

# **Skilled Readers Make Better Use of Anaphora: A Study of the Repeated-Name Penalty on Text Comprehension**

**Amy Shapiro<sup>1</sup> and Amy Milkes<sup>2</sup>**

---

<sup>1</sup>Psychology Department, University of Massachusetts, North Dartmouth, MA  
<sup>2</sup>Nemours Children's Clinic, Jacksonville, FL

---

USA

[ashapiro@umassd.edu](mailto:ashapiro@umassd.edu)

## Abstract

**Introduction.** The repeated-name penalty refers to the interference experienced by readers when pronouns are replaced with repeated names. The deletion of anaphoric referents has been shown to decrease reading speed (Gordon, Grosz, and Gilliom, 1993), but not comprehension. The present study sought to explore whether the repeated-name penalty would be evident in posttests of reading comprehension. Because good readers make use of textual cues differently than poor readers, we also investigated whether good and poor readers would be differentially affected by the absence of anaphoric reference.

**Method.** Subjects each read two texts, one with pronouns and one with repeated names, and took comprehension posttests.

**Results.** Regression analyses showed that the repeated-name penalty does affect comprehension but only for highly skilled readers, suggesting that more skilled readers were better able to make use of pronouns as cues to local coherence. Consistent with the construction-integration model of text comprehension (Kintsch, 1988), the effect was found for measures of fact acquisition but not deep understanding.

**Discussion.** These results suggest that the ability to make use of anaphora as a cue to create local text coherence is a factor in reading skill. Failure of past studies to detect the effect of repeated names on comprehension may be due to the short length of the test passages used in prior research.

**Keywords:** Anaphora, Construction-Integration Model, Pronouns, Repeated-Name Penalty, Text Coherence, Text Comprehension

## Introduction

The *repeated-name penalty* refers to an increase in reading time experienced by readers when pronouns are replaced with repeated names in sentences (Gordon, Grosz, and Gilliom, 1993). This finding is notable because it provides a window on an important aspect of reading processes. Specifically, it points to readers' use of reference cues as they decode sentences. To date, the repeated-name penalty has only been studied in the context of short, isolated paragraphs and not full texts. Further, the effect has only been shown to affect reading rate and not comprehension or learning measures. The general purpose of the present study was to explore the repeated-name penalty as a factor in text comprehension. Since a great many studies have shown the relevance of reader characteristics to learning, we also explored the interaction between pronoun use and some reader characteristics. Specifically, this study examines the interaction between reading skill, prior knowledge, and the presence of pronouns or repeated names on text recall and comprehension. In order to motivate our choice of reader characteristics as variables in the present experiment, a number of studies are reviewed below. That section is followed by a discussion of the repeated-name penalty.

### Reader and Text Characteristics Affecting Comprehension

Both prior knowledge and reading comprehension skill are highly related to reading and learning from texts. It is well documented in text comprehension research that prior knowledge in a domain facilitates the comprehension process (Recht & Leslie, 1988; Means & Voss, 1985; Spilich, Vesonder, Chiesi, & Voss, 1979; Chiesi, Spilich, & Voss, 1979). In fact, Dochy, Segers, and Buehl (1999) reviewed 183 books, articles, and papers on prior knowledge and found that 91.5% of them reported positive effects of prior knowledge on performance. When readers have knowledge of the subject they are reading about, they are better able to retain and recall information from the text.

Of course, reading comprehension skill also influences how much information is learned and/or retained from reading a text (Laing & Kamhi, 2002; Long & Chong, 2001; Perfetti, 1989; Schmidt, Rozendal, & Greenman, 2002; Voss & Silfies, 1996). Perfetti (1989) argued that reading is a generalized ability that is independent of knowledge. While he suggests that both knowledge and reading ability contribute to the process of learning from text he has shown that, knowledge levels being equal, those individuals who are strong readers will be able to comprehend and recall information from texts better than poor readers.

Another factor that affects text comprehension is a characteristic of the text itself: coherence. Coherence generally refers to the degree to which the relationships between ideas in a text are made clear. Several researchers have demonstrated the effects of increased text coherence on comprehension (Beck, McKeown, Sinatra, & Loxterman, 1991; Britton & Gulgoz, 1991; McNamara, Kintsch, Songer, & Kintsch, 1996; Voss & Silfies, 1996). For example, Britton and Gulgoz (1991) took an unfamiliar text about the Vietnam War, which had causal gaps between ideas, and revised it so that ideas and sentences linked together more explicitly. In other words, they increased the local coherence of the text. They found that college students who read the revised (coherent) version recalled significantly more information in a free recall test and answered significantly more inference questions correctly than those who read the original version.

