DOCUMENT RESUME

ED 230 918 CS 007 193

AUTHOR Bruce, Bertram; And Others
TITLE Vocabulary Bias in Reading Curricula. Technical Report No. 280.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Jun 83
CONTRACT 400-76-0116; 400-80-0031
NOTE 58p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Blacks; Content Analysis; Language Acquisition; Language Research; Language Usage; Lower Class Students; Middle Class Students; Primary Education; *Racial Bias; Racial Differences; *Readability Formulas; *Reading Materials; Reading Research; *Social Bias; Social Differences; *Vocabulary Skills; Whites; *Word Lists; Working Class

ABSTRACT

A study examined the match between the vocabularies of children of different ethnic and socioeconomic status groups and the school vocabulary revealed by readability formulas and word lists. The Spache 1040 and the Dale 769 readability formula word lists were used as indicators of school vocabulary in the early primary grades, and a corpus of talk involving thirty-nine 4.25- to 5-year-old children grouped according to race and social class served as the indicator of the children's vocabularies. A comparison of the two vocabularies showed two significant biases in the readability formula word lists: (1) against working-class as opposed to middle-class children (which was evident on both word lists), and (2) against black was opposed to white children (which was more pronounced on the Spache than on the Dale list). (Author/FL)

CENTER FOR THE STUDY OF READING

Technical Report No. 280

VOCABULARY BIAS IN READING CURRICULA

Bertram Bruce
Andee Rubin
Kathleen Starr
Cheryl Liebling

Bolt Beranek and Newman Inc.

June 1983

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

This research was supported by the National Institute of Education under Contract Nos. HEW-NIE-C-400-76-0116, and 400-80-0031. We would like to thank Paul Horwitz, John Frederiksen, Dave Getty, Ed Smith, Bob Linn, Bill Nagy, Bill Hall, Marilyn Adams, and Jeanne Chall for discussions and comments on the analyses presented here and Cindy Hunt for manuscript preparation.

Abstract

The research reported here addresses the question: How well does the increasingly standardized vocabulary of the school match the words familiar to children of different social class and ethnic groups? Readability formula word lists (Spache 1040, Dale 769) were used as indicators of school vocabulary in the early primary grades. A corpus of talk involving 39 children (ages 4-1/2 to 5) grouped according to race and social class served as an indicator of the children's vocabularies. Comparisons of the vocabularies show two significant biases in readability formula word lists. The first bias, against working-class as opposed to middle-class children, is evident on both the Dale and Spache lists. The second bias, against Black as opposed to White children, is most pronounced for the Spache revision of the Dale list--a revision that was designed to make the word list reflect the school vocabulary more accurately.

Vocabulary Bias in Reading Curricula

A child's vocabulary is indicative of his or her cultural background, interests, and personal experiences. In an analogous way, the vocabulary of a text is indicative of its subject matter, point of view, and so on. Although the vocabulary match between a child and a particular text may be only a small factor in any one reading experience, the match of the vocabulary of a group of texts with the child's vocabulary is a good measure in general of how easy those texts will be to read.

In order to assess this match one needs accurate knowledge of the words children are exposed to and use at home and in school. In addition, one needs an estimate of the essential school vocabulary that children are expected to master. We have been fortunate in both respects. The Hall corpus (Hall, Linn, & Nagy, in press) is an excellent gauge of the vocabulary knowledge of children of different SES and ethnic groups as they are about to enter school. The school vocabulary manifested in basal readers, workbooks, tests, supplementary materials, and textbooks is reflected in the word lists used in formulas designed to assess readability. Since these word lists were compiled in part from the very sources they are now used to measure and modify, they both reflect and influence the vocabulary found in school materials.

In this report, we look at the match between the vocabularies of children of different ethnic and socio-economic status groups and the school vocabulary revealed by readability formula word lists. Our data indicate two sources of bias in readability formula word lists. One significant bias is in favor of middle-class, as opposed to working-class, children. The second significant bias is in favor of White, as opposed to Black, children. These biases exacerbate the problems that working-class and/or Black children encounter in school. An understanding of how these biases have evolved may help in countering their effects.

Readability Formula Word Lists as a Window on School Vocabulary

A readability formula is a method of assigning a numerical estimate of "readability," variously defined as "ease of reading," "interest" or "ease of understanding" (Gilliland, 1972), to a text. Because readability formulas are intended as quick and convenient measurements, they typically take into account only easily-measurable aspects of a text such as word difficulty and average sentence length. A weighted combination of these measurements yields a number for each text. The resulting estimate is usually intended to represent a grade level.

One of the most popular readability formulas in current use for primary-grade materials was devised by Spache

(1978). Using it requires choosing three to five 100-word selections from a book, measuring the percentage of uncommon words (based on a 1040-word list of familiar words) and the average number of words per sentence in the passages, then combining the two numbers according to the equation:

Reading grade = .082(% uncommon words) +
.121 (average number words per sentence) + .659

For example, consider the beginning of the story Frog and Toad: Down the Hill from a children's book by Arnold Lobel (Lobel, 1976).

> Frog knocked at Toad's door. "Toad, wake up," he cried. "Come out and see how wonderful the winter is!" "I will not," said Toad. "I am in my warm bed." "Winter is beautiful," said Frog. "Come out and have fun." "Blah," said Toad. "I do not have any winter clothes." Frog came into the house. "I have brought you some things to wear," he said. Frog pushed a coat down over the top of Toad. Frog pulled snowpants up over the bottom of Toad. He put a hat and scarf on Toad's head. "Help!" cried Toad. "My best friend is trying to kill me!"

In this 104-word passage, there are 16 sentences, for an average sentence length of 6.5. According to the Spache

1040 list, there are 4 or 3.8 percent unfamiliar words, so by the Spache formula:

$$\text{Reading grade} = .082(3.8) + .121(6.5) + .659$$
$$= .311 + .787 + .659 = 1.8$$

With two other samples from this story the Spache grade level estimate is 1.7.

