NOTICE: THIS MATERIAL MAY BE
PROTECTED BY COPYRIGHT LAW
(TITLE 17, U.S. CODE)

Context clues are unreliable predictors of word meanings

ELINORE KRESS SCHATZ
Greater Miami Hebrew Academy
R. SCOTT BALDWIN
University of Miami

THREE STUDIES were conducted to determine the extent to which context helps students infer the meanings of unknown words. In Experiment 1, students in Grades 10 and 11 were randomly assigned to either a context or a no-context condition. The no-context group read low-frequency words in isolation. The context group read the same words embedded in passages taken from novels. Experiment 2 was a repeated-measures study in which 39 students in Grade 11 read sets of words in isolation and also in passages taken from four different content areas. Experiment 3 was a systematic replication of Experiment 1 in which subjects were required to write definitions for the low-frequency words instead of choosing the definitions in a multiple-choice format. In none of the three experiments was there any statistically significant effect due to the context: $t(99) = .552, p > .10$; $F(1, 38) = .227, p > .10$; and $t(83) = -.29, p > .10$, respectively. The conclusion is drawn that instructional strategies that prioritize context clues should be reexamined.

On ne peut se fier aux indices contextuels pour prédire le sens des mots

ON A MENÉ trois enquêtes en vue de déterminer jusqu'à quel point le contexte pouvait aider les élèves à attribuer une signification aux mots inconnus. Dans la première expérience, des élèves de dixième et onzième années ont été soumis au hasard à l'une de ces situations: avec contexte, sans contexte. Le groupe dans la situation sans contexte a lu une liste de mots isolés peu fréquents. L'autre groupe, avec contexte, a pu lire les mêmes mots, présentés dans des extraits de romans. La deuxième expérience consistait en une évaluation répétitive et s'effectuait sur un groupe de 39 élèves de onzième année où ceux-ci devaient lire une série de mots présentés isolément puis dans des passages portant sur quatre sujets différents. La troisième expérience était une reprise systématique de la première dans laquelle les sujets devaient donner une définition aux mots peu fréquents plutôt que d'en trouver une à partir d'un choix multiple. On n'a remarqué aucun effet considérable statistiquement, attribuable au contexte dans l'une ou l'autre des trois expériences. On en tire donc la conclusion que les stratégies d'enseignement privilégiant les indices contextuels méritent d'être examinées à nouveau.

Las claves de contexto son predictores poco confiables del significado de las palabras

SE LLEVARON a cabo tres estudios para determinar hasta que punto el contexto ayuda a los estudiantes a inferir significados de palabras desconocidas. En el experimento 1, estudiantes de los grados 10 y 11 fueron asignados al azar, ya fuera, a una condición de contexto o a otra de no contexto. El grupo de no contexto leyó palabras aisladas de baja frecuencia. El grupo de contexto leyó las mismas palabras embebidas en pasajes tomados de novelas. El experimento 2 fue un estudio de medidas repetidas en el que 39 estudiantes del grado 11 leyeron conjuntos de palabras, aisladas y también en pasajes tomados de cuatro áreas de contenido diferentes. El experimento 3 fue una réplica sistemática del experimento 1 en el que se pidió a los sujetos que escribieran definiciones para las palabras de baja frecuencia en vez de que escogieran las definiciones en un formato de opción multiple. En ninguno de los 3 experimentos se encontró ningún efecto estadísticamente significativo debido al contexto. La conclusión obtenida es que las estrategias de instrucción que dan prioridad a las claves de contexto deben ser reexaminadas.

Zusammenhangs-Anhaltspunkte sind unzuverläßig als Voraussagen von Wortbedeutungen

ES WURDEN drei Studien durchgeführt, um festzulegen, inwieweit Zusammenhänge den Schülern beim Verständnis unbekannter Worte behilflich sind. In Experiment 1 wurden Schüler im 10. und 11. Schuljahr wahllos in Zusammenhangs- und Nicht-Zusammenhangs-Lagen versetzt. Die Nicht-Zusammenhangs-Gruppe las nicht häufig vorkommende Worte vereinzelt. Die Zusammenhangs-Gruppe las dieselben Worte als zugehörig zu Abschnitten in Romanen. In Experiment 2 wurde eine Wiederholungsmaßnahmen-Studie vorgenommen, in welcher 39 Schüler im 11. Schuljahr Wortzusammenstellungen einzeln und auch als Abschnitte aus verschiedenartigen Bereichen lasen. Experiment 3 war eine systematische Nachbildung, in welcher die Versuchspersonen angehalten wurden, Erklärungen zu finden für selten vorkommende Worte, anstatt die Erklärungen in Art eines Vielwahl-Formates auszuwählen. In keinem der drei Experimente beobachtete man einen statistisch bedeutungsvollen Einfluß, der auf den Zusammenhang zurückzuführen wäre. Daraus ist die Schlußfolgerung zu siehen, daß Unterrichtungs-Methoden, die sich auf Zusammenhangs-Anhaltspunkte berufen, sorgfältig untersucht werden sollten.

Based on an inspection of instructional materials and reading methods texts, the positive value of context clues appears to be an almost unquestioned assumption in the field of reading. We will argue in this manuscript that the utility of traditional context clues, such as synonym clues and comparison and contrast clues, has been vastly overestimated. Moreover, the results of our research demonstrate that context does not usually provide clues to the meanings of low-frequency words, and that context clues actually inhibit the correct prediction of word meanings just as often as they facilitate them.

The idea that the use of context clues is an effective strategy for inferring word meanings is a venerable one (e.g., Barnum, 1906; Huey, 1908). The current pedagogical status of instruction in use of context clues as a method of vocabulary acquisition and as a word-recognition strategy is one of profound authority (e.g., Durr, Le Pere, Pescosolido, Bean, & Glaser, 1983; Johnson & Pearson, 1984). Moreover, publishers and authors of methods texts also view the use of context clues as an effective strategy for inferring *word meanings* (e.g., Burns, Roe, & Ross, 1984; Niles, Deffenbaugh, Hynes-Berry, Lamberg, & Savage, 1983; Wiener & Bazerman, 1985). The standard philosophy regarding the priority of context clues in any strategy for identifying semantically unfamiliar words is clearly stated by Nist (1985): "Try context first, structure second, sounding it out third, and, if all else fails, then, and only then, consult a dictionary" (p. 86). Today, almost eight decades after the publication of Huey's classic text (1908), publishers, teachers, and the authors of reading methods textbooks have essentially the same perception of context as an efficient mechanism for inferring word meanings.