Based on this research, it seems logical to assume that making a text easier and more coherent would improve comprehension for all readers. However, research has shown that improving text coherence, which makes the text easier to read, is not beneficial to all readers (McNamara et al., 1996; McNamara & Kintsch, 1996; Voss & Silfies, 1996). McNamara et al. (1996) found that prior knowledge interacts with text coherence. The researchers used four versions of a biology text, each of which was high or low on local coherence and high or low on global coherence. Their measures were designed to test both fact retention and deeper learning of their subjects. They found that on the free recall test, prior knowledge was not related to performance. However, on the problem-solving questions, readers with high knowledge performed better with the low coherence text and low knowledge readers performed better with the high coherence text. McNamara and Kintsch (1996) replicated McNamara et al.'s (1996) findings with a history text and an adult sample.

McNamara et al. (1996) ground their work in Kintsch's (1988; 1994) construction-integration model of text comprehension. Kintsch draws the distinction between simply remembering a text and actually learning from it. A textbase representation is sufficient to facilitate recall, and in some cases, summary of a text. However, a situation model allows a reader to have a deeper understanding of the text and actually learn from it. A textbase is formed when the reader creates a mental representation of the material in the text itself. A situation model, however, is formed when the textbase is integrated with prior knowledge, resulting in a deeper, more complex understanding of the topic. McNamara et al. (1996) interpret their results as an indication that high knowledge readers learn better from less coherent text because it forces them to make inferences and actively apply their prior knowledge to

the information they are reading. In other words, it helps them develop a situation model. However, readers with low prior knowledge benefit from high coherence texts. The added text content is helpful because they do not have the knowledge to make inferences about what is not in the text. These readers benefit from the expanded text because it is fully coherent and rich with information, so it allows them to form a solid textbase representation.

Voss and Silfies (1996) explored all three of the text and reader attributes discussed above (reading skill, prior knowledge and text contents/coherence). The researchers developed two sets of fictitious history texts. Each set included an expanded and an unexpanded version. The expanded version was fully developed, with explicit causal relationships between events and ideas. The unexpanded version was not well developed; it simply described the events without explaining how they related to each other or the final outcome of the story. The researchers hypothesized, based on Kintsch's (1988) model of text comprehension, that the development of a textbase representation would be related to reading skill and the development of a situation model would be related to prior knowledge. In addition, they predicted that prior knowledge and reading skill would influence reading comprehension differentially, depending on the text contents. They found exactly that. Prior knowledge was significantly correlated with the learning outcome when the unexpanded text was read and reading skill was not; reading skill was significantly correlated with the learning outcome when the expanded text was read and prior knowledge was not.

Voss and Silfies (1996) explained that when an unexpanded text is read, those readers with prior knowledge can fill in the gaps (make the text more coherent), so they can develop a situation model and fully comprehend the text. High knowledge readers do not learn as much from high coherence texts because all the information is provided for them. Therefore, they do not have a reason to relate the text information to what they already know, or form a situation model. Readers with low prior knowledge do not learn as much from unexpanded texts because inferences are required for understanding. They cannot make inferences because they have no prior knowledge on which to base them. On the other hand, when the text is well developed, prior knowledge is not necessary for comprehension because all of the information is included in the text. In addition, reading skill is a factor in comprehension of expanded texts because they include more information and are harder to read than less developed texts.

In sum, there is a good amount of evidence showing that readers' skill and prior knowledge are related to comprehension, as is the text's coherence. In light of these results, it is important to study text comprehension within the context of these variables.

### The Repeated-Name Penalty

Gordon et al. (1993) conducted a series of self-paced reading experiments in which they gave students individual blocks of several sentences, with the presence and sentence position of pronouns manipulated, and measured the reading time and comprehension of each sentence set. They found that, under certain conditions, the use of repeated names resulted in slower reading times, as compared to matched sentences with pronouns. Gordon et al. (1993) refer to this effect as the repeated-name penalty. Interestingly, the repeated-name penalty occurred only when the name was in a prominent position, such as the subject of the sentence, and not when the repeated name was in a less prominent position, such as the object of the sentence (Gordon et al., 1993; Yang et al., 1999). Additionally, the repeated-name penalty is dependent on the structural relation between sentences. When a sentence continued a critical referent from the preceding sentence, the repeated-name penalty occurred. However, when the position of the critical referent was shifted between sentences, the repeated-name penalty did not occur (Gordon et al., 1993).

