
2-1-1991

The Dolch List Revisited - An Analysis of Pupil Responses Then and Now

Robert E. Leibert
University of Missouri

Follow this and additional works at: https://scholarworks.wmich.edu/reading_horizons



Part of the Education Commons

Recommended Citation

Leibert, R. E. (1991). The Dolch List Revisited - An Analysis of Pupil Responses Then and Now. *Reading Horizons: A Journal of Literacy and Language Arts*, 31 (3). Retrieved from https://scholarworks.wmich.edu/reading_horizons/vol31/iss3/5

This Article is brought to you for free and open access by the Special Education and Literacy Studies at ScholarWorks at WMU. It has been accepted for inclusion in Reading Horizons: A Journal of Literacy and Language Arts by an authorized editor of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



The Dolch List Revisited – An Analysis of Pupil Responses Then and Now

Robert E. Leibert

A little more than a half century after its introduction, the *Dolch List of 220 Basic Sight Words* (1936) continues to be used in classrooms and cited in reading textbooks. In addition to its instructional application in both classroom and clinic (Fry, 1977), the list is also used for assessment, and published as a test (Dolch, 1942).

To aid in the interpretation of results, Dolch (1948) proposed a rough correlation between word knowledge and text difficulty. He estimated that pupils who could pronounce all 220 words should be able to read books of third grade difficulty. Pupils knowing half of the words should be able to read second grade texts, and those pupils who could identify less than half of the words he considered to be first grade readers. Dolch provided another performance indicator by dividing his list into two groups according to word difficulty. Using the results of a study based on 6,000 pupils, 3,000 in each of grades one and two, he formed an easier and more difficult 110 word grouping (Dolch, 1948).

However, the most specific means cited by reading texts for interpreting pupil performance on the Dolch List is a scale developed by McBroom, Sparrow and Eckstein (1944)

which converts a raw score from the Dolch List to an estimated reading level from preprimer to third grade; i.e., 0–75, preprimer; 76–120, primer; 121–170, first reader; 171–210, second reader; above 210, third reader and above (Zintz, 1981; Zintz and Maggart, 1988).

Purpose

One reason that this versatile list remains relevant is that it still reasonably represents the vocabulary of primary materials (Palmer, 1986). Interpretive data, on the other hand, have not been revalidated. The purpose of this investigation was to provide more recent information about the performance of pupils on the Dolch Test, and to provide some observations to assist teachers in interpreting the result of pupil responses to these words.

The following questions guided the data analysis:

- 1. What was the average number of words mastered at each reader level for the pupils in the study?*
- 2. How does the McBroom-Sparrow-Eckstein (MSE) scale compare with the average score of pupils in this study?*
- 3. How does the difficulty of words for the current population compare with the subdivision of the Dolch List?*
- 4. What are the most common trends for incorrect responses to words on the Dolch List?*

Background

A study (Neff and Leibert, 1981), in the spring of 1981, provided a means for collecting performance data on the Dolch Words. Responding to the need to improve the learning of urban youth, several principals in a Midwestern

urban school district decided to incorporate the learning of basic information in a game-like environment. In common with many other urban schools, achievement in this district was below national norms and the project schools performed below the district average. The staff designed games and activities which appeared in school yard graphics and in-school activities. Common games like hop-scotch were revised so information such as the names of the states and the Dolch Words were worked into the games. Special school recognition was accorded children for improving their mastery of the targeted information. After the project had been in effect for more than a year, a study was conducted to determine the efficacy of the incidental practice. This report focused upon the examination of pupil responses for the three grades in this investigation.

Procedure

Pupils in the original study were selected from grades two through four in five comparable urban schools. Twenty pupils were randomly drawn from each grade level and from each school for a total of 100 pupils each from grades two and three. Four pupils moved during the data collection in grade four, resulting in a total sample of 296 pupils. Testing was performed individually by a team of graduate students who used a common system for administering the lists and recording responses. As each list of 20 words was completed any word missed or omitted was presented for a second trial. Reading level was defined as the placement of each pupil in the district reading series by the classroom teacher. At the time of the study the 1974 edition of the Houghton Mifflin Series was the basic instructional program.

For the current investigation, incorrect responses were recorded and tabulated for each word. The resulting data allowed comparisons to be made with existing estimates of

word list difficulty. Pooled responses permitted a study of the variability of incorrect responses as well as determining, for this population, the ranking of word difficulty.

Research question 1: What was the average number of words mastered at each reader level for the pupils in the study? Pupils in this investigation were being instructed in basals ranging from preprimer to the fifth reader level. The average number of correct responses by reader level was computed for each grade.