Research on Context Clues

The literature is replete with empirical studies that demonstrate the facilitating effects of context on *word recognition*—that is, the identification of a word that is unfamiliar in print but is in the reader's oral vocabulary (Stanovich & West, 1981; West, Stanovich, Feeman, & Cunningham, 1983-1984). Most of the studies investigating context facilitation at the word-recognition level seem to deal with the following topics: words in context versus words in isolation (Biemiller, 1977-1978; Goodman, 1965; Samuels, 1970; Weaver, 1977), good and poor readers' use of context (Allington & Fleming, 1978; Krieger, 1981; Perfetti, Goldman, & Hogaboam, 1979; Schvaneveldt, Ackerman, & Semlear, 1977), and the interaction of context with word frequency (Allington, 1980; Pearson & Studt, 1975). Three key generalizations are derived from such studies: (a) Readers recognize words faster in context than in isolation; (b) poor readers are not only able to use context, but benefit at least as much as good readers from context; and (c) context significantly helps readers pronounce high-frequency words, but

has only a minimal effect in helping readers pronounce more difficult, less frequent words.

Studies of context clue strategies in the area of *vocabulary instruction* (Crist & Petrone, 1977; Gipe, 1978-1979; Madison, Carroll, & Drum, 1982) typically attempt to show the effects on vocabulary acquisition of presenting semantically unknown words in specifically selected defining contexts, as in this example:

> The judge said the lawyer gave a *cogent* argument because she presented her facts in a convincing way.

It appears from these studies that the more context clues are provided in the form of direct definitions, synonyms, or precise descriptions, the better the students will be able to learn the meanings of unknown words (Carnine, Kameenui, & Coyle, 1983-1984; Carroll & Drum, 1983). In addition, if context is reinforced by teacher-directed instruction and student repetition of definitions, then context effects operate maximally (Jenkins, Pany, & Schreck, 1978; Pany & Jenkins, 1978).

The value of context clues as part of a *comprehension strategy* also is supported by the majority of studies on this topic (Carroll & Drum, 1983; Crist & Petrone, 1977; Duffelmeyer, 1984; Madison, Carroll, & Drum, 1982). Readers frequently encounter difficult, unfamiliar words, such as *intrepid* and *inveterate,* during generic reading. According to the current research literature, context clues should help readers to infer the meanings of such words and thereby facilitate comprehension without the need for readers to interrupt the reading act with diversions to glossaries, dictionaries, or other external sources of information.

Most of the studies through the years supporting the facilitating effects of context on word comprehension use high-frequency words in a cloze-type task (Carroll & Drum, 1982; Quealy, 1969; Seibert, 1945). In these studies, words are surrounded by highly constrained contexts which considerably reduce the number of possible word choices. The following sentences are typical of the context conditions in these studies:

> He drank a _____ of coffee.
>
> As _____ as ice.
>
> Bring him back dead or _____.
>
> A recent study by the AMA has shown that child abuse is likely to be a more frequent cause of death than such diseases as leukemia, cystic fibrosis, and _____ dystrophy.

These studies do show that readers can infer the meanings of words. However, the use of high-frequency words in a cloze task means that the words are unknown only in the sense that they are missing from the passage, and not in the sense that the meaning of the deleted word is unknown to the reader. The problem is that the results of studies using high-frequency words have been extrapolated to situations where low-frequency words occur. The fact that readers can guess the meaning of a missing high-frequency word in sentences such as (1) does not mean that they can guess the meaning of a low-frequency word which is semantically unfamiliar, as in sentence (2).

(1) I like ham and _____ for breakfast.

(2) Mike is a *scurrilous* person, and I have never liked him.

Not every study supports use of context clues as a sound strategy for identifying semantically unfamiliar words. Looby (1939) studied the question of children's comprehension of word and phrase meanings through a series of objective and essay tests. She also used the individual interview for an affective evaluation. She found that although context was widely used to determine the meanings of words and phrases, the responses on the tests indicated a full range of interpretations—from adequate understanding to gross misunderstanding. Most crucial was the finding that the subjects were very positive in their own minds that their understandings of word meanings were correct; however, they were as certain they were correct when these understandings turned out to be wrong as when they were right. This effect may be the general case when subjects are exposed to low-frequency words in naturally occurring prose, instead of in contrived contexts that are largely redundant with the word in question.

Word-recognition and vocabulary studies have shown context to have facilitating effects. However, with respect to the identification of meanings of low-frequency words, the research literature is flawed because the studies have evaluated the effects of context with high- instead of low-frequency words, or they have used contrived passages instead of naturally occurring prose. The purpose of this study was to discover whether context clues help readers to identify the meanings of low-frequency words in naturally occurring prose.

EXPERIMENT 1

Method

Subjects

The sample consisted of 101 students in Grades 10 and 11. The 53 10th-grade and 48 11th-grade students were enrolled in a private school in Fort Lauderdale, Florida, whose students were mainly middle- to upper-middle-class and Caucasian. Their mean verbal percentile scores on the Preliminary Scholastic Aptitude Test in October were 45.88 for the 10th-grade class and 65.87 for the 11th-grade class.

Materials

Two 25-item tests, a words-in-context test and a words-in-isolation test, were designed to assess the extent to which context helps students infer the meanings of unknown words. The test items were field-tested on a group of college-bound 10th-grade students in order to determine the extent of the subjects' prior knowledge of the target words. The effects of context would be obscured if students already knew the definitions of most of these words. Therefore, any test item that was defined correctly by 65% or more of the students was discarded.