Of importance to our purposes here, Gordon et al. (1993) found no effect of pronoun manipulation on comprehension, only on reading time. Their comprehension measure, however, consisted of single true/false questions for sets of 3-5 sentences.<sup>1</sup> With this small amount of information to decode and remember, we suspect that the nonsignificant effect of pronoun use on test performance in that study may be due to ceiling effects. Indeed, the decrease in reading speed associated with repeated names does suggest that pronouns facilitate comprehension.

The type of coherence that pronouns provide seems to be inherently different from the text coherence discussed earlier (Britton & Gulgoz, 1991; McNamara et al., 1996). The coherence manipulated in most studies involves *clauses or phrases* added to explain relationships. We refer to the coherence created by anaphora as referential coherence because anaphora increases coherence by signaling relationships between ideas, rather than actually provid-

---

<sup>1</sup> The design of that study was aimed at testing predictions of Centering Theory. The present investigation was not concerned with Centering Theory, but with exploring the effects of coherence created by anaphoric reference on learning from text.

ing information that explains relationships. In other words, pronouns help readers bridge sentences by indicating a common subject, thus connecting them conceptually. Indeed, the increased reading time in the absence of pronouns observed by Gordon et al. (1993) suggests that comprehending texts without pronouns necessitates some additional processing on the part of readers.

Theories of learning from text serve to explain why repeated names in a text would interfere with learning. Van Dijk (1980) proposed that learned information is incorporated or stored in macrostructures. His theory suggests that information units are linked together to form these macrostructures which serve to both organize and reduce complex information. Van Dijk argues that macrostructures allow us to form larger "chunks" of information that "have their proper meaning and function" (p. 14). Further, he contends that assigning a macrostructural unit to a series of independent facts (which he calls "units at the microstructural level") defines the thematic relationship between them. Since pronouns signal readers that the subject of a sentence is the same as a prior sentence, they serve to help readers to connect information between sentences. In van Dijk's terminology, they signal readers to connect information units within a common macrostructure.

Kintsch (1988; 1994) extended van Dijk's theory with his construction integration model. As discussed earlier, Kintsch proposes that such macrostructures are formed in memory when information contained in a text is stored. These structures, which he calls textbases, mirror the text's organization and are constructed from its semantic content. Within the construction-integration model, then, pronouns may also be seen as a cue to readers that ideas between sentences should be joined within a common structure. The construction-integration model also proposes that another kind of structure, called a situation model, may also be formed during reading. Situation models contain the information from the textbase and additional information from permanent memory. The situation model, then, may be thought of as the storehouse for our deeper understanding of written material. A more involved form of processing is required to create a situation model, as the new information must be integrated with prior experiences and knowledge.

The construction-integration model predicts that the repeated-name penalty would occur for textbase construction (e.g., factual information from the text itself). It does not, however, predict differences in situation model measures (i.e., deep understanding) as a result of repeated names versus pronoun texts because anaphoric reference does not help readers con-

nect the text with prior knowledge. Pronouns only cue readers to create local coherence within a text, thus facilitating textbase construction.

To test these predictions, the present study presented subjects with texts containing either repeated names or pronouns. Posttests included factual knowledge (short answer) and deeper understanding (essay) measures. There were three main hypotheses of this study. The first was that the repeated-name penalty would occur, shown by significantly better posttest performance in the pronoun condition than in the repeated name condition. Our second hypothesis was that there would be an interaction between reading ability and pronoun use. Good readers make use of textual cues differently than poor readers (Oakhill & Yuill, 1986). Hence, good and poor readers should be affected differentially by referential coherence. Consistent with the construction-integration model of text comprehension (Kintsch, 1988), both of these hypotheses are made for the factual knowledge posttest but not the essay posttest. That is, these variables are expected to affect the construction of textbases, but not situation models. Our third hypothesis was that prior knowledge would not interact with pronoun use, as a text low in referential coherence may not be “remedied” by prior knowledge; we view referential coherence as a reading issue alone and not a content issue. We did predict, though, that prior knowledge would be a significant factor by itself in the quality of subjects’ essays. Because many prior studies have shown that prior knowledge predicts subjects’ deeper understanding of texts (Dochy, Segers, and Buehl, 1999), we expected to find a significant contribution of prior knowledge to the quality of subjects’ essays.