Readability formulas are used in a variety of situations where estimates of text complexity are thought to be necessary. Educational publishers use them in designing basal and remedial reading texts; some states, in fact, will consider using a basal series only if it fits certain readability formula criteria. Oregon, for example, demands that basal publishers provide the average readability for each book, the highest and lowest readability scores in each book, the number of samples on which each score is based and the actual readability worksheets (Robert Tierney, Note 1). Standardized reading comprehension test manufacturers use readability formulas to rate and modify the grade level of test passages.

Readability Formula Word Lists

Readability formulas were first developed in the 1920's for use by textbook writers; in the past fifty years hundreds have been proposed (Klare, 1976). An important measure in many of these formulas is the vocabulary load, or

percentage of hard words in a text. In studies by Lorge (1944), Flesch (1943), and Dale and Chall (1948), the measure of vocabulary load was found to be the most important factor in determining reading difficulty. To calculate this load, formula designers have compiled and used a number of different word lists.

Most of the early lists were based on frequency counts of words sampled from texts. For example, the Teacher's Word Book (Thorndike, 1921 and later revisions in 1932 and 1944) listed the most frequently used words found in a wide range of sources from the Bible and English classics to popular adult magazines and children's books. Certain sections of the Thorndike list - especially the first thousand most common words - have been used both as the base list of easy words in readability formulas and as a source in the development of other word lists.

Criticism of word counts sampled only from printed materials led to lists based on studies of the writing vocabularies of both children and adults (Tidyman, 1921; Horn, 1926). Other lists were compiled from spelling lists, vocabulary found in primary reading series, and counts of the spoken vocabulary of young children (Horn, 1925; International Kindergarten Union, 1928). These lists have been adapted, revised, and combined in various forms (Lorge, 1944). Buckingham and Dolch (1936) included words from

their own and nine other word counts in their Combined Word List. Dale (1931, 1943) used many of these same counts to compile his two lists of common words.

Word lists for currently used readability formulas are still based on the word counts done in the 1920's and 1930's. Though some researchers have stated that familiarity in the spoken language is one of the principal factors involved in making words easy for beginning readers (Stone, 1956), revisions of the early lists have been based almost entirely on vocabulary counts from written texts, primarily basal reading series. To show the course of development of a particular list used in one readability formula, we will trace the history of Dale's list of 769 Easy Words and its use in the Spache formula (Spache, 1978).

To compile the 769 Easy Word list, Dale compared the International Kindergarten Union List (1928) and Thorndike's first 1000 words (1921), and selected words common to both (Dale, 1931). Spache used the Dale 769 list in his original formula (Spache, 1953). Later, Stone produced a revision of the Dale 769 list that he claimed increased the accuracy of the Spache formula (Stone, 1956). Stone chose two new sources for easy primary reading words: his own study of twenty-one primary reading series published in the 1930's (Stone, 1936), and a list by Krantz (1945) based on a study of words used in 369 primary reading books. Both of these

studies rated words on the basis of the grade level at which each word was introduced in different reading series. Stone revised the Dale 769 list by replacing 173 of the original words with 173 words rated easier in both studies.

Spache adopted Stone's Revised List as the base word list for his formula, and continued to use it for almost twenty years. When he revised his formula in 1978, Spache believed that the Stone list no longer represented the vocabulary found in school books. To modify the list again, he used three sources: a sample of supplementary reading materials published for first and second grades, a study of the meaning vocabulary of first graders (Dale & Schuh, 1970), and a frequency count of words in six basal reading series and six other textbook series (Harris & Jacobson, 1972). Based on these lists, 94 words were deleted from the Stone List and 365 new words were added for a total of 1040 words on the new Spache list. Ironically, nearly 30 percent of the 365 words added to the Stone list had originally been on the Dale 769 list. Spache believed this new list to be a better reflection of the vocabulary present in basal readers and supplementary books for the primary grades (Spache, 1978), and thus a better measure of reading difficulty. Since the Spache formula is so widely applied to primary grade materials, we have used the Dale 769 and the Spache 1040 lists as the basis of our investigations. We have also examined the 365 words Spache added as a way of separating

the characteristics of the older list and the newer additions. We consider the Spache 1040 list to be composed of the Dale 769 list and the Spache 365 added list, even though it is clear that they don't "add up" to 1040. The 94-word difference is accounted for by the words Spache deleted from the Stone list before he added his own 365 words. The results of our comparisons of these three lists with the vocabularies of the children in the Hall corpus are discussed below.

The Hall Corpus Word Lists as a Window on Children's Vocabularies

The Hall corpus is an ambitious study of the words children produce and perceive; as such it provides us with a view of the oral and aural linguistic environments of children of different social class and ethnic groups--of the words with which they are "surrounded." The Hall corpus contains all the word tokens which were not only spoken by the children under study but also spoken to them by adults within specified situations of language use. The total corpus of words in the children's linguistic environments contains some 1,058,943 tokens. As described in detail in (Hall et al., in press, Chapter 1), the children in the Hall study are categorized by race and socio-economic status. Our analysis is based upon the vocabularies of the four groups determined by varying both of these characteristics and referred to as G_1 through G_4.

G_1 = Black middle-class vocabulary

G_2 = White middle-class vocabulary

G_3 = Black working-class vocabulary

G_4 = White working-class vocabulary

In fact, the word "vocabulary" is somewhat misleading in this context. Definitions of vocabulary abound, each with its own strengths and methodological problems. Lorge and Chall (1963) present a well-organized, thoughtful discussion of some of the major methodological difficulties in estimating vocabulary size. Our particular concept of vocabulary, however, might better be termed "familiar words" since we are focusing not on the edges of children's vocabularies, but on the more central parts. We define the relevant sets of familiar words for each group of children using the frequency with which each word was spoken by a child (even though the corpus also includes words spoken in the child's linguistic environment by other members of the family and the experimenter). In our analysis, we have considered both the relative and absolute frequency with which children used words.