The words-in-context test consisted of 25 passages selected from 10 novels from the school's reading lists for 10th- and 11th-grade students (e.g., *Grapes of Wrath, Cry, the Beloved Country, The Scarlet Letter,* and *The Pearl*). Sample passages were selected from each book by choosing every second paragraph of every third page. The paragraph thus located was selected if it contained one or more low-frequency words. A low-frequency word was defined as a word which appears four times or less in a million running words, as reported in the Kucera and Francis (1967) word frequency list. The following are examples of low-frequency words which fit this criterion and therefore were among the test items:

cozened	recondite
imperious	salient
ignominiously	cogently
perambulating	inexorable

If a low-frequency word was not present in a paragraph or if a paragraph contained more than three low-frequency words, that paragraph was eliminated, and the next paragraph was examined.

This procedure generated 256 paragraphs. From this item pool, 25 paragraphs were selected using a table of random numbers. The following is an example of a paragraph used in this study:

> He thanks the simple woman, and tells her to go well. He stands for a moment, then turns swiftly and goes to his room. He takes out an envelope from a drawer, and takes paper money from it. He looks at it *ruefully,* and then with decision puts it into his pocket, with decision takes down his hat. Then dressed, with indecision looks out of the window to the house of Mrs. Lithebe, and shakes his head. But he is too late, for as he opens his door, Kumalo stands before him. (Paton, *Cry, the Beloved Country,* p. 81).

The first low-frequency word to occur in the paragraph became the target word and was underlined. After the target word was identified, a three-sentence test passage was constructed from it using the following procedures: If the target word occurred in the first sentence of the paragraph, then that sentence became the first one in the passage, followed by the next two sentences as they appeared in the paragraph. If the low-frequency target word appeared at the end of the paragraph, then the two sentences preceding it were used to form a

three-sentence passage. If the target word appeared anywhere else in the paragraph, then the sentence that preceded the one containing the low-frequency word, the sentence containing the low-frequency word, and the sentence that followed the target word constituted the three-sentence passage. In four cases, a paragraph comprised two lengthy sentences, one of which contained the target low-frequency word. In those instances the entire two-sentence paragraph served as the test passage. Passages ranged in size from 22 words to 134 words, with a mean of 65 words. In light of recent research on the capacity of cloze to measure intersentential comprehension (Shanahan, Kamil, & Tobin, 1981-1982), we considered this passage size sufficient to give context a reasonable chance to reveal the meaning of the target word.

After each target word had been underlined and the passages had been constructed, passages were then randomly ordered. For the words-in-context test, the target word and five items in a multiple-choice format were typed beneath each passage. Of the five choices, one was a synonym, and the other four were of the same word frequency level and part of speech as the synonym. An example of a test item follows:

> He takes out an envelope from a drawer, and takes paper money from it. He looks at it *ruefully,* and then with decision puts it into his pocket, with decision takes down his hat. Then dressed, with indecision looks out of the window to the house of Mrs. Lithebe, and shakes his head.
>
> RUEFULLY
> (A) sorrowfully
> (B) thankfully
> (C) fearfully
> (D) casually
> (E) longingly

The words-in-isolation test was identical to the words-in-context test except that the passages were not included. In effect, the words-in-isolation test was simply a multiple-choice vocabulary test.

The tests were placed in booklets which included the following sets of instructions:

> (1) *Words-in-isolation test.* Find the definition that is *most similar to or closest in meaning to* the CAPITALIZED word. Use the answer sheet to record your answers. Answer all 25 items. If you do not know an answer, then guess. Use the answer sheet to record your answers.
>
> (2) *Words-in-context test.* Each of the passages below contains an underlined word. After reading each passage, find the definition of the underlined word that is *most similar or closest in meaning* to it by choosing from one of the five lettered alternatives which follow each passage. Answer all 25 items. If you do not know an answer, then guess. Use the answer sheet to record your answers.

Procedure

We randomly assigned test booklets to subjects. The subjects filled in NCS trans-optic answer sheets, which required them to "bubble in" (i.e., blacken the circles in order to record) their answers. The words-in-isolation test required approximately 10-15 minutes for completion, and the words-in-context test took approximately 30-40 minutes for completion. We told the students to read the instructions and proceed with the task. The students were also told that they would finish at different times because they were doing different tasks. As the students finished, they turned in their test booklets and answer sheets and were given a reading or writing assignment provided by their teachers. All students finished in the allotted time.

Results

For the combined groups, the mean percentage correct was 35.8. A t test for independent observations was used to compare the mean scores on the words-in-isolation test and the words-in-context test. There was no statistically significant difference between the means of the no-context group ($M = 9.14$, $SD = 2.08$) and of the context group ($M = 8.76$, $SD = 3.72$), $t\ (99) = .552$, $p > .10$. The test for homogeneity of variance was nonsignificant. A post hoc power analysis was also performed. The results of this analysis indicated that with

50 subjects per cell, there was an 89% chance of finding an effect large enough to account for 6% of the variance in test performance.

Discussion

The absence of any statistically significant difference between the scores on the words-in-context test and on the words-in-isolation test suggests that students either could not or chose not to use context to infer the meanings of unknown words. Prior to this study, which used naturally occurring prose, most research showing the effectiveness of context used passages which were specifically designed to enhance the effects of context or to control for certain contextual factors such as form and proximity (Carnine, Kameenui, & Coyle, 1983-1984; Crist & Petrone, 1977; Gipe, 1980; Madison, Carroll, & Drum, 1982). Given that reading materials in the world outside the classroom are not artificially contrived to suggest the meanings of unfamiliar words, the results of the above-cited studies have questionable applicability to reading materials from naturally occurring prose.

Nagy, Herman, and Anderson (1984-1985) cite two other limitations of previous studies which result in inflated context effects: (a) the prompting of students to devise word meanings from context, and (b) asking students to supply a new label for a familiar concept, as opposed to the more difficult task of supplying a new label for a new concept. In spite of the fact that both of these circumstances that were expected to inflate context effects occurred in Experiment 1, no such effects were found.