## **Method**

### *Subjects*

Participants were fifty-seven undergraduate students from the University of Massachusetts, participating to fulfill an introductory psychology course requirement. There were roughly the same number of men and women.

### *Materials*

The materials used in this study were based on those created by Voss and Silfies (1996). They developed two texts, each of which described two fictional countries. We used these because the two texts had already been equated by Voss and Silfies (1996) and we were



interested in using a within-subjects design. Furthermore, the posttests Voss and Silfies developed for these texts had already been used successfully in text processing research. One text is called the “Anchad” text and one is called the “Padria” text, each title being the name of one of the countries discussed in the text. The Anchad pronoun text was 742 words in length and the repeated names text was 802 words in length. The Padria texts were 753 words and 778 words in length, respectively. They are referred to as “full-length” texts because they are complete text passages as compared to the isolated sentence sets studied previously. In each text, a series of events are described that led to the outbreak of a war.

All participants received both an Anchad and a Padria text, one of which made use of pronouns and one that did not. The order was counterbalanced so that half of the participants received an Anchad text first and half received a Padria text first. Also, it was counterbalanced so that all four possible combinations of versions and orders of the two texts were presented to subjects.

To test the repeated-name penalty, the Anchad and Padria texts were slightly modified so that the same sentence subject appeared at the beginning of several consecutive sentences. This modification did not affect the text contents, it only altered the order of some words already in the text. One version of each text contained subject names repeated in consecutive sentences and another version used a subject name in the first sentence and a subject pronoun in the following sentence, as the following example illustrates:

Initial sentence:

*John Lerner* is the leader of the resistance movement.

Sentence 2 in the subject repeated names condition:

*John Lerner* is a peasant whose parents were employed as crop laborers.

Sentence 2 in the pronoun condition:

*He* is a peasant whose parents were employed as crop laborers.

In both the Anchad and Padria texts there were 14 sets of repeated subjects in consecutive sentences.

## *Procedure*

*Pretesting.* Participants first completed the reading comprehension portion of the Nelson-Denny reading test. The Nelson-Denny is a validated test of reading comprehension and has been used in hundreds of published articles related to reading research. It has been shown to be a good predictor of a variety of measures of academic success (Hawes, 1982; Wood, 1982; Zimmer, Glover, Ronning, & Petersen, 1979). This test is timed, but subjects were allowed to work at their own pace for the remainder of the study. Participants then completed a demographics questionnaire and Voss and Silfies (1996) 20-question test designed to measure their world history knowledge.

*Reading and Posttesting.* Participants were told to carefully read the first text and that they would be answering questions about it after they were finished. When they were done reading, they received the short answer questions. Participants could not refer to the text to answer the questions. They were then asked to write the essay. They were allowed to refer to the text to answer this question. When finished, subjects were given the second text and the procedure was repeated. The total time needed to complete the study ranged from one hour to one hour and 45 minutes.

## *Statistical Analysis*

*Dependent Measures.* The short-answer tests for each text were developed by Voss and Silfies' (1996). The answers to each of these questions were found in the text itself. As such, they are measures of subjects' knowledge of factual information retained from the text itself; a textbase measure.

Subjects were also asked to answer an essay question that asked about possible underlying causes of the conflict described in the text. The same question was asked for both the Padria and Anchad texts. These questions required subjects to think beyond the factual information presented in the text and engage in analysis of the causes of the conflict. Answers to this question reflect subjects' deeper understanding of the text and thus were used as measures of situation model development.

*Data Scoring.* Essays were scored on a 6-point scale designed to reflect the quality of the contents of the essay. The rating scale ranged from 0 to 5, see Table 1 for the actual rating criteria. The experimenters independently scored sets of both Anchad and Padria essays. To avoid reader fatigue, no more than five essays were scored in one sitting. We scored a num-

ber of Anchad and Padria essays as practice and met to discuss our scores, which helped hone the rating scale and make the scoring more accurate. As a final test of the reliability of our rating method, we then each scored 10 Padria and 10 Anchad essays independently. The reliability of our ratings was  $r = .795$  for the Anchad essays and  $r = .742$  for the Padria essays. The short answer questions each had single words or phrases as answers. Any answer that did not appear exactly as it was written on the answer sheet was scored as incorrect. The only exception was the use of singular versus plural nouns or verb tense.

