DOCUMENT RESUME

ED 093 486 PS 007 343

AUTHOR Rozin, Paul; And Others
TITLE Do Children Understand the Basic Relationship Between Speech and Writing? The Mow-Motorcycle Test.
SPONS AGENCY National Institutes of Health (DHEW), Bethesda, Md.; National Science Foundation, Washington, D.C.
PUB DATE 15 Apr 74
NOTE 16p.; Revised edition

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Caucasians; Disadvantaged Youth; *Early Childhood Education; Grade 1; Grade 2; Kindergarten; *Miscue Analysis; Negroes; *Reading Diagnosis; *Reading Readiness; *Reading Tests; Speech Skills; Suburban Youth; Written Language

ABSTRACT

School children (N=218) who have not yet attained moderate reading fluency were tested for their awareness of a fundamental relationship between our writing system and speech: that the sounds of speech are represented in writing. Children were shown a long and short word written on a card (e.g., mow and motorcycle), and asked which word corresponded to a spoken word (e.g., mow). The word choices were always grossly different in length, so that a nonreader could perform perfectly if the relationship between written and spoken length was understood. Children were also asked about their basis for responding. Most inner-city kindergarteners in a reasonably representative sample did not perform well on this test. A majority of suburban kindergarteners and inner city first and second graders performed well, but many did not. Controls suggested that failure on this test cannot be attributed to the specific form of presentation of the materials or to misunderstanding of the question being asked. (Author)

DO CHILDREN UNDERSTAND THE BASIC RELATIONSHIP BETWEEN SPEECH AND WRITING?

THE MOW-MOTORCYCLE TEST

Paul Rozin, Beth Bressman, and Mark Taft

University of Pennsylvania

September 11, 1973

Abstract

School children who have not yet attained moderate reading fluency were tested for their awareness of a fundamental relationship between our writing system and speech: that the sounds of speech are represented in writing. Children were shown a long and short word written on a card (e.g., mow and motorcycle), and asked which word corresponded to a spoken word (e.g., mow). The word choices were always grossly different in length, so that a non-reader could perform perfectly if the relationship between written and spoken length was understood. Children were also asked about their basis for responding.

Most inner-city kindergarteners in a reasonably representative sample did not perform well on this test. A majority of suburban kindergarteners and inner city first and second graders performed well, but many did not. Controls suggested that failure on this test cannot be attributed to the specific form of presentation of the materials or to misunderstanding of the question being asked.

DO CHILDREN UNDERSTAND THE BASIC RELATIONSHIP BETWEEN SPEECH AND WRITING? THE MOW-MOTORCYCLE TEST[1]

Most existing writing systems, including the English alphabetic system, are ultimately based on a visual "code" that tracks the sound-stream of speech. In the course of learning to read, a child may come to a realization of this relationship at some point. On the other hand, it is logically possible that a child may show some progress in reading without a clear realization of the fundamental sound-orthography relationship, as for example, if he is taught by the "whole word" method. The abstractness of the sound correspondences of alphabetic elements (Liberman, et. al., 1967; Savin, 1972; Mattingly, 1972; Gleitman and Rozin, 1973) makes this possibility quite plausible.

In the course of developing a syllable-oriented reading curriculum (Gleitman and Rozin, 1973), which makes the sound-tracking aspect of our orthography more apparent, we became interested in determining the extent to which children in the early grades grasp this fundamental fact about reading. We developed an objective test of the child's knowledge of sound-tracking, by determining whether he understood that longer written words generally take longer to say. Children from kindergarten to second grade were shown pairs of long words (9-11 letter, e.g., MOTORCYCLE) and short words, (3-4 letter, e.g., MOW). For each pair, each word was read by E ("One of these words is MOW, the other is MOTORCYCLE") and then E asked the child which written word corresponded to one of the spoken words (e.g., "which word is MOW?"). A set of eight such pairs was followed by a brief interview to determine how the child made his choices.