Because this experiment was restricted to naturally occurring prose in fiction, it could be argued that context is ineffective only in this particular literary genre. Experiment 2 was designed to increase the generalizability of the findings and to serve as a systematic replication of Experiment 1.

EXPERIMENT 2

Method

Subjects

The sample consisted of 39 students in Grade 11 from a private school in Miami, Florida. The students came from a predominately middle- to upper-middle-class background. The average Preliminary Scholastic Aptitude Test verbal percentile score for this sample in October was 57.6.

Materials

As in Experiment 1, two tests were devised to determine the effects of context. However, in this experiment, test passages and target words were chosen from four content areas in order to determine whether the effects of context are different in different content areas. Passages and target words were identified by the same procedure detailed for Experiment 1. Typically, the science passages contained many low-frequency words per paragraph. Because of the high density of these low-frequency technical words, we did not apply the stipulation of only one to three low-frequency words per paragraph (as listed for Experiment 1) to science paragraphs.

Literature passages were taken from the same novels that were used in Experiment 1. In addition, passages were taken from popular reading (magazines and newspapers), history textbooks, and science textbooks (biology and physics). Fifteen passages from each of the four content areas were sampled, using the same procedure as in Experiment 1, for a total of 60 items. The following are examples of the passages for each of these content areas:

Newspapers

> The viewers, to their delight, will discover a *glib*, articulate man who might otherwise remain anonymous if he were not dominating the Superstars competition the way he has dominated the javelin throw this past year. *(Miami Herald)*

Magazines

No one would deny the *pragmatic* advantages, both cultural and civic, of such arts districts. Atlanta's Midtown Business Association of real estate interests and businessmen points out that 4 million visitors flock to the district each year to attend arts-related events. "The arts are up on a level now with religion and education," says association president Hiram Wilkinson. *(Newsweek)*

History textbooks

Just about the time when the Allied troops were liberating Europe from the Nazi *yoke*, American forces were routing the Japanese from their captured Pacific possessions. By the end of February, 1945, the Philippines were liberated by General MacArthur.

Science textbooks

Like worms, mollusks have three cell layers. A *coelom* is present. However, the anatomy of members of this group of animals is quite different. On the basis of both internal and external features, adult mollusks do not seem to be closely related to annelids.

The words-in-isolation test consisted of the same 60 words as the words-in-context test. The instructions were written at the top of both the words-in-isolation and the words-in-context test booklets and were worded the same as in Experiment 1.

Procedure

Each subject participated in each of the eight treatment conditions (context/no-context x 4 types of subject matter). Because we wanted to determine prior word knowledge without the possibility of context contamination, it was not considered appropriate to counterbalance the test items by including the words-in-isolation test items with the words-in-context test items. Carroll and Drum (1983) noted the interference effect of context when they expressed concern that even minimally explicit context clues could facilitate word definition. Therefore, all students took the words-in-isolation test first. All students were able to finish the words-in-isolation test within one 45-minute class period.

We administered the words-in-context test over a period of 2 days: 40 items were given on the first administration, and 20 items on the second administration. Testing was divided over a period of 2 days to eliminate subject fatigue and to ensure that each subject would have enough time to complete the test. Students were not able to complete more than 40 items on the first testing session, as the test booklets were taped closed after Item 40 to insure that the students did not look ahead at the words for the following day. As the students finished the test, they turned in their test booklets and answer sheets to us and were given a reading or writing assignment by their teacher. All tests were given in March of the school year.

Results

Table 1 shows the cell means and standard deviations for the scores on the words-in-isolation test and the words-in-context test. The data were analyzed using a 2 x 4 (treatments x treatments x subjects) design; the two factors were

Table 1 Means and standard deviations for two context conditions across four subject areas

Context Condition	Literature		Newspapers/ Magazines		History Textbooks		Science Textbooks	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Isolated	5.89	2.56	5.61	2.83	7.05	3.07	6.15	2.35
Passages	6.05	2.60	5.43	2.55	7.05	2.81	5.74	2.83
Mean	5.97		5.52		7.05		5.94	

Note. Maximum score possible was 15 for each subject area.

context condition (isolation or passage) and content area (novels, magazines and newspapers, history textbooks, or science textbooks). The criterion variable was the number of correct choices on the vocabulary test.

An analysis of variance was performed on the total number of correct responses for each subject. No statistically significant difference was found between scores on the words-in-isolation test and on the words-in-context test, $F(1, 38) = .227, p > .10$. No interaction was found to be significant. A statistically significant effect was found for type of text, $F(3, 114) = 7.94, p < .001, d = .04$. Students apparently knew more words from the history textbooks used in this experiment than from the other content area textbooks. As in Experiment 1, the sample means for the isolation conditions are higher than the context means.

Discussion

The absence of significant difference between the scores on the words-in-context test and on the words-in-isolation test across four different content areas suggests that context is an ineffective or little-used strategy for helping students infer the meanings of low-frequency words. The fact that these low-frequency words came from a greater diversity of passage types than in Experiment 1 gives this experiment greater generalizability. The only statistically significant finding was that students knew more low-frequency words from the history passages than from the other three content areas.

Both Experiment 1 and Experiment 2 relied on the use of multiple-choice tests of word meaning. Because in constructing the tests we had selected distractors in a nonrandom manner, it seemed possible that we had inadvertently biased the outcome of both studies in favor of the no-context group. Experiment 3 was a systematic replication of Experiment 1 designed to eliminate possible bias in the criterion measure.

EXPERIMENT 3

Method

Subjects

The sample comprised 84 students in Grades 10 and 11 from a private Hebrew day school in Miami Beach, Florida.

Materials

The materials used in this experiment were identical to the materials in Experiment 1 except for the vocabulary test. Instead of taking a multiple-choice test, all subjects were simply asked to write out a definition for each target word.

Procedures

The procedures were identical to those for Experiment 1 with the following exceptions: Student answer sheets were unmarked with respect to treatment group. The context and no-context answer sheets were placed in folders marked simply *A* and *B*. Both of us then judged the adequacy of the definitions in a double blind procedure: Neither of us knew the scores assigned by the other, and we did not know whether we were scoring the context or the no-context condition. Because we were interested only in full denotative meanings or accurate synonyms, all items were scored as either *right* or *wrong*. This approach has been used in other context studies (e.g., Carroll & Drum, 1982). The interrater reliability was .91.