**Table 1: The Guidelines Used to Standardize Essay Scoring.**

SCORE	CRITERIA*
0	No response to the question; or the content is either all wrong and/or all irrelevant.
1	Addresses the question but speaks only in generalities. The response may mention something from the text but there is only a small amount that is accurate and/or relevant.
2	Response to the question is somewhat relevant. Specifically, there is a mixture of accurate and inaccurate/relevant and irrelevant points made. It can neither be easily classified as mostly accurate/relevant or inaccurate/irrelevant.
3	Response to question is mostly relevant and accurate (there may be one sentence or so that is problematic). The response may be mostly drawn directly from the text itself but has definite bearing on the question at hand. While relevant points are included in the answer, little or not explicit attempt is made to explain how they contributed to the conflict. The student may offer a thesis/hypothesis for the cause of the conflict but this thesis is either (1) not supported by facts included in the essay or (2) wrong.
4	Response to question is mostly relevant and accurate (there may be one sentence or so that is problematic). The response may be mostly drawn directly from the text itself but has definite bearing on the question at hand. Further, the student is starting to explain the relationship between multiple points as a way of addressing the root of the problem. However, the explanation is limited to the text or "surface" explanations.
5	Response to the question is relevant and thoughtful with a good amount of substance. Some insight is present in the answer and multiple causes are cited. The student has engaged in some analysis of the causes of the events cited, going beyond the text itself.

\* Ignore all references to the military action itself and anything that occurred after the military action. Count these references neither in favor nor against the response. In the Padria text, ignore Padria being bombed and subsequent events. In the Anchad text, ignore buildup of troops on the borders and subsequent events.

## Results

### *Correlation Analyses*

As a validity check of our measures, correlations were computed between the knowledge and comprehension variables, and the dependent variables. Since both of these measures have repeatedly shown to predict learning outcomes, the same should generally be true here. Table 2 presents these results. As we predicted, history knowledge score was significantly correlated with performance on both short answer and essay posttests in both the pronoun and repeated names conditions. Our prediction that reading comprehension scores would significantly correlate with posttest performance was also confirmed, as reading comprehension was significantly correlated with performance on all measures except the essay measure for the pronoun condition.

**Table 2: Correlations of Comprehension-Set and Knowledge-Set Variables with Dependent Variables.**

<i>Variables</i>	<b>PRONOUNS</b>		<b>REPEATED NAMES</b>	
	<i>Short Answer</i>	<i>Essay</i>	<i>Short Answer</i>	<i>Essay</i>
<b>Knowledge Set</b>				
History Knowledge	.41**	.44**	.53**	.43**
Interest in history	.25	.01	.32*	.25
Number of history classes	.14	.08	.08	.14
Number of political science classes	.01	-.11	-.10	-.05
Interest in current events	.07	-.11	.03	.15
<b>Comprehension Set</b>				
Reading Comprehension	.44**	.26	.36**	.40**
Reading Rate	.18	.29*	.09	.00
Grade Point Average	.04	.27*	.01	.12
SAT Verbal Score	.34*	.27	.44**	.38**

Note. \* $p < .05$ . \*\* $p < .01$

### *Simple Comparisons*

The prediction of a main effect of pronoun condition was not supported by the data. The mean short answer scores for the repeated names and pronoun conditions were 4.79 and 4.78, respectively. The mean essay scores for those conditions were 2.79 and 2.84, respectively. The differences were nonsignificant in both cases,  $t(57) < 1$ . Contrary to our prediction, then, we replicated Gordon et al.'s (1993) original finding. They reported only an increase in reading time associated with repeated names and no main effect of pronoun manipulation on their comprehension measure. The next section explores our second hypothesis that the repeated-name penalty is present when considered within the context of reading skill. It may be that the pronoun manipulation only affects comprehension when interacting with reading skill.