From the total number of spoken words contained in the Hall data, we selected the 1000 most frequently spoken tokens for each of the four groups of children (based on a measure which takes into account how evenly spread the occurrences are within the group). After deleting token

duplications such as pronunciation variants, proper names, regular verb and noun parts as well as nonsense syllables, letters and numerals, we arrived at four sets of different sizes. The smallest of these contained 732 words, so we considered our analysis set to be the 732 most familiar words in each group. For each group, we considered words familiar to one child to be familiar to the group as a whole. This constituted our measure of relative frequency.

We also considered absolute frequency. Using a threshold that meant that, on the average, a given word was used at least 5 times by each child, we arrived at four sets of different sizes.

Insert Table 1 about here

The most basic piece of comparative information about these word sets is their relative size, as displayed in Table 1. (Notice the totals refer to "types," not "tokens." In other words, several occurrences of the same word are counted only as one word). The most noteworthy fact about this table is that G_1 and G_2 produced an almost identical number of word types with absolute frequencies ≥ 45. In addition, the middle-class vocabulary contains (substantially) more word types in everyday situated language than does the working-class. The pattern of numbers suggests that class is a more potent determiner of

vocabulary size than race and that for Black children, class makes more of a difference than for White children.

The larger size of the middle-class familiar everyday vocabulary of frequently used words suggests that, by virtue of its size alone, chances for a match with readability formula words is necessarily greater. The rest of our analysis is based on the entire list of 732 words for each categorical group. Note that this, in effect, gives the working-class vocabulary a built-in "advantage"; because we are including words whose absolute frequency is lower for the working-class vocabulary, we are more likely to obtain a match with readability word lists. Any inequities in matches between middle-and working-class vocabularies, then, should be taken even more seriously.

The data represented in Table 1 plus other preliminary analyses of the four sets of familiar words indicate that comparisons across class and race were the most significant. Therefore, the rest of our analysis is based on data combined in the following way, yielding four comparison groups.

$G_1 + G_2$ = Middle-class vocabulary (Black and White)

$G_3 + G_4$ = Working-class vocabulary (Black and White)

$G_1 + G_3$ = Black vocabulary (Middle-class and Working-class)

$G_2 + G_4$ = White vocabulary (Middle-class and Working-class)

In the next section, we will compare the vocabularies for each of these categorical groups to the "ideal" vocabulary implied by readability word lists.

The number of familiar word types as reflected in Table 1, of course, tells only a small part of the story. We would like to be able to answer questions about the relationships between the most frequent spoken words of categorical groups. For example, what types of words can be considered common to the middle-class (G_1 and G_2) and working-class (G_3 and G_4) children's most frequently spoken words? Or do these categorical groups have very few words in common? Are there words middle-class children use which working-class children don't, and vice-versa? Likewise, we could ask similar questions in comparing the most frequently spoken words of Black and White children. Table 2 spells out the relationships that answer these questions.

Insert Table 2 about here

The first column demonstrates that the middle-class and the working-class as well as Whites and Blacks share approximately 700 words of their most frequently spoken words. This pattern suggests the notion of a core vocabulary, i.e., a set of words familiar to children regardless of class and/or race. It seems likely that these core words are essentially the words anyone would suggest as common words for five-year olds; the actual list supports this general impression. By extension, we can conceptualize a common language spoken by five-year-olds. In general, they can communicate with one another using words familiar to all, even though each child brings a somewhat different vocabulary to the communicative situation.

The differences among the most frequently spoken words are equally illuminating. While the class groups share a core vocabulary as do the race groups, there are 180-200 words which can be considered distinctive to each categorical group.

Insert Figure 1 about here

The second and third columns in Table 2 begin to define by example the notion of a distinctive vocabulary. Roughly, a distinctive vocabulary is a set of words included in one vocabulary but not in the other vocabulary (or vocabularies) with which it is being compared. Figure 1, for example,

illustrates the distinctive vocabularies that result when two vocabularies--middle-class and working-class--are compared. The striped area represents the middle-class distinctive vocabulary with respect to the working-class, since it excludes from the middle-class just those words which the middle-class and working-class share: those in the intersection of middle-class and working-class, represented by the blank area in the diagram. The last column in Table 2, for example, reports the size of the distinctive vocabulary of the second vocabulary as compared to the first. Formally, we can define the distinctive vocabulary of A with respect to B (DV(A;B)), where A and B are both vocabularies, as

$$DV(A;B) = A-(A\ B).$$

In other words DV(A;B) are all those words in A which are not also in B.

While this definition is unambiguous for the case of two vocabularies, it is more complex when more than two are involved. We could extend this definition of distinctive vocabulary to more than two groups, but for the purposes of this analysis, we will limit our definition to comparisons of two groups.

The Match Between Children's Oral Vocabularies and Readability Formula Word Lists

In this section we discuss first a relatively simple measure of the match between vocabularies and word

lists: the overlap (intersection) of each readability formula word list with the vocabularies of the class or race groups. Next, we list some of the specific words that readability formula lists assume are common but that are infrequently used by at least one group of children. Finally, we present a more detailed statistical analysis of the match between the two sources of familiar words.

Passive Bias: Simple Mismatches Between the Two Lists

Our first analysis consists of determining how many words in the three readability formula word lists described above are not in children's vocabularies. The consideration of these words as "familiar" in the calculation of readability scores will make texts appear easier than they actually are for children who do not frequently use the words.

Insert Table 3 about here

Table 3 shows the number of words in each of the three lists which are distinctly familiar to each of the categorical groups of children. For example, 75 of the words on the Spache 1040 list are distinctly used by the middle-class children (G_1 & G_2), while 62 of them are distinctly used by the working-class children (G_3 & G_4). Looking at all columns, we see that with respect to class, the Dale 769 list contains the most words which are

distinctly used by the middle-class as compared with the working-class. At least insofar as the most frequently spoken words are concerned, the Dale 769 words are more familiar to the middle-class. Similarly, with respect to race, the Spache 1040 list contains the most words which are distinctly used by White children when compared with Black children; the Spache 1040 words are more familiar to the White children, at least with respect to most frequently spoken words. These observations imply that readability estimates will be more accurate for middle-class children than for working-class children and for White children than for Black children; or, in other terms, that a formula using the Dale 769 list is biased against working-class children and one using the Spache 1040 list is biased against Black children.