Results and Discussion

For the combined groups, the mean percentage correct was 14. A *t* test for independent samples was used to compare the mean scores of the context and no-context groups. Again there was no statistically significant difference between the means of the no-context group ($M = 3.36$, $SD = 2.95$) and the context group ($M = 3.53$, $SD = 2.70$), $t(83) = -.29, p > .10, d$

= .06. This outcome reconfirms the results of Experiments 1 and 2 and suggests that the multiple-choice format did not bias the results.

GENERAL DISCUSSION

A standard strategy given to students for determining the meanings of unknown words has been to use the surrounding context to infer meaning. This strategy has been perpetuated from Artley (1943) to Carnine, Kameenui, and Coyle (1983-1984). Probably one of the major reasons for this perpetuation has been that researchers have used contrived or unrepresentative passages on which to base their conclusions about context.

An outstanding example of the use of contrived passages to demonstrate the effectiveness of context can be found in Duffelmeyer's study (1984). Duffelmeyer modified the vocabulary subtest of the Gates-MacGinitie Reading Tests, Level E, by replacing each isolated target word with a context-rich sentence. Here is an example (p. 105):

> When you are driving, be careful not to *exceed* the speed limit.
>
> A. go beyond
> B. do well
> C. proceed
> D. stumble
> E. approve

In addition to constructing contrived sentences, Duffelmeyer apparently made the distractors syntactically as well as semantically inappropriate. It is hardly surprising that Duffelmeyer's context group performed much better than subjects who took the vocabulary test without the benefit of any sentence context.

Carroll and Drum (1983) used more natural prose in the form of high school textbooks to investigate the effects of "explicit" and "implicit" context clues on the acquisition of word meaning. Explicit clues offer more or less precise definitions to the reader, whereas implicit clues do not provide any type of definition. All subjects were given a 60-item words-in-isolation test for low-frequency words selected from five major subject areas. There were 20 target words and 40 distractors. Subjects were asked to (a) circle *YES* or *NO* to indicate whether or not they knew the word, and (b) write a definition for the word. Definitions were scored using Drum's (1982) eight-level hierarchy of word knowledge. Two weeks later, the experimental group read the same words embedded in passages that offered explicit context clues. A control group read a modified version of the passages which provided only implicit clues for the target words.

Carroll and Drum found that the group receiving explicit context clues wrote better definitions for the target words than the control subjects, who received only implicit context clues. They also found that both the experimental and the control groups performed better on the posttest than on the pretest. In general, the study presents good evidence that people *can* use context clues to identify the meanings of low-frequency words. However, the study begs the more salient questions concerning the reliability of context clues:

1. Do traditional context clues occur with sufficient frequency to justify them as a major element of reading instruction?
2. Does context *usually* provide accurate clues to the denotations and connotations of low-frequency words?

Carroll and Drum and others have failed to come to grips with these questions by not randomly selecting low-frequency words from a variety of texts and *then* measuring the extent to which context facilitates the acquisition of word meaning. In non-contrived, naturally occurring prose passages, using context clues may be an unreliable means of inferring word meanings (Beck, McKeown, & McCaslin, 1983). The real issue is not whether or not children can use context clues, but whether or not difficult words in naturally occurring prose are usually amenable to such analysis.

Nagy, Herman, and Anderson (1984-1985) seem to have come the closest to approximating the normal reading situation while investigating

the issue of the effects of context on word meaning. They used low-frequency target words embedded in passages taken from junior high texts and found a statistically significant—but minute—effect for context in both narrative and expository prose. Superficially, it would appear that the results of their experiment are in direct contrast to the results of ours. However, because their design was very powerful, it was possible to find an extremely small effect. Our experiments actually replicate their findings. In both the study by Nagy et. al. and our study, the measures of association were extremely small—that is, for any given instance, the probability of identifying a low-frequency word was very low. In another study, Gough, Alford, and Holley-Wilcox (1981) found that the predictability of content words in sentences taken from articles in *Reader's Digest* was only .10.

It may be that on some occasions, context will reveal the meaning of an unknown word. However, as Looby (1939) noted, as often as not, context will suggest a meaning the author did not intend, and the reader has no good way of verifying that meaning without recourse to external references.

Item 2 from Experiment 1 (see Table 2) provides an example of the facilitative effects of context:

> True, the same thing happened before. Without warning, after a long season of *dearth*, three or four customers, lost faces, straggled in one day, as if they had been let out of their poor rooms with a few pennies in their pockets. And others who had skimped on food, began to buy more.
>
> DEARTH
> (A) robberies
> (B) scarcity
> (C) doubt
> (D) predicaments
> (E) bloodshed

Seventy-six percent of the subjects in the context group chose the correct answer, *scarcity,* whereas only 47% of the no-context group selected the correct answer. A test for differences between two proportions (Bruning & Kintz, 1977) indicated a statistically significant difference between the two groups. This suggests that context did help students to identify the meaning of *dearth.* Here, the low-frequency target word *dearth* receives contextual support from such words as *poor, few, pennies,* and *skimped,* all of which are associated with *scarcity,* the synonym for *dearth.* Likewise, the low-frequency word *ameliorating* (Item 20) finds its contextual support through such positive connoting words as *strength, security, civilizing, steadily, climax, triumph,* and *united.*

> Strength is the outcome of need; security sets a premium on feebleness. The work of *ameliorating* the conditions of life—the true civilizing process that makes life more and more secure—had gone steadily on to a climax. One triumph of a united humanity over Nature had followed another.
>
> AMELIORATING
> (A) investigating
> (B) improving
> (C) reporting
> (D) revolutionizing
> (E) financing

In contrast, the confounding effects of context can be seen in the results of Item 3: Every subject in the context group selected one of the wrong answers. Here, the target word *waning* receives little or no support from the surrounding context. In fact, phrases such as "the waning moon arose" and "it came up" suggest the concept of enlargement, a meaning which the author clearly does not intend. Stylistically, the author juxtaposed the idea of the anticipated prominence of the moon with its current state of inconspicuousness.