### *Regression Analyses*

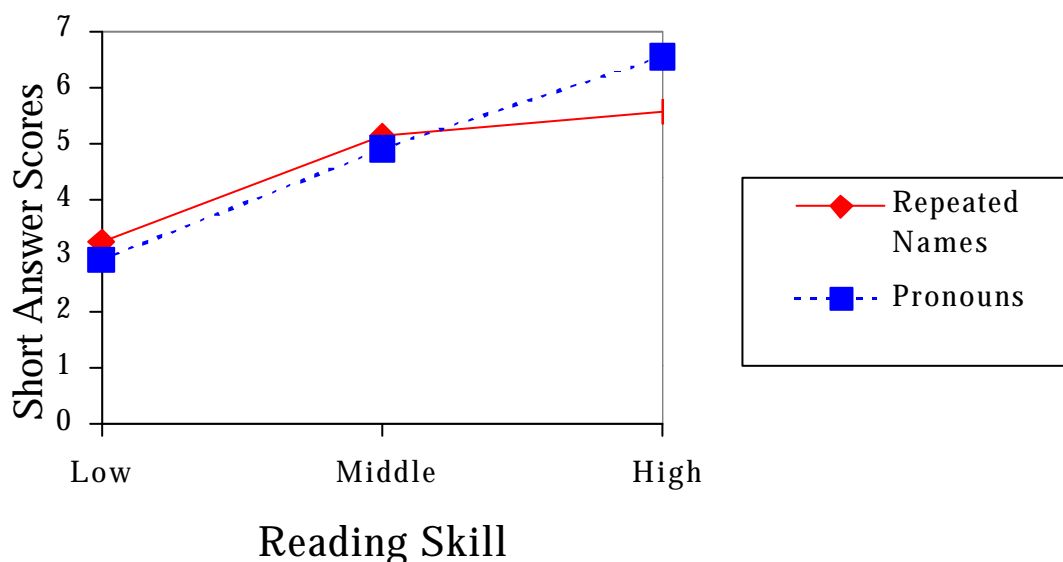
Regression analyses were conducted to explore the contribution of prior knowledge and reading skill to text recall and comprehension, for both the repeated name and pronoun texts. Not all of the knowledge set and comprehension set variables could be used in the analyses because the number of participants available for the study would not support this large number of independent variables. Therefore, reading comprehension pretest score and history knowledge pretest score were used as the measures of reading skill and prior knowledge in the multiple regression analyses. These were chosen as they are the most direct measures of reading skill and history knowledge included in this study.

Two sets of multiple regressions with repeated measures were conducted. Each analysis included two continuous independent groups factors (reading score and history score) and one repeated measures factor (pronouns vs. repeated names). One analysis was conducted with short answer scores (number of questions correct) as the dependent variable and one with essay scores as the dependent variable. To do a repeated measures multiple regression it is necessary to do two separate multiple regressions, one in which the dependent variable is averaged across the repeated measures conditions (pronouns and repeated names) and a second in which the repeated measure is entered as a dummy variable (Cohen & Cohen, 1983). The data were satisfactorily screened for the regression assumptions, with the exception of one outlier in the short answer data (subject 7), which was removed from the short answer analysis. The regression analysis that used short answer scores as the dependent variable will be discussed first, followed by the regression that used essay scores as the dependent variable.

*Short Answer Questions.* The reading scores and history scores together were significant,  $R^2 = .48$ ,  $F(2, 55) = 25.36$ ,  $p < .01$ . Considered separately, history score was significant ( $sr = .46$ ,  $p < .01$ ) and reading score was also significant ( $sr = .29$ ,  $p < .01$ ). The repeated measures factor was not significant,  $R^2 = .00$ ,  $F(1, 55) = .00$ . The two-way interactions, considered as a whole, were significant ( $R^2 = .09$ ,  $F(3, 50) = 2.84$ ,  $p < .05$ ). Considered separately, the reading by repeated measures interaction was significant ( $sr = .35$ ,  $t(54) = 2.76$ ,  $p < .01$ ), but the history knowledge by repeated measures interaction was not significant. Finally, the three-way interaction was not significant,  $R^2 = .03$ ,  $F(1, 49) = 1.50$ .