Since the Spache formula developed from two distinct sources, it is instructive to examine the two parts of its associated word list with respect to the observed bias against Black children. The numbers in the third column, representing the vocabularies' match with the Spache added 365 list, display large differences when comparing White and Black children. In fact, only 18 of the 365 words on the Spache added list appear in the Black vocabulary of most frequently spoken words, while only 34 of these words are frequently used by White children. Comparing Black vs. White children's distinctive vocabulary with the Dale 769

and Spache 1040 lists, we can see that more of the lopsided quality of the Spache 1040 is traceable to the Spache additions than to the original Dale list.

Insert Table 4 about here

In part, history explains this discrepancy since the Dale list had "used up" many of the most common words familiar to all children and Spache, in venturing outside this core vocabulary, was more likely to choose words unfamiliar to at least some children. Table 4 provides numerical support for this argument. While 49% of the Spache 1040 list is in the core vocabulary for race comparisons (i.e., frequently used by both Black and White children), only 20% of the Spache added words are in the core for race. The large majority of the core words in the Spache list come from the Dale list. Thus Spache, in choosing words to add to the list, had to rely more on non-core words and, in essence, he chose more words from the White children's most frequently spoken words than from the Black children's most frequently spoken words.

It is possible, of course, that this lack of balance occurred because there are more words to begin with in the middle-class and White children's vocabularies of the most frequently spoken words. The reader will recall from Table 1 that the middle-class vocabulary, irrespective of race,

comprises an average of 212.5 words (of 732) and that the White children's vocabulary, irrespective of class, comprises on the average 203.5 words. The working-class vocabulary, however, contains an average of 185 words and the Black vocabulary an average of 195 words. The middle-class and White vocabularies of words with absolute frequencies ≥ 45, thus, are larger to begin with.

Thus, even a "fair" algorithm (a notion we will define precisely below) for adding new words to a readability list may have resulted in the kind of bias we see here. For this reason, we call the picture of bias we have sketched "passive bias" since it may be due to naturally-occurring differences in the size of different groups' most frequently spoken words. A contrasting view of "active bias" will be presented below.

Example Words

The statistics just presented characterize the match between the readability formula word lists and the children's oral vocabularies, but only numerically. To make more specific observations about the types of words that differentiate the four groups one needs to look at the lists themselves. In Table 5 we show a small subset of all the words under consideration, namely, those words that are on the Dale 769 list and on the distinctive spoken words lists for the middle-class and working-class groups. These are

the words referred to numerically in the first two rows of the second column of Table 3. In Table 6, we show a second subset of words, namely, those words that are on the Spache 1040 list and on the distinctive spoken words lists for the Black and White groups of children. These are the words referred to numerically in the last two rows of the first column of Table 3. It would be informative as well to look at those readability words that are in the core vocabularies, those that never intersect with the most frequently spoken words, and frequently spoken words which do not appear on any readability word lists. The subsets we are presenting, however, offer some important insights into the structure of both the children's spoken vocabularies and the word lists.

Insert Tables 5 and 6 about here

There are a number of observations about the patterns one sees in Table 5. It is worth noting, first of all, that there are only 109 words listed, that is, about 14% of the Dale 769 list. Given the observation noted in Table 4 that 61% of the Dale 769 list is frequently used by middle-class and working-class groups, we know that the Dale 769 is in one sense fair; most of its words are familiar across class. The words in Table 5 are those which are distinctly familiar to the middle-class or working-class groups.

Further examination of Table 5 supports the observations made on a numerical basis above. When a word is distinctively familiar to one group, it is more likely to be frequently used by children of the middle class. This can be seen by inspection of the list for the working class with respect to the list for the middle class. In effect, the list of distinctive vocabulary is skewed towards words that White children frequently use.

Table 6 lists 144 words, about 14% of the Spache list, which are distinctively familiar to White and Black children. Here too, a definite pattern emerges: When a word is distinctively familiar to one group, it is more likely to be frequently used by White children. As with Table 5 for class, the lists of distinctive vocabulary with respect to race show that the Spache 1040 is skewed towards words that White children frequently use.

Although the number of words in Tables 5 and 6 is relatively small, it is interesting to note some patterns in their distribution. These are, then, hypotheses which might be investigated in further vocabulary studies. The middle-class distinctive vocabulary contains a group of words related to emotion or thought--"afraid," "dream," "knew," "laugh," "surprise" and "wonder." The working-class list contains only "cry." (This is consistent with Hall, Nagy and Nottenburg's (1981) analysis of internal state

words). Another contrast is in animal and outdoors words. The middle-class list contains "animal," "bee," "butterfly," "feed," "grass" and "land"; the working-class list contains only "sheep." These differences may be a reflection of the children's experiences (e.g., trips to the country), or their home environments.

The lists in Table 6 hint at other patterns. Words referring to emotion or thought are comparably represented on the two lists, but the White list contains more animal words which may come from books: "elephant," "tiger," "sheep," "wolf" and "turkey." In both of these lists, the patterns are only suggestive: No definitive statements can be made without further study.

A Case for Active Bias

In contrast to the definition of "passive bias" given above, we will now discuss the notion of "active bias" and make a case for its existence in the construction of the Dale 769 and Spache 1040 lists. The basic idea is this: In "passive bias" the differential representation of various groups' vocabularies in word lists is attributed to the varying vocabulary sizes; in "active bias" there is an additional claim that beyond the effect of different vocabulary sizes, words are more frequently chosen from some groups' vocabularies than from others'. To assess the possibility of active bias across class in these lists, we need to know for both middle- and working-class:

(a) how many distinctive words we would expect to be chosen from their vocabulary based on the relative sizes of their distinctive vocabularies

(b) how many distinctive words were actually chosen based on the readability formula word list.

A major discrepancy between these two values would indicate active bias. The corresponding values for Black and White vocabularies could be used to assess active bias across race.