The primary purpose of this paper is to report the incidence of understanding of the sound-orthography relationship as tapped by the "mow-motorcycle test". Some further studies to assist in the interpretation of this data are also reported.

Method

Subjects. Eight groups of kindergarten to second grade children were tested for a total of 218 Ss in all. The groups are listed and described in the left hand portion of Table 2. One group of kindergarteners (S-K) was drawn from a suburban school that is located in a middle to upper-middle class neighborhood and is almost entirely white. This group was tested towards the end of the school year. The other groups were taken from the kindergarten, first grade and second grade classes of inner-city schools that are located in a lower class neighborhood and whose student body is almost 100% black. Three of these urban (inner-city) groups (U-K, U-2, and U-2C) were drawn from regular classes and were tested towards the end of the school year. Four others were sampled from children participating in a Project-Follow-Through sponsored program conducted during the subsequent summer session. This involved students from four inner-city schools including the one whose regular classes were tested. Two of these groups contained children who were about to enter the first grade in the fall (U-PK, U-PK-B), and the other two contained children one year older, who had just completed the first grade (U-P1, U-P1-B).

Children were selected as Ss either by taking the first N from alphabetic class lists (where, as in U-2, we sampled from all second grade classes) or by running an entire class. All classes were heterogeneously grouped except U-2C, which was an average reading ability class in its school.

Word-test. This test was administered to six groups (S-K, U-K, U-2, U-2C, U-PK and U-P1). All Ss were individually tested by one of the Es. The child was presented sequentially with nine 5" by 8" cards with two words written on each card, one below the other. Some of the words were presumably familiar (e.g.,

ash, asparagus), some were almost surely unknown (e.g., erg, ergonomic) and some were not words at all (e.g., autt, fezagolomiz). The spacing of letters and the placement of words were identical on all cards. The first card used very simple material as a practise-trial and the child was given the correct response if he failed to give it himself.[2] For each card shown, E said the following: "One of these words says 'mow' (or another short word). Say 'mow'. The other word says 'motorcycle' (or another long word). Say 'motorcycle'. Now point to the word that says 'motorcycle'."[3] E's response was always "O.K.", whether the child pointed to the correct target-word or not.

The word-pairs and the target words are listed in Table 1. As the table indicates, the list was balanced both for word-position and target-selection: the long word was the target half the time; on half of these occasions it was on top. After making his last choice (either "fez" or "fezagolomiz") each S was asked why that word said "fezagolomiz", the target word. He was scored as having given a correct response if he (a) chose correctly, and (b) made some mention of the fact that "fezagolomiz" sounded long, big, etc., and that there was a relationship between its written and spoken length. Credit was not given if he merely said it was longer (which is an obvious difference between the visual components and was frequently noted, even when the child had been guessing randomly). When the answer did not meet the established criterion, E probed further. "Could this word (pointing to "fez") say 'fezagolomiz'?" If the child answered no, he was asked "Why not?" He was scored correct if he made some association between temporal and spatial elements.

Box-test. The purpose of this test, administered individually to the Ss in groups U-PK-B and U-P1-B, was to exclude the possibility that the letters themselves may have been a distracting influence. The Ss might have been trying so hard to decode the individual letter symbols that they lost sight of the more general concept of tracking the sound stream. We therefore tried to tap the

child's ability to deal with this concept in a less distracting manner. Two boxes were placed in front of the child: the shorter one (5" long x 2" wide x .5" high) on his right, and the longer one (12" long x 2" wide x .5" high) on his left. The boxes had removeable tops. Inside the shorter box was placed the word "gub" in large plastic letters of about the same height and depth as the box. In the larger box, E placed the word "calotemin". The child was asked to repeat each word. The boxes were designed to clearly demonstrate that long words could only fit in the long box and that short words fit perfectly in the short box. The covers were replaced and the child was asked to point to the box which contained the word "gub". He was corrected if necessary, S was told to close his eyes while E rearranged the letters to make new words. E then said, "One of these boxes has the word 'mow', say 'mow', the other box has the word 'motorcycle'. E's response to both correct and incorrect answers was "O.K.". This procedure was followed for the eight word pairs shown in Table 2. They were presented in the same order. After the practice-trial, the child never saw the letters, and only saw the covered boxes.