To determine the direction of the effects in the reading by repeated measures interaction, the subjects were first divided into three groups based on their reading comprehension score. The mean of the reading scores was 22.3, with a standard deviation of 5.5. The low group consisted of those who scored one standard deviation below the mean, the high group scored one standard deviation above the mean and the middle group scored somewhere in between those two scores. Then, the averages of the short answer scores in the pronoun condition and in the repeated names condition were plotted on a graph for each reading group (low, middle, and high). This graph is presented as Figure 1.

**Figure 1: Short Answer Performance of Low, Middle, and High Skill Readers in the Repeated Name and Pronoun Conditions.**



*Essay Questions.* The results from the multiple regression showed that reading score and history score, entered together, was a significant factor ( $R^2 = .32$ ,  $F(2, 55) = 13.10$ ,  $p < .01$ ). Considered separately, the history score was significant ( $r = .40$ ,  $p < .01$ ), but the reading score was not. The repeated measures factor was not significant,  $R^2 = .001$ ,  $F(1, 55) = .07$ . The two-way interactions, considered as a whole, and the three-way interaction were also not significant,  $R^2 = .005$ ,  $F(3, 50) = .24$  and  $R^2 = .008$ ,  $F(1, 49) = 1.154$ , respectively.

## Discussion

Our first hypothesis was that pronoun use would be a significant, independent factor for the short-answer but not the essay regressions, with scores in the pronoun condition being significantly higher than their counterparts in the repeated name condition. However, the repeated measures factor was not significant by itself in either analysis. This finding shows that, contrary to our initial thinking, repeated names do not generally “penalize” readers with respect to comprehension when considered as an isolated variable. This finding, then, is a replication of Gordon et al.’s (1993) recall posttest results and disproves our first hypothesis.

Considered within the context of reading skill as a variable, however, the repeated-name penalty was evident, supporting our second hypothesis. In the pronoun condition, participants with higher reading ability performed significantly better than the other participants on the short-answer questions. In the repeated name condition, however, participants with high reading skill performed more comparably with their less skilled counterparts. As predicted, the interaction between reading skill and pronoun use only occurred in the short-answer posttest.

The significant interaction on the short-answer test reveals that referential coherence is important for highly skilled readers’ textbase construction. This result is similar to that of Voss and Silfies (1996), who found that reading skill predicted performance on a factual posttest after reading a highly coherent text but not a less coherent text. Like the present results, their coherence manipulation affected tests of textbase development and not the situation model. The parallel between Voss and Silfies’ results and those presented here suggests that what we have referred to as referential coherence, obtained through pronoun use, affects comprehension in a way that is similar to other manipulations of text coherence. Specifically, referential coherence affects readers in a way that is similar to the addition of phrases or clauses that serve to explain relationships between ideas.

At the beginning of this investigation, we proposed that Gordon et al. (1993) did not observe an effect of their pronoun manipulation on comprehension because of the short paragraphs they used as stimuli. The lack of a main effect of pronoun manipulation on comprehension measures in the present study suggests that this idea was incorrect. Instead, the interaction with reading skill suggests that, for comprehension measures, the repeated-name penalty exists only for high skilled readers. Since Gordon et al. did not explore their comprehension measure within the context of reading skill, they were unable to detect the repeated-name penalty for comprehension. While Gordon et al.'s work suggests that readers generally take more time to read a repeated names text, we have extended their findings to show that only higher skilled readers suffer a relative decrement in comprehension.

One likely explanation for this interaction between reading skill and pronoun use is that skilled readers make better use of cues like pronouns to aid them in comprehension. As a consequence, when highly skilled readers read a pronoun text, they benefit from their effective use of the pronoun cues. However, when they read a repeated name text, that important tool is missing and it negatively affects their decoding of the text. Individuals with low and moderate reading skill may have more difficulty using pronouns as cues for comprehension, so they perform the same whether they are reading a pronoun or repeated names text. This explanation is supported by Oakhill and Yuill (1986), who studied 7- and 8-year-old English speakers' ability to understand pronouns. They gave a series of 2-sentence clauses to the children and asked them to identify the pronoun referents. They found higher error rates on the part of poor readers and suggest that difficulty in pronoun interpretation and comprehension may contribute to the difficulty experienced by poor readers.