Table 1

Average number of words pronounced correctly on the first trial by grade and by reader level.

	PP	P	1	2	3	4	5
2	79	129	188	197	214		
3		135	166	204	211	218	
4			140	198	208	217	218

Scores at each grade level show incremental growth in word mastery for each level of reading. This mastery is associated with both grade and reader level. Greater variability of scores is observed among pupils reading at the primer and first reader level than pupils at levels two through four.

Research question 2: How does the McBroom-Sparrow-Eckstein (MSE) scale compare with the average score of pupils in this study? The MSE data are reported as ranges for each grade level. Average scores for the current sample were compared with the top of the score range for the MSE scale for each reader level. Pupils in both studies demonstrated almost complete mastery of the list words by

second reader level, 90 and 95 percent respectively. Pupils in the current study who received instruction at the preprimer through first reader levels evidenced slightly higher average scores than the top score for each range of the MSE scale. Pupils assigned to levels two and three performed slightly lower than the top scores reported on the MSE scale. That is, pupils reading at the beginning levels in the current study appeared to know more Dolch Words than the pupils in the MSE sample. For example, the MSE scores are reported as a continuous scale so scores in the range of 76 to 120 would be estimated to be equivalent to the primer level. Using the current data this means that a considerable number of pupils who were placed at the preprimer through first reader levels in the Houghton Mifflin series would have been estimated as reading a level higher by the MSE data.

Table 2

Comparison between the average scores of the current population and the upper score of the McBroom, Sparrow & Eckstein scale by reader level.

	PP	P	1	2	3	4	5
C	79	131	176	199	210	217	218
P	75	120	170	210	210+		

C = Current Data - Combined scores for all three grades.

P = Published Data - High end of range reported in the MSE scale.

Research question 3: How does the difficulty of words for the current population compare with the subdivision of the Dolch List? By arranging the words in the current study according to their difficulty, it was possible to compare the easiest 110 words for this urban group with the list reported by Dolch (1948). This comparison indicated close

agreement in word difficulty then and now. These two lists differed by only 15 words or by about 14 percent.

Another observation made while analyzing the results was that pupils were able to correct a number of their first trial responses. For example, fourth grade pupils made a total of 1305 incorrect responses on the initial trial, but corrected 461 or about 35 percent of these responses on the second trial. Directions for administering the Dolch List (Zintz, 1981) indicate that only immediate self corrections are counted as correct.

	PP	P	1	2	3	4	5
2	02	08	17	23	50		
3		08	11	34	35	100	
4			13	29	50	69	100

Table 3 confirms that pupils corrected words at each level tested, indicating that the initial score represents a lower estimate of actual word knowledge. Improved performance because of a second trial increases in significance as reading level increases. A second trial also shows that some of the words were still not correctly identified by pupils reading in 4th and 5th grade books.

Research question 4: What are the most common trends for incorrect responses to words on the Dolch List? Some words proved to be exceptionally prone to error even for pupils reading at levels above third grade. Does this mean that certain words are more difficult to learn, or are

some other factors involved? One possibility is that some words are not encountered in written materials until later levels. A partial answer to the question was explored by addressing the relationship between the degree of word mastery and the level where five of the most difficult words on the Dolch List were introduced by the Houghton Mifflin program (Durr, 1974).

Table 4

The relationship of the grade level of word introduction to the percent of word mastery for five Dolch list words

Word	Grade introduced	Book level		
		1	2	3
FAR	2	40	55	82
FULL	3	57	58	89
GAVE	2	40	57	74
THANK	P	43	62	84
WERE	1	43	36	50

The data shown in Table 4 do not provide conclusive evidence to link mastery with the level of word introduction. These data also support improvement in mastery associated with higher reading levels which is consistent with the data from Table 1. The word *were*, for the pupils in this study, was exceptionally difficult. Mastery is slow and apparently neither strongly connected with either level of introduction nor continued practice.

Incorrect responses were also examined for trends. As would be expected, words missed by a sizable number of pupils produced a variety of responses. Nevertheless,

many words produced a high number of similar responses. A listing of the most common incorrect responses for the five most difficult words provides an example of this observation.

Table 5

Most common response for the five most difficult words

<u>List word</u>	<u>Pupil response</u>	<u>%</u>
full	fall	58
far	fare/fair	72
were	where	82
thank	think	88
gave	give	88

Except for the word *were*, the most difficult word for this population, the other four responses involved a vowel shift. This type of response accounted for most errors including those made by pupils reading at the fourth and fifth reader levels. Further inspection of responses revealed that *were* was also the most frequently given incorrect response for *where*, another word on the Dolch List. Considering the total list, there were several notable exceptions to the vowel shift response. *Must* was the most frequent response for *much* (74%), while *himself* was often identified as *myself* (58%), *who* was substituted for *how* (49%), *way* for *why* (60%), and *lunch* for *laugh* (39%).