Following the child's last choice, (either the long or short box), he was asked why he thought "fezagolomiz", the target word, was in that box. He was scored correct if he in some way stated that long words fit or go in long boxes. Simply saying "because the box is longer" was not sufficient. E questioned the child further by pointing to this shorter box and asking if "fezagolomiz" could fit inside it. If S responded "no" he was asked "why not?" To get credit, S had to convey the idea that a long word could not fit in a small space.

On both word-test and box-test, eight correct choices would represent a perfect score, while four would be the score expect from random guessing. We established a strict success criterion as a score of ≥ 7 as well as an acceptable reason, and a weak criterion of ≥ 6 correct responses without including the reason.

Training procedure. It is conceivable that the child might have understood

the relevant principle, but had not demonstrated his ability because he did not see its applicability in the test context. Therefore, a short training protocol was developed in an attempt to teach the child how the critical concept applies to the test-situation. Although this protocol clearly explicates the basic relationships needed to perform the test, it is probably too brief to teach the relationship de novo, though more than adequate to remind or orient children who already had the basic idea. It quickly reviewed visual discrimination of length, auditory discrimination of length and the integration of these two factors. Practice was given, showing the child how to apply this knowledge to the format of the test (either word or box), followed immediately by a MOW-MOTORCYCLE re-test. Only children who failed to score at least six on the initial test served as subjects (from Summer Urban Group U-PK, U-PK-B, U-P1 and U-P1-B) and training and retesting occurred from a few days to a few weeks after the initial test. Nineteen subjects were given this training procedure, while nineteen others were merely retested to control for "spontaneous" improvement. Children were given the training procedure appropriate to the test (Box or Word) they originally took, and were retested on the same type of test.

The details of this procedure were as follows. In all examples, correction was provided if a child guessed incorrectly. The child was first shown a card with the words "cup" and "cafeteria" printed on it and asked to "point to the word that is longer" (or if necessary, to the "word that has more letters.") This is continued with three further examples (e.g., tan and tangerine). Children originally tested by the box-test were shown the two boxes and asked to point to the longer one.

The next step was an attempt to emphasize discrimination of temporal length of spoken words. E said: "Some words take longer to say than other words. Listen. Fififififi takes longer to say than fi (similarly for two other examples). Now you're going to do it. Which word takes longer to say, nah nah nah nah nah

or nah? Similarly for two others). Which is shorter to say? Kah or Kahkahkah-kahkah?" (And again for two others). The same procedure was used for box or word groups.

The next stage was an attempt to integrate visual and temporal components. The specific procedure depended upon whether the child previously took the word-test or the box-test. If the former, E said: "Words that take longer to say have more letters. The longer a word sounds the longer it looks. Listen--which word sounds longer--cup or cafeteria?" E then showed the child a card with the two words on it and asked which words look longer. After this the child was informed that "one of the words says 'cup' the other 'cafeteria'," and asked to "point out the word that says 'cafeteria'." Similarly for three other pairs, always correcting errors if required.

If tested by the box-test, E said: "Words that take longer to say have more letters. If a word has a lot of letters, it must fit in a long box. A long word fits in a long box--a short word fits in a short box." "Listen. Which word sounds longer--cup or cafeteria?" "Which box would the word cafeteria fit in?" "The longer word fits in the longer box." Similarly for three other pairs. Note again, that throughout this teaching procedure, correction was provided for any response, if necessary.

After the training procedure was completed, the child was given the same form of the test (word or box) which he took originally.