As predicted, there was no independent contribution of pronoun use on essay performance. In addition, neither subject variable interacted with the pronoun manipulation on that measure. These results indicate that the deeper understanding measured by the essays was unaffected by referential coherence. Instead, consistent with our third hypothesis, only prior knowledge by itself significantly contributed to essay performance. This pattern of results is explained by the construction-integration model (Kintsch, 1988). Pronouns serve to help readers create local coherence and, thus, strong, accurate textbase representations. Pronouns do not help readers integrate information with prior knowledge, the process through which a situation model (deeper understanding) is constructed. Since the essay questions were designed to assess deeper understanding, the nonsignificant effect of pronoun manipulation was predicted for that measure. On the other hand, it makes sense that greater prior knowledge



was associated with better performance on the essay test. The more prior knowledge a subject has, the more he or she will be able to integrate with the text content and enhance understanding. Indeed, the essay results parallel those reported in a number of prior reports (see Dochy et al., 1999, for a review). While support for hypothesis 3 does not serve to further understanding of pronoun use in text comprehension, it does speak to the validity of our prior knowledge and deep comprehension measures.

Until now the repeated-name penalty, as Gordon et al. (1993) defined it, has not been studied in a full-length, naturalistic text passage, nor in combination with reader characteristics. We found no evidence that the length of a text is relevant to the presence of a repeated-name penalty in comprehension. We did find, though, that the repeated-name penalty does indeed extend to measures of text retention and learning, but only for better readers. Indeed, the differential effect of this type of coherence on readers of differing ability may be a window on an important factor in reading skill. What makes good readers good? The present results suggest that one factor is the ability to make effective use of referential cues to create local coherence. Others have shown that local coherence in a text significantly affects comprehension (Britton & Gulgoz, 1991). The present results suggest that skilled readers create this type of coherence, in part, with the aid of such referential markers.

## References

- Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly, 27*, 251-276.
- Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology, 83*, 329-345.
- Chiesi, H. L., Spilich, G. J., & Voss, J. F. (1979). Acquisition of domain-related information in relation to high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior, 18*, 257-273.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69*, 145-186.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*, 311-347.
- Hawes, K. (1982). Correlations of scores on the Nelson-Denny Reading Test, Form E, and measures of academic aptitude, academic achievement, readers' attitude, and students' self-concept. *Psychological-Reports, 51*, 229-230.
- Kintsch, W. (1988). The use of knowledge in discourse processing: A construction integration model. *Psychological Review, 95*, 163-182.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*, 294-303.
- Laing, S. P., & Kamhi, A. G. (2002). The use of think-aloud protocols to compare inferencing abilities in average and below-average readers. *Journal of Learning Disabilities, 35*(5), 436-447.

- Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1424-1429.
- McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse Processes*, 22, 247-288.
- McNamara, D. S., Kintsch, E., Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1-43.
- Means, M. L., & Voss, J. F. (1985). Star wars: A developmental study of expert and novice knowledge structures. *Journal of Memory and Language*, 24, 746-757.
- Oakhill, J., & Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. *Language and Speech*, 29, 25-37.
- Perfetti, C. A. (1989). There are generalized abilities and one of them is reading. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 307-336). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Recht, D. R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, 80, 16-20.
- Schmidt, R. J., Rozendal, M. S., Greenman, G. G. (2002). Reading instruction in the inclusion classroom: Research-based practices. *Remedial and Special Education*, 23, 130-140.
- Shapiro, A.M. (2002). *The Relationship between Text Structure and Reader Characteristics: A Search for Predictors of Successful Learning*. Paper discussion at the Annual Meeting of the American Educational Research Association, New Orleans, April.
- Spilich, G. J., Vesonder, G. T., Chiesi, H. L., & Voss, J. F. (1979). Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 18, 275-290.
- van Dijk, T. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

- Voss, J. F., & Silfies, L. N. (1996). Learning from history text: The interaction of knowledge and comprehension skill with text structure. *Cognition and Instruction, 14*, 45-68.
- Wood, P. (1982). The Nelson-Denny Reading Test as a predictor of college freshmen grades. *Educational-and-Psychological Measurement, 42*, 575-583.
- Yang, C. L., Gordon, P. C., Hendrick, R., & Wu, J. T. (1999). Comprehension of referring expressions in chinese. *Language and Cognitive Processes, 14*, 715-743.
- Zimmer, J., Glover, J., Ronning, R., & Petersen, C. (1979). On the utility of the Nelson-Denny Reading Test as a covariate in research on prose-processing. *Perceptual-and-Motor-Skills, 48*, 641-642.