> The wind cried and whisked in the brush, and the family went on monotonously, hour after hour. They passed no one and saw no one. At last, to their right, the *waning* moon arose, and when it came up the wind died down, and the land was still.
>
> WANING
> (A) picturesque
> (B) brilliant
> (C) conspicuous
> (D) diminishing
> (E) everlasting

Table 2 shows 12 pairs of proportions of items correct that are different for the two

Table 2 Proportions of correct responses and tests for differences between proportions in Experiment 1

Item	Condition		z value
	Context	No-Context	
1	.76	.29	4.70***
2	.76	.47	2.99**
3	.00	.18	-3.33***
4	.24	.59	-3.54***
5	.44	.55	-1.10
6	.06	.37	-3.78***
7	.44	.63	-1.90*
8	.34	.31	.32
9	.28	.51	-2.35**
10	.26	.55	-2.96**
11	.46	.51	- .50
12	.24	.39	-1.62
13	.16	.16	.00
14	.44	.20	2.58**
15	.44	.02	5.00***
16	.12	.10	.32
17	.46	.61	-1.50
18	.52	.41	1.10
19	.50	.25	2.58**
20	.74	.39	3.54***
21	.08	.10	- .35
22	.58	.53	.51
23	.56	.55	.10
24	.36	.31	.53
25	.14	.16	- .28

*$z > 1.65, p < .10$ **$z > 1.96, p < .05$ ***$z > 3.30, p < .01$

groups at traditional levels of statistical significance. In 6 cases, the context group outperformed the no-context group; the reverse was true for the other 6 items. Given this population of low-frequency words, context was facilitating 24% of the time, but also had a confounding effect 24% of the time. These results are consistent with Looby's (1939) observation that context often provides misleading information about the meanings of unknown words. Until someone can demonstrate that context works better than this with larger samples than ours of randomly selected words from naturally occurring prose, we ought to be a little more guarded about class time spent teaching children to guess word meanings.

There are, of course, potential limitations in this series of experiments. First, we did not control for the subjects' formal knowledge of how to use context clues. It is possible that if the subjects had been given adequate training in using context clues, the context groups in these experiments might have performed better. We think such a result would be unlikely because the subjects were normal, fairly sophisticated senior high school students. If students don't have contextual skills by this point in time, they probably are not going to get them at all.

It is also possible to criticize the experiments on the grounds that the linguistic sample was small. A larger sample of words would certainly be desirable. However, 70 items drawn at random from five different types of naturally occurring prose passages offer a larger and more representative sample than most studies of context clues (for example, Carroll and Drum, 1983, used 20 target words).

Why Do Context Clues Work—Sometimes?

According to Finn (1977-1978), some words are easy to supply in a cloze task,

whereas others are more difficult or impossible to supply in such a task. In like fashion, some words are critical to the context, whereas others are redundant. Finn's transfer feature theory (1977-1978) explains why cloze words that are difficult to supply are also the most critical contributors to the meaning of the surrounding context. This relationship, in turn, provides a sound theoretical explanation for the results of the current study.

Finn (1977-1978) used Bormuth's original data (1966) from five alternate forms of fifth-word-deletion cloze tests. These cloze tests had been administered to 675 subjects in Grades 4 through 8. Finn reanalyzed this data base of 5,185 words according to a variety of linguistic variables such as cloze easiness (the percentage of correct cloze responses) and standard frequency index (SFI) of the cloze words. A multiple-regression analysis was run with cloze easiness as the dependent variable, and SFI as one of three independent variables.

Finn's transfer feature theory suggests an inverse relationship between the amount of information a word carries and cloze easiness. If a word is *high-information,* it contains many lexical markers that give information about itself, but there are few clues to its identity that can be gathered from the words around it—that is, the context provides little transfer feature support. Therefore, if a word is high-information, it has low cloze easiness—it would be difficult to fill in on a cloze exercise. In contrast, a *low-information* word would be easy to fill in on a cloze exercise.

Using context to guess the meaning of a semantically unfamiliar word is essentially the same as supplying the correct meaning in a cloze task. The following calculus was derived from Finn's cloze rules to illustrate the basic relationship between context and the information content of words.

Rule 1: $pC = 0$ when $x \cap y$

The probability that context will identify the meaning of a low-frequency word is zero when the word and its surrounding context share no transfer features (see Figure 1, Situation A).

Rule 2: $pC \doteq 1$ when $x \in y$

The probability that context will identify the meaning of a low-frequency word is unity when the word and its surrounding context share all transfer features, i.e., when the word is a semantic subset of or is completely redundant with its surrounding context (see Figure 1, Situation B).

Rule 3: $vC = 1/pC$

The value of a context clue as an aid to comprehension is inversely proportional to its probability of identifying the meaning of the low-frequency word—that is, there is a perfect negative correlation between the applicability of a context clue and its functional value as a reading strategy. (See Figure 1, Situation C).

Where

pC = probability of using context to identify the meaning of a low-frequency word

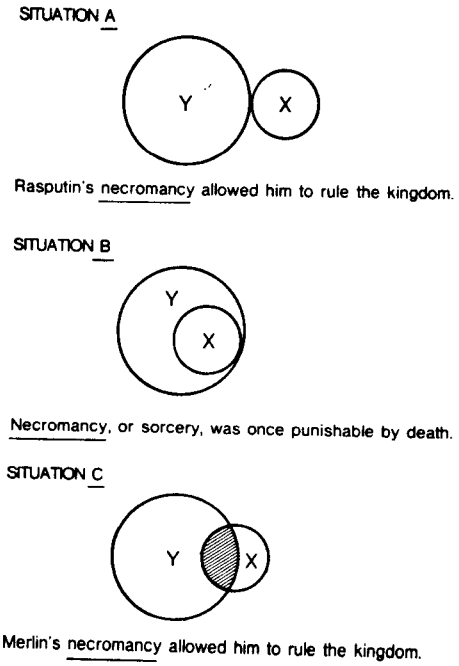
y = all semantic features of the context surrounding a low-frequency word

x = all semantic features of a low-frequency word

vC = value of a context clue as a facilitator of comprehension

Figure 1

Transfer feature relationships between low-frequency words and surrounding contexts



In Situation *A*, there are virtually no transfer features between *necromancy* (set *X*) and its surrounding context (set *Y*). The sets are mutually exclusive. The meaning of the word would be extremely difficult to predict in a cloze situation (low cloze easiness), and its meaning is critical to the meaning of the sentence—that is, it constrains the context by adding a lot of information (high-information). Our research suggests that this is the standard case for low-frequency words in naturally occurring prose.