Passive and active bias: An analogy. An analogy to clarify the distinction between passive and active bias might go as follows: Suppose you were choosing bulbs for your garden out of a large sack which contained different numbers of tulip, crocus, daffodil and hyacinth bulbs. If you chose bulbs randomly (with your eyes closed), you would end up with a batch of bulbs in which the distribution among tulips, crocuses, daffodils and hyacinths mirrored the distribution in the sack. If the sack had 75% crocus bulbs, your selection would similarly be overloaded with crocuses. This situation is one of passive bias. If, however, you opened your eyes and picked out some extra crocus bulbs, the number of crocus bulbs you had would be due both to their preponderance in the sack and to your choosing extra crocuses. This latter situation represents the addition of active bias. Although this analogy mirrors the two types of bias in word lists correctly, there is one crucial

difference. Nowhere are we claiming that any active bias detected in readability formula word lists is intentional. In the word list scenario, there is no counterpart of "opening your eyes" and deliberately choosing particular types of words.

An example

The remainder of our discussion will present evidence for active bias in the Dale 769 and Spache 1040 lists. We will go through one example in detail, then just present the results of the other analyses. Suppose we wanted to compare middle-class and working-class vocabularies. The first step is to calculate the relative sizes of the two distinctive vocabularies. As shown in Table 2, the middle-class distinctive vocabulary contains 182 words and the working-class distinctive vocabulary contains 196 words. Thus, the total "distinctive vocabulary pool" from which words could be drawn is 182 + 196 = 378.

If words were chosen from these two groups of distinctive words strictly on the basis of their size, we would expect 182/378 or .482 to come from the middle-class distinctive vocabulary. This is the expected probability.

The essence of the calculation consists of comparing the expected probability with the actual ratio between words drawn from the distinctive vocabulary of the first group and

those drawn from the distinctive vocabulary of either group. For this step, we need to consider the intersections of the two groups' vocabularies with the word list in question. By a calculation similar to those above, we find that the number of words, say, in the Spache added 365 list, which are also in the distinctive vocabulary of either middle- or working-class is 42. Of these 42 words, 19 of them come from the middle class. Thus, the actual ratio we need to compare with the expected probability calculated above (.482) is 19/42 =.452. By inspection, it seems clear that these two fractions are not significantly different. Using the standard binomial probability comparison formula, we get a Z-value of -.377, which supports the null hypothesis of no significant difference between the two ratios. We interpret this as saying that any discrepancy in the number of words chosen from the distinctive vocabulary of middle-class children (19) and that of working-class children (23) can be linked to the difference in their relative sizes (182 to 196).

Class and race comparisons. Comparing the expected probability of middle-class vocabulary words and their actual ratio in word lists suggests a bias in favor of middle-class words. As shown in Table 7, even though we expect only 48% of the distinctive vocabulary words on the Dale 769 list to come from the middle-class distinctive vocabulary, 57% actually come from that source. This

translates into an increased tendency above and beyond the difference in vocabulary size between the two groups for the Dale list to contain middle-class words. The only other word set on the list which shares this indication of bias is the words which are common to both the Dale and Spache lists. It appears that most of the bias in the Dale list is due to words it shares with the Spache list. Words that are only on the Spache list, in fact, contain fewer than expected middle-class words (but not significantly); the Spache list thus does not appear to be significantly biased, although the non-significant trend is in that direction.

Insert Tables 7 and 8 about here

Table 8 shows similar comparisons across race, but this time the Spache list exhibits significant bias. Further consideration shows that the words ccmmon to the Dale and Spache lists were relatively unbiased, but that the words Spache added to the Dale list were biased enough in favor of the White distinctive vocabulary that the resulting Spache list was also biased.

Consolidation of results on bias. Returning to our flower bulb analogy, what have we discovered about the flowers in the readability formula word list garden? First, we have found that the distribution of bulbs in the sack is not uniform--that there is a preponderance of hyacinths and

tulips (middle-class vocabulary words), but not as many crocuses or daffodils (working-class vocabulary words). Thus, as we would expect, there are more hyacinths and tulips in the garden. We called this phenomenon passive bias.

Second, we have found that the gardener is not choosing from the sack at random, but is occasionally picking an extra tulip or hyacinth (middle-class vocabulary word or White vocabulary word), so that the tulips and hyacinths are even more plentiful than they would be by virtue of their larger numbers in the sack. We called this process active bias.

It is important to reiterate a crucial difference here between our flower garden analogy and readability formula word list construction. While the gardener could be conceived of as purposely choosing additional hyacinths, there is no implication that list designers are intentionally favoring middle-class children or White children. Spache, after all, used published educational materials in updating the Dale list to the Spache 1040 list. Those materials, however, as part of the same educational culture, reflect the same bias and were based, in fact, on older word lists. Using them to update word lists is not only circular; it also perpetuates any (unintentional) bias present in the original lists. Spache, Dale and other word

list designers unintentionally but effectively build class and race bias into their lists.

Implications

The results presented in this paper are from a well-balanced sample of children's oral vocabularies and two popular readability formula word lists, and we believe they have important general implications. These range from those pertaining to the use and interpretation of readability analyses to those concerned with an emerging picture of a school reading curriculum biased against working-class and Black children.

Consider the beginning of "The Little Knight," a story from Scott Foresman's Reading Unlimited series.

> Once upon a time a king and a queen lived in a big old castle. The king and the queen were sad because their castle was so cold. Sometimes the queen had to put on a blanket to keep warm. And the king had to put on an old rug. Then they didn't look like a king and a queen.
>
> Something else made the king and queen sad. They couldn't sleep because a dragon kept them awake. Every night the dragon sat in his cave on the top of the hill. And he roared and roared and roared.

According to the Spache list there are three unfamiliar words in this passage--"rug," "couldn't," and "cave." Applying the Spache formula to the whole story, we get a grade level estimate of 1.9, with a total of 8 unfamiliar words. However, some of the words considered familiar by Spache we found in a preliminary analysis to be relatively unfamiliar to working-class children in the Hall corpus--"queen," "castle," "dragon," and "awake". If we count these as unfamiliar words when applying the formula to the whole story, the grade level jumps to 2.4, with a total of 18 unfamiliar words. This estimate would be a better reflection of the difficulty of this story for many working-class children.

