for matching than mismatching pictures \([F(1,144) = 8.71, p < .01, MSe = 3.037E-03; F(1,52) = 5.54, p < .05, MSe = 6.774E-03]\). In addition, there was a three-way interaction among match, comprehension skill, and ISI, whereby the more skilled comprehenders showed the match effect at the 0ms ISI but not at the 750ms ISI, whereas less skilled comprehenders only showed the match effect at the 750ms ISI and not at the 0ms ISI \([F(1,144) = 2.80, p = .096; F(1,52) = 6.32, p < .05, MSe = 1.870E-03]\). At the 0ms ISI there was a main effect of match \([F(1,72) = 4.12, p < .05, MSe = 3.690E-03; F(1,52) = 4.07, p < .05, MSe = 5.309E-03]\). Also, there was an interaction between match and comprehension skill (did not reach conventional levels of significance by subjects) \([F(1,72) = 2.20, p = .14; F(1,52) = 4.07, p < .05, MSe = 2.809E-03]\), indicating that the match effect was only present for the more skilled comprehenders. At the 750ms ISI there was only a main effect of match \([F(1,72) = 4.78, p < .05, MSe = 2.384E-03; F(1,52) = 4.71, p < .05, MSe = 3.424E-03]\). Within-group comparisons showed significant match effects only for the more skilled comprehenders at the 0ms ISI \([F(1,36) = 5.09, p < .05, MSe = 4.474E-03; F(1,52) = 7.12, p < .05, MSe = 4.526E-03]\), and for the less skilled comprehenders at the 750ms ISI \([F(1,36) = 3.59, p = .066, MSe = 3.012E-03; F(1,52) = 6.68, p < .05, MSe = 2.260E-03]\).

The overall mixed analysis of variance (ANOVA) for Passage Comprehension with list as a between subjects factor showed an effect of match whereby responses were more accurate for matching than mismatching pictures \([F(1,144) = 16.54, p < .001, MSe = 3.176E-03; F(1,52) = 11.15, p < .01, MSe = 6.714E-03]\). In addition, there was a main effect of ISI whereby responses were more accurate at the 750ms ISI than the 0ms ISI \([F(1,144) = 4.63, p < .05, MSe = 3.825E-03; F(1,52) = 10.05, p < .01, MSe = 1.881E-03]\).
Table 8  
Mean % (Standard Deviation %) for Accuracy Data in Experiment Three  
Using Reading Span  

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>98 (05)</td>
<td>94 (10)</td>
</tr>
<tr>
<td>750ms</td>
<td>98 (05)</td>
<td>97 (05)</td>
</tr>
</tbody>
</table>

Table 9  
Mean % (Standard Deviation %) for Accuracy Data in Experiment Three  
Using Passage Comprehension  

<table>
<thead>
<tr>
<th>ISI</th>
<th>More Skilled Comprehenders</th>
<th>Less Skilled Comprehenders</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Match</td>
<td>Mismatch</td>
</tr>
<tr>
<td>0ms</td>
<td>97 (06)</td>
<td>95 (08)</td>
</tr>
<tr>
<td>750ms</td>
<td>98 (04)</td>
<td>96 (06)</td>
</tr>
</tbody>
</table>
Date: 7/15/2003

Carol J. Madden
MC: 1270

From: David Quadagno, Chair

Dept.: Psychology

Re: Reapproval of Use of Human subjects in Research:
Pictures and Language

Your request to continue the research project listed above involving human subjects has been approved by the Human Subjects Committee. If your project has not been completed by 7/8/2004 please request renewed approval.

You are reminded that a change in protocol in this project must be approved by resubmission of the project to the Committee for approval. Also, the principal investigator must report to the Chair promptly, and in writing, any unanticipated problems involving risks to subjects or others.

By copy of this memorandum, the Chairman of your department and/or your major professor are reminded of their responsibility for being informed concerning research projects involving human subjects in their department. They are advised to review the protocols of such investigations as often as necessary to insure that the project is being conducted in compliance with our institution and with DHHS regulations.

Cc: R. Zwaan
HSC No. 2003.377-R
Informed Consent Form

I freely and voluntarily consent to participate in the research project entitled "Pictures and Language". This project is being conducted under the direction of Dr. Rolf Zwaan, who is a professor in the Department of Psychology at FSU. This is a project that the Zwaan Lab is conducting in accordance with a "Language Representation" grant funded by the National Institute of Mental Health. I understand that this experiment does not present any risk to me. I also understand that I will receive course credit for participation in this experiment, and that I may withdraw my participation at any time without penalty or loss of credit. I understand that I will be reading short narrative passages/sentences and responding to pictures during the hour-long experiment. I understand that no individual results will be reported, but only group means. I understand that my results will be coded by an arbitrarily assigned participant number only (not by name) and kept in strict confidentiality to the extent allowed by the law. I understand that I may contact Carol Madden (644-1425) if I have any questions about this project, even after the experimental session is over. I understand that my consent may be withdrawn at any time without prejudice or loss of credit. I have read and understood this consent form.

Participant Signature ____________________________ Date ____________________________

It is helpful for our research if we can use SAT or ACT scores as a separate variable in our analyses. However, we need your permission to access these scores that the university keeps on file. Of course, this data will be coded with your subject numbers rather than your names, and will be kept anonymous and confidential. Please provide your initials below if you permit the use of your scores in our analyses.

Initials ______________
ENDNOTES

1 A total of 184 students actually participated in Experiment 1 to yield 160 data sets that could be used in the analysis design; ten more skilled and ten less skilled comprehenders for each of the four stimulus lists at each of the two ISI presentations. Similarly, more than 160 students participated in Experiments 2 and 3 to obtain the desired number of ten more skilled and ten less skilled participants per design cell.

2 The same procedure was also applied to the within-group comparisons by subjects and by items for Experiment 1 and the new pattern of results did not differ from the original. Thus, there did not appear to be a problem with skill group crossover in the middle of the comprehension skill distribution in Experiment 1.
REFERENCES


BIOGRAPHICAL SKETCH

Carol Joy Madden was born in a suburb of Philadelphia on January 30, 1976 to James and Claire Madden. Carol graduated magna cum laude from the University of Notre Dame, receiving her Bachelor of Arts degree in Psychology in May of 1998. She entered the graduate program in Cognitive Psychology at Florida State University in the Fall of 1998 to work with Dr. Rolf A. Zwaan. She received her Masters of Science degree in this program in 2000. Her research interests include: mental representations during language comprehension, concept representations, situation models and perceptual symbols, lexical access of stored representations, context effects in language comprehension, individual differences in working memory and comprehension.