In Situation *B* all transfer features of *necromancy* (set *X*) are included in its surrounding context (set *Y*). Set *X* is inclusive in set *Y*. The meaning of the word would be quite easy to predict in a cloze situation (high cloze easiness), and its meaning is redundant to the meaning of the sentence—that is, it fails to constrain the context (low-information). Our research suggests that this case is relatively rare in naturally occurring prose. The reason for this rarity is straightforward. When novelists and journalists, for example, compose passages that contain low-frequency words, they do so with a minimum of semantic redundancy. It is difficult to imagine Hemingway or Faulkner using a low-frequency word such as *ambidextrous* in the following manner:

> He is ambidextrous because he uses both hands equally well.

One assumes that authors use low-frequency words not for the purpose of teaching word meanings, but for the purpose of adding information or constraining the text. The following sentences serve to illustrate this concept:

> The *dog* was walked on the leash.
> The *dachshund* was walked on the leash.
> The *toddler* was walked on the leash.

In the first sentence, *dog* is redundant because it verifies but does not contribute to the text. In the third sentence, *toddler* is almost totally unpredictable and yet is critical to the meaning of the sentence. Notice also that if you did not know the meaning of *toddler*, you would almost certainly guess the wrong meaning. In the second sentence, *dachshund* carries greater information than *dog*, and the context clue will tell something about the meaning of the word; this case is identical to Situation *C* in Figure 1 and corresponds to Rule 3 in the calculus. In some cases a word will be partially predictable based on the context, but then the amount of information it contributes to the passage will be correspondingly less. The general implication is that context clues work well only when the word in question is relatively unimportant in adding meaning to the passage.

Conclusions

1. In general, context clues do not reveal the meanings of low-frequency words in naturally occurring prose. Deighton's dictum (1959) that "context always *determines* the meanings of a word, it does not always *reveal* it" is probably an overstatement. Context clues probably do not work as often as most reading educators believe.

2. Context clues appear to be just as likely to result in confusion as in the correct identification of word meaning.

3. Context clues work best when the target word is redundant with the rest of the context and contributes little new information to the passage. Or, stated another way, the more information the word contributes to the passage, the less likely it is that context clues will work.

The most obvious remaining question, then, is this: What strategy should students adopt when they come to unknown words in their reading materials? Could students be made aware of the importance of using the dictionary and/or glossary as a *first* step in the determination of word meanings early in the upper elementary grades? That is an empirical question which this study suggests is worth pursuing. Traditionally, dictionary use is the *last* step taken to get the meaning of a low-frequency word. This look-it-up-as-a-last-resort strategy may be appropriate for word recognition, but not for the comprehension of high-information words that are not in the reader's lexicon.

A dictionary-and-glossary strategy might be stressed even more as students progress through secondary school and come into contact with greater numbers of semantically unfamiliar

words. For many students graduating from high school, magazines and newspapers will constitute the bulk of their reading. If the major word-meaning strategy taught to students is to use context clues, these students may develop a misunderstanding about the meanings of many unknown words because they are likely to infer the wrong meanings. In addition, researchers have shown the importance of word knowledge to reading comprehension (Davis, 1968; Spearritt, 1972). For instance, Wittrock, Marks, and Doctorow (1975) showed that even so much as one unknown word in a passage can cause a sentence or even an entire passage to be incomprehensible.

When reading history and science textbooks, students might be encouraged to use the text's glossary. It seems that after a technical word appears once and is defined, if it reappears, it does so as a term that is undefined, yet is critical to the meaning of the text. For example, students coming across the word *lepton* for the first time in their physics textbooks may receive a thorough definition of this term. When students see it again a few paragraphs later, however, this word will probably not be supported by context clues, either intentionally or otherwise.

If one of the aims of reading instruction is to develop independent readers, then a strategy emphasizing reliance on context clues, in reality, may teach students to skip over unknown words or to infer the wrong meanings. No one study is ever definitive. However, given the results of our experiments, some of the traditional assumptions about the teaching of context clues should be rigorously reexamined.

REFERENCES

ALLINGTON, R.L. (1980). Word frequency and contextual richness effects on word identification of educable mentally retarded children. *Education and Training of the Mentally Retarded,* 118-121.

ALLINGTON, R.L., & FLEMING, J.T. (1978). The misreading of high-frequency words. *Journal of Special Education, 12,* 417-421.

ARTLEY, A.S. (1943). Teaching word-meaning through context. *Elementary English Review, 20,* 68-75.

BARNUM, E. (1906). *Teachers College Record.* New York: MacMillan.

BECK, I.L., MCKEOWN, M.G., & MCCASLIN, E.S. (1983). Vocabulary development: All contexts are not created equal. *Elementary School Journal, 83,* 177-181.

BIEMILLER, A. (1977-1978). Relationships between oral reading rates for letters, words, and simple text in the development of reading achievement. *Reading Research Quarterly, 13,* 223-239.

BORMUTH, J.R. (1966). Readability: A new approach. *Reading Research Quarterly, 1,* 79-132.

BRUNING, J.L., & KINTZ, B.L. (1977). *Computational handbook of statistics* (2nd ed.). Glenview, IL: Scott, Foresman.

BURNS, P.C., ROE, B.D., & ROSS, E.P. (1984). *Teaching reading in today's elementary schools* (3rd ed.). Boston: Houghton Mifflin.