Readability formulas would seem most needed in rating text difficulty for children not in the White middle-class. Unfortunately, it is in this situation that they are least reliable. While the standard error of estimate for the Spache formula is two months (i.e., the true grade level of a text could be as much as two months more or less than the estimate), the difference between the Spache grade level and our revised estimate is five months. Thus, for children of the appropriate background the formula may not be too far off, but for others the formula will merely assert that the story is readable and thereby put the blame on the children's "vocabulary problem," possibly causing them to be labeled "poor readers." In order to avoid this situation,

great care is needed both in determining when use of a readability formula is appropriate and in interpreting the formula scores.

Our example illustrates a major drawback of readability formulas: They do not reflect different readers' social and cultural backgrounds. This is hardly surprising, as the original compilers of readability formula word lists attempted to capture the vocabulary found in school materials and other texts that are strongly representative of White middle-class America. Thus, the extent to which the readability formula word lists fail to match the vocabularies in the Hall corpus reflects the failure of school texts to match the background, experience, and culture of many of the children who use them. These children must do more than learn new words; they must become familiar with a new culture. Revising the word lists would not be sufficient to correct the mismatch. School texts must also be revised to reflect the diversity of our society. If curricula are not changed, we must at least be aware that we are demanding much more of those children whose lives are not represented in the materials they use in school.

Conclusion

Vocabulary is a reflection of dialect, knowledge, experience, and interests, among other things; in short, it

is a measure of many of the factors that influence success on tests and in school. The reader's vocabulary knowledge is the most accurate single predictor of reading comprehension and IQ test scores (Anderson & Freebody, 1979). For these reasons, and because it is easy to quantify, vocabulary is one of the principal factors in readability formulas. No simple measures of vocabulary and sentence length, however, can account for other factors which do make a particular text difficult, such as discourse cohesion, the number of inferences required, the number of items to remember, the complexity of ideas, rhetorical structure, and the knowledge of literature assumed. Since these dimensions are much harder to measure, further research is needed to determine if the bias found against working-class and Black children in the readability formula word lists is indicative of a mismatch in other text dimensions as well.

Readability formulas are only one component of a complex system of educational materials. While their limitations have often been discussed,[1] readability formulas are widely used and play an important role in the educational system. They interlock with standardized tests and curricula to present a unified educational approach which does not address the needs of many children, especially those of lower socio-economic status. For example, standardized tests assert that some students lack

the aptitude for success in school. These tests are of the type once used to validate readability formulas. But now, readability formulas are used to adjust passage difficulty on the tests. Books for beginning readers (primarily, basal readers) served as a source for the word lists for readability formulas; now the formulas are used in the preparation and editing of basal readers. While basal publishers do not in general give authors explicit instructions to tailor their stories to readability formulas, the formulas are used to choose the most appropriate passages, adapt them to particular grade levels, and sequence them in order of increasing complexity.

Other investigations have provided evidence that complements the analyses presented here. Hall and Tirre (1979) discovered that the words used on four standard intelligence tests (including the Stanford-Binet) more closely reflected middle-class vocabulary than working-class vocabulary. In addition, they demonstrated that middle-class children produce even more "school words" at home than they do at school. For some of them, school may seem like a watered-down version of their home environment. For working-class children, on the other hand, school may present a bewildering package of new words and situations to master. And, it must be remembered that the biases evident in the composition of a "school vocabulary" are only the tip of the iceberg. The effect of the school environment itself

has been shown to influence children's vocabulary. Hall, Nagy and Nottenburg (1981) cite evidence that Black children use fewer internal state words in school than they do at home.

It should come as no surprise that talking in school is different from talking at home or on the street; Roger Shuy (1981) reminds us that "the language of the classroom is one context out of many possible daily language contexts" (p. 170). What is disturbing is the combination of emphasis placed on school language and culture by the society at large: "Educators single out the ability to talk effectively in schools as the norm for effective talking" (Shuy, p. 170) and the bias inherent in the definition of that culture. There is, in the final analysis, a complete circularity, from school talk to tests to curricula to readability formulas. The circular system strongly reflects the background and needs of White middle-class America. Thus, the bias found through our analysis may be indicative of a larger bias in our educational system, one that it is important to understand for the good of our children and our society.

Reference Note

1. Tierney, R. J. Personal communication, September 1981.

References

Bruce, B. C., Rubin, A. D., & Starr, K. L. Why readability formulas fail (Reading Ed. Rept. No. 28). Urbana: University of Illinois Center for the Study of Reading, August 1981.

Buckingham, B. R., & Dolch, E. W. A combined word list. Boston: Ginn and Co., 1936.

Child Study Committee of the International Kindergarten Union. A Study of the vocabulary of children before entering first grade. Washington, D.C.: International Kindergarten Union, 1928.

Dale, E. A comparison of two word lists. Educational Research Bulletin, 1931, 10, 484-489.

Dale, E. Easy word list. Columbus, Ohio: Bureau of Educational Research, Ohio State University, 1943.

Dale, E., & Chall, J. S. A formula for predicting readability. Educational Research Bulletin, 1948, 27, Jan. 11-20, and Feb. 37-54.

Dale, E., & Schuh, E. A list of 1400 Words Known by 75% of More of First Grade Children in the Enrichment Program of the Columbus (Ohio) Public Schools. Columbus, Ohio: Ohio State University, March 1970.

Davison, A., Kantor, R., Hannah, J., Hermon, G., Lutz, R., & Salzillo, R. Limitations of readability formulas in guiding adaptations of texts (Tech. Rep. No. 162). Urbana, Ill.: University of Illinois Center for the Study of Reading, March 1980.

Flesch, R. F. Marks of readable style. New York: Bureau of Publications, Teachers College, Columbia University, 1943.