Results

Not surprisingly, there is a close agreement between number of correct choices and appropriate reason offered for the choice. The mean number of correct choices for children who gave appropriate reasons was 7.5. The corresponding figure for those whose reason was inadequate was 5.6 ($P < .001$, χ^2 based on 6x2 distribution of score range [3-8] vs. adequate or no reason. $\chi^2=32.1$, $df=5$). Less than 4% of Ss scoring 5 or below gave adequate reasons.

The results of the MOW-MOTORCYCLE word test are shown in Table 2. Note that

while many suburban kindergarteners (43%) meet our strict criterion (≥ 7 correct with adequate reason) though not given explicit reading instruction, very few urban kindergarteners do (8 or 11%). Further, note that scores improve from kindergarten through second grade in the urban children (since almost all the suburban children acquire basic reading skills by the middle of the first grade, only suburban kindergarten children were tested), but that many urban second-graders (76% in one class, 40% in the other)[4] still fail to meet our strict criterion. The difference between the scores of urban and suburban kindergarteners is highly significant ($p<.001$, Mann Whitney U Test, $U = 176$, $Z = 3.77$), as is the improvement from kindergarten (8% meeting strict criterion) to second grade (60% meeting strict criterion) in urban children ($p<.001$, 2x2 χ^2 based on urban-suburban vs. meeting strict criterion or not. Group 2-U-C was not included in this comparison as it came from a different school than the other urban groups).

As Table 2 indicates, the box-test does not lead to better performance than the word-test. In fact, there is a weak effect in the opposite direction ($p<.05$, $F = 4.39$, $df = 1/96$), which indicates that attempts to decode the words cannot explain the poor performance on the test. As expected, post-first grade children performed better than post-kindergarten on box or word tests ($p<.001$, $F = 22.9$, $df = 1/96$). There was no significant interaction between grade and type of test ($F = .62$, $df = 1/96$).

The results of the training procedure are presented in Table 3. As this table indicates, there is some improvement from first to second test but this improvement is only slightly (and insignificantly) greater following explicit instruction. For Ss, the mean improvement in score following teaching was 1.63, compared to an increase of 1.33 in retested controls ($p>.10$, $t = .16$, $df = 35$). Given the relatively small N, the Box-Word and kindergarten-first grade groups have been combined, yielding 2 larger groups, teaching ($N = 19$) and no teaching ($N = 18$). The t test is based on the difference scores [post - pre] of individuals in these groups). Most important, only three of the

nineteen children given explicit training on the task achieved our strict criterion. The majority of "failures" evidently did not result from a simple failure to understand the task instructions.

DISCUSSION

It appears that a substantial number of children in inner-city early grades do not comprehend the basic nature of our writing system, even at the end of one year of reading instruction with a phonics-oriented method. The exact proportion of such children cannot be estimated from our sample, but it is clear that there is a problem here for many children, and that it extends through the second grade for some. The results of the box-test and the explicit training argue against various artifactual interpretations and suggest that the effect is real.

We do not know why so many children fail to learn the basic sound-tracking principle, or why suburban kindergarteners do so much better than their inner-city age-mates. The urban-suburban difference cannot be easily explained in terms of different kindergarten curricula, for, in fact, the inner-city curricula were more explicitly oriented to reading skills.

Common sense would hold that it might be useful for a child to grasp the nature of the writing system before delving into its detailed specifics (letter-phoneme mappings). It appears that partial mastery of the details does not guarantee appreciation of the basic system. It might be worthwhile to determine whether children are better able to understand the trees (and their relationships) if they are first helped to see the forest. To this end one might develop a small curriculum (perhaps an expanded form of our teaching procedure reported here) explicitly aimed at this goal. Another alternative is to design a reading curriculum around the idea of teaching "sound-tracking" independent of our alphabetic (phoneme-based) system which is so difficult to comprehend. The details of the English writing system could then be gradually introduced

after the child has achieved fluency in using a writing system in which the tracking relations between visual code and sound-stream is more transparent. We are attempting to do just this with an introductory reading program based on the syllable as a fundamental unit (see Gleitman and Rozin, 1973).