CARNINE, D., KAMEENUI, E., & COYLE, G. (1983-1984). Utilization of contextual information in determining the meanings of unfamiliar words. *Reading Research Quarterly, 19,* 188-204.

CARROLL, B., & DRUM, P. (1982). The effects of context clue type and variations in content on the comprehension of unknown words. In J. Niles & L.A. Harris (Eds.), *New inquiries in reading research and instruction. Thirty-first yearbook of the National Reading Conference* (pp. 89-93). Rochester, NY: National Reading Conference.

CARROLL, B., & DRUM, P. (1983). Definitional gains for explicit and implicit context clues. In J. Niles & L.A. Harris (Eds.), *Searches for meaning in reading/language processing and instruction. Thirty-second yearbook of the National Reading Conference* (pp. 158-162). Rochester, NY: National Reading Conference.

CRIST, R., & PETRONE, J. (1977). Learning concepts from contexts and definitions. *Journal of Reading Behavior, 9,* 301-303.

DAVIS, F.B. (1968). Research and comprehension on reading. *Reading Research Quarterly, 3,* 499-545.

DEIGHTON, L. (1959). *Vocabulary development in the classroom.* Study reviewed in W. Petty, C. Herold, & E. Stoll (Eds.). (1967). *The state of knowledge about the teaching of vocabulary* (pp. 54-56). Champaign, IL: National Council of Teachers of English.

DUFFELMEYER, F.A. (1984). The effect of context on ascertaining word meaning. *Reading World, 24,* 103-107.

DURR, W., LE PERE, J., PESCOSOLIDO, J., BEAN, R., & GLASER, N. (1983). *Houghton-Mifflin reading program.* Boston: Houghton-Mifflin.

FINN, P. (1977-1978). Word frequency, information theory, and cloze performance: A transfer feature theory of processing in reading. *Reading Research Quarterly, 13,* 508-537.

GIPE, J.P. (1978-1979). Investigating techniques for teaching word meanings. *Reading Research Quarterly, 14,* 624-644.

GOODMAN, K.S. (1965). A linguistic study of cues and miscues in reading. *Elementary English, 42,* 639-643.

GOUGH, P.B., ALFORD, J.A., JR., & HOLLEY-WILCOX, P. (1981). Words and contexts. In O.L. Tzeng & H. Singer (Eds.), *Perception of print: Reading research in experi-*

mental psychology (pp. 85-102). Hillsdale, NJ: Erlbaum.

HUEY, E. (1908). *The psychology and pedagogy of reading.* Cambridge, MA: Massachusetts Institute of Technology Press.

JOHNSON, D.D., & PEARSON, P.D. (1984). *Teaching reading vocabulary* (2nd ed.). New York: Holt, Rinehart and Winston.

KRIEGER, V. (1981). Differences in poor readers' abilities to identify high-frequency words in isolation and context. *Reading World, 21,* 263-272.

KUCERA, H., & FRANCIS, W.N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

LOOBY, R. (1939). Understandings children derive from their readings. *Elementary English Review, 16,* 56-62.

MADISON, J., CARROLL, B., & DRUM, P. (1982). The effects of directionality and proximity of context clues on the comprehension of unknown words. In J. Niles & L.A. Harris (Eds.), *New inquiries in reading research and instruction. Thirty-first yearbook of the National Reading Conference* (pp. 105-109). Rochester, NY: National Reading Conference.

NAGY, W.E., HERMAN, P.A., & ANDERSON, R.C. (1984-1985). Learning words from context. *Reading Research Quarterly, 20,* 233-253.

NILES, O.S., DEFFENBAUGH, S.A., HYNES-BERRY, M., LAMBERG, W.J., & SAVAGE, R.C. (1983). *Skills for reading.* Glenview, IL: Scott, Foresman.

NIST, S.L. (1985). A holistic approach to teaching college reading. *Reading World, 24,* 82-87.

PATON, A. (1948). *Cry, the beloved country.* New York: Scribner.

PEARSON, P.D., & STUDT, A. (1975). Effects of word frequency and contextual richness on children's word identification abilities. *Journal of Educational Psychology, 67,* 89-95.

PERFETTI, C., GOLDMAN, S., & HOGABOAM, T. (1979). Reading skill and the identification of words in discourse context. *Memory and Cognition, 7,* 273-282.

QUEALY, R.J. (1969). Senior high school students' use of context aids in reading. *Reading Research Quarterly, 4,* 512-533.

SAMUELS, S.J. (1970). Recognition of flashed words by children. *Child Development, 41,* 1089-1094.

SCHVANEVELDT, R., ACKERMAN, B.P., & SEMLEAR, T. (1977). The effect of semantic context on children's word recognition. *Child Development, 48,* 612-616.

SHANAHAN, T., KAMIL, M., & TOBIN, A. (1981-1982). Cloze as a measure of intersentential comprehension. *Reading Research Quarterly, 17,* 229-252.

SPEARRITT, D. (1972). Identification of subskills of reading comprehension by maximum likelihood factor analysis. *Reading Research Quarterly, 8,* 92-111.

STANOVICH, K.E., & WEST, R.F. (1981). The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 658-672.

WEAVER, C. (1977). Using context: Before or after? *Language Arts, 54,* 880-886.

WEST, R.F., STANOVICH, K.E., FEEMAN, D.J., & CUNNINGHAM, A.E. (1983-1984). The effect of sentence context on word recognition in second- and sixth-grade children. *Reading Research Quarterly, 19,* 6-15.

WIENER, H.S., & BAZERMAN, C. (1985). *Reading skills handbook* (3rd ed.). Boston: Houghton Mifflin.

WITTROCK, M.C., MARKS, C., & DOCTOROW, M. (1975). Reading as a generative process. *Journal of Educational Psychology, 67,* 484-489.

Footnotes

We wish to thank the students and staff at Ransom-Everglades School in Coconut Grove, the Palmer School in Miami, the University School in Fort Lauderdale, and the Greater Miami Hebrew Academy in Miami Beach, Florida for participating in this study.