Gilliland, J. Readability. London: University of London Press Ltd., 1972.

Hall, W. S., Linn, R., & Nagy, W. (Eds.), Spoken words. Hillsdale, N.J.: Erlbaum, in press.

Hall, W. S., Nagy, W. E., & Nottenburg, G. Situational variation in the use of internal state words (Tech. Rep. No. 212). Urbana: University of Illinois, Center for the Study of Reading, August 1981.

Hall, W. S., & Tirre, W. C. The communicative environment of young children: Social class, ethnic, and situational differences (Tech. Rep. No. 125). Urbana: University of Illinois, Center for the Study of Reading, May 1979.

Harris, A. J., & Jacobson, M. D. Basic reading vocabularies. New York: Macmillan, 1972.

Horn, E. The commonest words in the spoken vocabulary of children up to and including six years of age. 24th Yearbook of the National Society for the Study of Education (Part I, Chapter 7), 1925.

Horn, E. A basic writing vocabulary. University of Iowa Monographs in Education, First Series, No. 4, 1926.

Kintsch, W., & Vipond, D. Reading comprehension and readability in educational practice and psychological theory. In Lars-Goran Nilsson (Ed.), Proceedings of the Conference on Memory. Hillsdale, N.J.: Erlbaum, 1979.

Klare, G. R. A second look at the validity of readability formulas. Journal of Reading Behavior, 1976, 8, 129-152.

Krantz, L. L. The author's word list. Minneapolis: Curriculum Research Co., 1945.

Lobel, A. Frog and Toad all year. New York: Harper & Row, 1976.

Lorge, I. Word lists as background for communication. Teachers College Record, 1944, 45, 543-552.

Lorge, I. & Chall, J. Estimating the size of vocabularies of children and adults: An analysis of methodological issues. The Journal of Experimental Education, 1963, 52, 147-157.

McLaughlin, G. H. Proposals for British readability measures. Third International Reading Symposium (Eds.: A.L. Brown, & J. Downing) Cassell, 1968.

Shuy, R. Learning to talk like teachers. Language Arts, 1981, 58, 168-174.

Spache, G. D. A new readability formula for primary-grade reading materials. Elementary School Journal, 1953, 53, 410-413.

Spache, G. D. Good reading for poor readers (revised tenth edition). Champaign, Ill.: Garrard, 1978.

Stone, C. R. A graded vocabulary for primary reading. St. Louis, Missouri: Webster Publishing Co., 1936.

Stone, C. R. Measuring difficulty of primary reading material: A constructive criticism of Spache's measure. Elementary School Journal, 1956, 57, 36-41.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. Journalism Quarterly, 1953, 30, 415.

Thorndike, E. L. The teacher's word book. New York: Teachers College, Columbia University, 1921.

Thorndike, E. L. A Teacher's Word Book of 20,000 Words. New York: Bureau of Publications, Teachers College, Columbia University, 1932.

Thorndike, E. L., & Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Bureau of Publications, Teachers College, Columbia University, 1944.

Tidyman, W. F. A Survey of the writing vocabularies of public school children in Connecticut. U.S. Bureau of Education, Teacher's Leaflet No. 15, November 1921.

Footnote 40

[1] Criticisms have been leveled at readability formulas or their misuse from many quarters. For critiques see Bruce, Rubin, & Starr (1981), Davison et al. (1980), Gilliland (1972), Kintsch & Vipond (1979), McLaughlin (1968), and Taylor (1953).

Table 1

Total Number of Familiar Word Types

with Absolute Frequencies $\geq$ 45

Class	Race	
	Black	White
Middle	212 (G_1)	213 (G_2)
Working	178 (G_3)	194 (G_4)

Total number of word types per group = 732

Table 2

Intersections Between Categorical Groups'

732 Most Frequently Spoken Words

	Categorical Groups	Intersections[1] Core Vocabulary	Distinctive to 1st , not 2nd	Distinctive to 2nd, not 1st
Class	Middle-class, Working-Class ($G_1 + G_2$), ($G_3 + G_4$)	707	182[2]	196[2]
Race	Black, White ($G_1 + G_3$), ($G_2 + G_4$)	688	203[2]	179[2]

[1]Represents the number of words within the 732 most frequently spoken words per group that is in the intersection of the comparison groups.

[2]Represents the number of words within the 732 most frequently spoken words per group that is unique to the group. This number will be used for our analysis.

Note: The total vocabulary for each categorical group is greater than 732 because we have combined the original groups.

Table 3

Number of Readability Formula Words Appearing In Distinctive Vocabularies of Class and Race Groups

Group	Spache 1040	Dale 769	Spache added 365
Middle-class ($G_1 + G_2$)	75	62	19
Working-class ($G_3 + G_4$)	63	47	23
Black ($G_1 + G_3$)	63	51	18
White ($G_2 + G_4$)	81	53	34

Total number of most frequently spoken words per group = 732.

Table 4

Intersection of Each Word List with the Core Vocabulary of Most Frequently Spoken Words by Categorical Groups

Class

Word List	Intersection with Core Vocabulary=707	Distinctive to Middle class	Distinctive to Working class	No Intersection
Spache 1040	50%	7%	6%	37%
Dale 769	61%	8%	6%	25%
Spache Added 365	23%	5%	6%	66%

Race

Word List	Intersection with Core Vocabulary=688	Distinctive to Black children	Distinctive to White children	No Intersection
Spache 1040	49%	6%	8%	37%
Dale 769	61%	7%	7%	25%
Spache Added 365	20%	5%	9%	66%

Table 5

Dale 769 words distinctively used by middle-class and working-class children

Class	Words
Middle ($G_1 + G_2$)	afraid, ago, along, animal, arm, bee, bell, board, butterfly, captain, cent, children, choose, clear, company, corner, cover, double, dream, dust, either, except, feed, fly, follow, gone, grass, hall, heavy, hide, its, knew, land, laugh, letter, mark, moon, music, near, page, past, quick, roof, sea, short, skin(ny), sky, soon, spot, star, step, straight, surprise, sweet, teach, though, town, until, warm, without, wonder, year (62)
Working ($G_3 + G_4$)	across, basket, beside(s), bottom, carry, Chinese, circle, city, clock, cook, corn, cost, cross, cry, die, dress, drive, fair, fill, fruit, hang, hundred, instead, laid, lay, neck, none, pay, pan, present, quarter, race, ring, sand, seat, seen, self, sheep, shop, size, tie, tongue, uncle, weak, wild, wood, yard (47)