Summary and conclusions

Responses by 296 urban pupils in grades two through four provided the data to make comparisons with past indi-

cators of the Dolch List difficulty, and to provide insights about pupils responses to the list words.

The relative difficulty of the Dolch Words for pupils today and about 50 years ago was accomplished by comparing the easiest 110 words for the current pupils with a similar listing provided by Dolch (1948). These two listings differed by only fifteen words indicating that the relative difficulty of the Dolch List was highly similar for a current sample of urban pupils.

A second comparison was made by equating the average word mastery for the three grade levels for each reader level with the top score at each level of the McBroom-Sparrow-Eckstein (MSE) scale. Overall, pupils in both studies showed a command of the Dolch Words by the end of the second reader level. However, if the MSE scale had been used to estimate reading levels for the current population, a sizable number of pupils would have been placed a book level higher than their assigned levels. This is not surprising considering the increase in basal difficulty during the intervening years. Difference in results may also be related to variations in the manner in which teachers assigned pupils to books. The current data suggest that the MSE scale may overestimate reading level by as much as a book level for pupils reading from the preprimer through second reader levels. This is striking considering that the pupils in this study score below national norms on standardized tests of reading achievement.

As might be expected there was a variability among the incorrect responses for each word, but for about half of the words there was one predominant incorrect response. A vowel shift was evident in four of the five difficult words used as examples, as well as for many of the other words

for which a single word dominated the responses given. In a few cases the most frequent response involved some form of substitution, as the *st* for *ch* in *much*. With the possible exception of the word *why*, responses could not be attributed to words immediately preceding or following the stimulus word. Rather, the overall consistency of incorrect responses for any given word intimates that decoding knowledge alone may not account for these responses. The response of *give* for *gave*, for example, suggests that the answer may be triggered by some other factor. Perhaps because the responses often bear great similarity to the stimulus word, some combination of visual and auditory cues may explain these responses in a manner comparable to the auditory explanation for spelling substitutions offered by Read (1975).

Because of the relatively high level of response correction, recording second trial responses should be considered in any screening process using the Dolch List. While the scores on the MSE scale are based on first trial scores, assessment to discover actual word knowledge should include second trial information.

Frequency of errors also suggests some cautions about teaching words on the Dolch List. It is clear that for this urban population of pupils, certain words were more difficult than others to master. While these are considered basic words, it is also evidence that it takes several years for pupils to master them. While the comparison of overall mastery of words for this population was similar to that done years ago, it is possible that word difficulty may vary from population to population. Oral language and types of text used for reading instruction and practice may have an effect on the relative difficulty of the Dolch Words. At the very least, the observations about word difficulty and response

variability suggest caution about inferring decoding knowledge from word errors. Knowledge about local population trends with the list should prove helpful in using the Dolch List more effectively. Finally, developing a performance data-base appears to be a useful project for schools or clinics that wish to use a word list like the Dolch 220 for assessment and instruction.

References

- Dolch, E.W. (1936). A basic sight vocabulary. *Elementary School Journal*, 36, 456-460.
- Dolch, E.W. (1942). *The Dolch basic sight word test*. Champaign IL: Garrard.
- Dolch, E.W. (1948). *Problems in reading*. Champaign IL: Garrard.
- Durr, W.K., et al. (1974). *Houghton Mifflin Reading Program*. Boston: Houghton-Mifflin.
- Fry, E.B. (1977). *Elementary reading instruction*. New York: McGraw-Hill.
- McBroom, M., Sparrow, J., & Eckstein, C. (1944). *A scale for determining a child's reader level*. Iowa City: Bureau of Publications, Extension Division, University of Iowa.
- Neff, F.W., & Leibert, R.E. (1981). *Student knowledge of Dolch words in play and learn schools and comparison schools: Final report*. The University of Missouri-Kansas City.
- Palmer, B.C. (1986). Is the Dolch list of 220 basic sight words still relevant? *Reading Improvement*, 23, 227-230.
- Read, C. (1975). *Children's categorization of speech sounds in English*. NCTE Research Report No. 17. Urbana IL: National Council of Teachers of English.
- Zintz, M.V. (1981). *Corrective reading*, 4. Dubuque IA: William C. Brown.
- Zintz, M.V., & Maggart, Z.R. (1989). *The reading process; the teacher and the learner*. Dubuque IA: William C. Brown.

Robert E. Leibert is a faculty member in the School of Education at the University of Missouri, in Kansas City Missouri.