References

Gleitman, L. R. and Rozin, P. Teaching reading by use of a syllabary. Reading Research Quarterly, 1973, in press.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. Perception of the speech code. Psychol. Rev., 1967, 74, 431-457.

Mattingly, I. G. Reading, the linguistic process, and linguistic awareness. In, Kavanagh, J. F. and Mattingly, I. G. (Eds.) Language by Ear and by Eye. Cambridge: The M.I.T. Press, 1972.

Savin, H. B. What the child knows about speech when he starts to learn to read. In, Kavanagh, J. F. and Mattingly, I. G. (Eds.) Language by Ear and by Eye. Cambridge: The M.I.T. Press, 1972.

Footnotes

1. We thank Lila Gleitman for her invaluable comments and suggestions on this work at all stages, from inception to final paper, and Henry Gleitman for many valuable suggestions on both the experiments and manuscript, as well as statistical advice. We also thank Bonnie Snyder for her assistance in data collection, and the Philadelphia and Lower Merion School Systems for their cooperation. Supported by NIH Grant 23505-01 and NSF Grant GB 8013. Send requests for reprints to Paul Rozin, University of Pennsylvania, Philadelphia, Pennsylvania, 19174.

2. The practice trial for groups S-K, U-K, U-2C, and U-2 consisted of the presentation of the words "I" and "A". ("One of these is 'A', and one is 'I'. Which one is 'I'?") This example was in no way related to the association of visual and temporal cues with regard to word-length and might thus have served to confuse Ss on later trials. For this reason, groups U-PK, U-PK-B, U-P1 and U-P1-B were presented with "A" and "AAA" for their practice trial, thus receiving the desired contrast between shorter and longer words from the outset.

3. In the first group tested (2-UC), which was in a different urban school than 2-U and K-U, and in a somewhat poorer neighborhood, the child was not required to repeat each word presented by E. In all other groups, the procedure was as indicated.

4. The disparity in performance between the two urban grade groups is probably attributable to the differences in the population in the two school districts.

TABLE 1

Stimulus List for word-test*

Trial	Stimulus on top	Stimulus on bottom
1	motorcycle*	mow (1)
2	ash* (1)	asparagus
3	ter	terminality (1)*
4	ubiquitous	ub (pronounced yoob; (1)*
5	sick*	sychedelic (1)
6	automobile (1)*	autt
7	ergonomic (1)	erg*
8	fez (1)	fezagolomiz*

* In each word-pair, the word followed by (1) was read first, the other word was read second. The word marked by the asterisk was the target-word. Actual stimuli were typed in capital letters.

TABLE 2

Performance on word-test and box test

Groups	N	Procedure	Mean Scores	Percent who meet strict criterion*	Percent who meet weak criterion**
Tested at end of school year					
Suburban kindergarten (S-K)	35	Word-test	6.48	43	74
Urban kindergarten (U-K)	24	"	4.62	8	21
Urban second grade (U-2)	30	"	7.20	60	87
[3]Urban second grade (U-2C)	25	"	6.48	24	72
Tested during summer "Follow-through"					
Urban post-kindergarten					
Group U-PK	27	Word-test	4.92	11	26
Group U-PK-B	26	Box test	4.23	8	23
Urban post-first grade					
Group U-P1	25	Word-test	6.68	48	68
Group U-P1-B	26	Box test	5.88	27	62

* At least 7 out of 8 correct choices, with adequate reason.

** Scoring at least 6 correct choices.

TABLE 3

Effect of Instruction on Performance

Group	N	Mean number correct		Number meeting strict criterion on second test*
		1st test	2nd test	
Post-Kindergarteners				
Instructed	12	3.92	5.33	2
Controls	11	3.45	4.55	1
Post-First-Graders				
Instructed	7	4.14	6.14	1
Controls	7	4.00	5.63	2

*At least 7 out of 8 correct choice, with adequate reason.