Table 6

Spache 1040 words distinctively used by White and Black children

Race	Words
Blacks ($G_1 + G_3$)	able, ago, alone, balloon, basket, board, boot, breath, butterfly, cake, captain, carrot, children, circle, city, clock, coat, corner, dress, drive, either, feed, follow, frog, grass, king, knew, lady, laid, land, loud, matter, mud, music, must, pan, past, person, potato, present, promise, quiet, ran, sad, scream, seat, seen, shop, short, sister, size, skip, sky, spill, sun, surprise, sweet, teach, ugly, uncle, upstairs, wake, wonder (63)
Whites ($G_2 + G_4$)	afraid, air, airplane, also, angry, animal, arm, bee, bell, beside(s), best, bother(ing), broken, brush, build, cage, clown, company, cry, dream, dust, each, elephant, fill, flower, fruit, giant, half, hall, heavy, hang, hop, idea, instead, its, key, letter, machine, magic, mark, near, one, pack, park, pay, penny, pie, pot, quick, race, rest, roof, rope, sand, scratch, sea, secret, shot, snap, spot, star, straight, sheep, supper, swallow, swing, threw, tiger, tight, tooth, town, trick, turkey, until, warm, wind, wolf, wood, yard, year, zoo (81)

Table 7

Comparisons of the Words Included on the Dale 769 and Spache 1040 Lists from Distinctive Vocabularies Across Class

List	Proportion from Middle-class Distinctive Vocabulary	Z-value	Significance
Dale 769	.569	1.82	$p \leq .05$
Spache 1040	.544	1.46	N.S.
Intersection of Dale 769 and Spache 1040	.583	1.99	$p \leq .05$
Spache only	.452	-.377	N.S.
Dale only	.462	-.1443	N.S.

Note: Expected Probability of Middle-class distinctive vocabulary words: .482

Table 8

Comparisons of the Words Included on the Dale 769

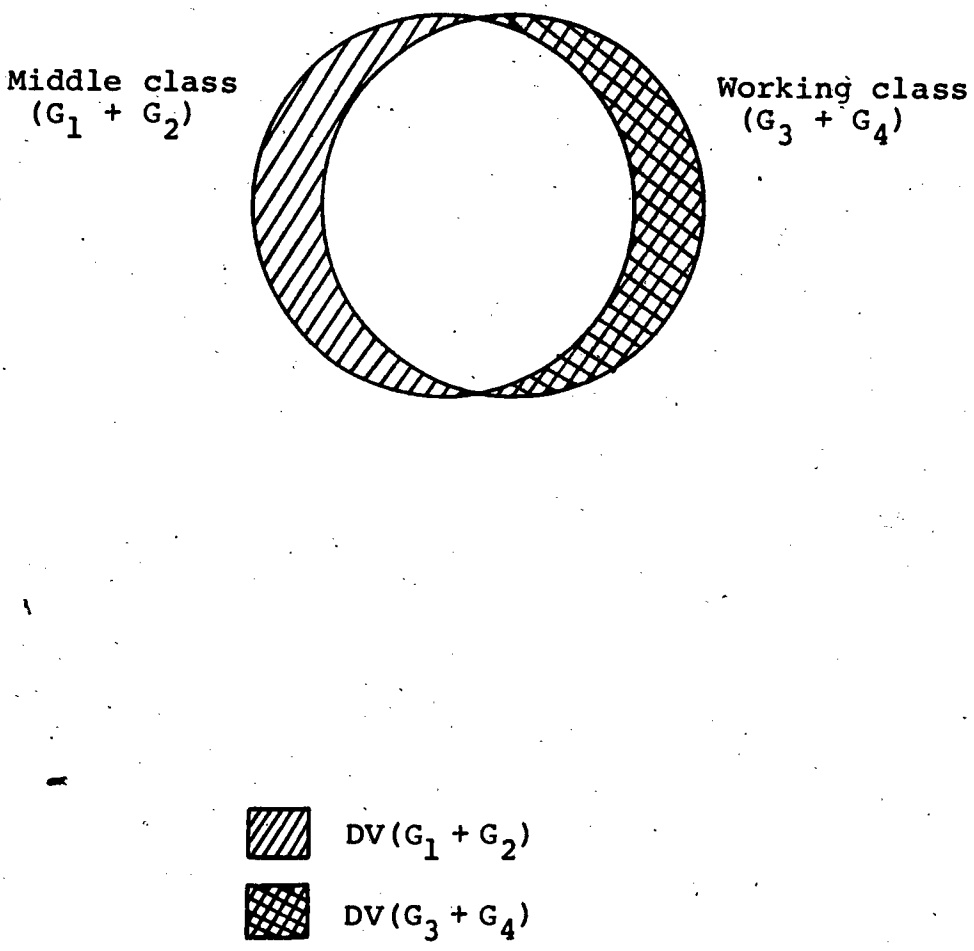
and Spache 1040 Lists from Distinctive Vocabularies Across Race

List	Proportion from White Distinctive Vocabulary	Z-value	Significance
Dale 769	.510	.838	N.S.
Spache 1040	.563	2.26	$p \leq .05$
Intersection of Dale 769 and Spache 1040	.511	.814	N.S.
Spache only	.654	2.68	$p \leq .01$
Dale only	.500	.218	N.S.

Note: Expected Probability of White distinctive vocabulary words: .469

Figure Caption

Figure 1. Distinctive vocabularies in two intersecting sets.


Middle class
(G_1 + G_2)
Working class
(G_3 + G_4)
DV(G_1 + G_2)
DV(G_3 + G_4)