Journal of Educational Psychology
1992, Vol. 84, No. 3, 353–363

0022-0663/92/$3.00

Predicting Progress in Beginning Reading: Dynamic Assessment of Phonemic Awareness

Janet E. Spector
College of Education, University of Maine

This study investigates the ability of a dynamic measure of phonemic awareness to predict progress in beginning reading. Thirty-eight kindergarteners who were nonreaders were assessed in the fall on receptive vocabulary, letter and word recognition, invented spelling, phoneme segmentation, phoneme deletion, and dynamic phoneme segmentation. They were retested near the end of the school year on reading, spelling, and phonemic awareness. The results of the multiple-regression analyses supported the hypothesis that dynamic assessment enhances the predictive utility of a phonemic awareness measure. Performance on dynamic phoneme segmentation was the best predictor of end-of-year reading scores and of growth in phonemic awareness. The study demonstrates the applicability of principles of dynamic assessment to the measurement of phonemic awareness and provides further evidence regarding the relationship between phonemic awareness and reading acquisition.

This study investigates the ability of a dynamic measure of phonemic awareness to predict progress in beginning reading. The dynamic approach was compared with a more conventional static approach to assessing phonemic awareness. I hypothesized that the dynamic measure would more accurately predict progress in beginning reading than would a static measure. The study was influenced by theory and research on two questions: (a) the relationship between phonemic awareness and reading acquisition and (b) the effectiveness of dynamic versus static assessment.

Phonemic Awareness and Reading Acquisition

One of the most consistent relationships to emerge from the past decade of research on reading is the relationship between phonemic awareness and reading acquisition (see reviews by Adams, 1990; Ehri, 1979; Golinkoff, 1978; Jorm & Share, 1983; Liberman & Shankweiler, 1985; Wagner & Torgesen, 1987; Williams, 1984). Although there is some variation across studies in the definition of *phonemic awareness*, the term generally is used to denote the ability to perceive spoken words as a sequence of sounds (Lewkowicz, 1980). Phonemic awareness has been measured by performance on a wide range of tasks, including rhyming (e.g., Calfee, Chapman, & Venezky, 1972); isolating beginning, medial, and ending sounds (e.g., Williams, 1980); breaking words into their component sounds (e.g., Fox & Routh, 1975; Goldstein, 1976; Helfgott, 1976); saying words with target sounds deleted (e.g., Bruce, 1964; Rosner & Simon, 1971); and producing invented spellings (e.g., Mann, Tobin, & Wilson, 1987; Morris & Perney, 1984; Read, 1971).

The results of both correlational and experimental studies generally have indicated that students who enter reading instruction unable to perform phonemic awareness tasks experience less success in reading than students who score high in phonemic awareness when instruction commences (e.g., Bradley & Bryant, 1983; Calfee, Lindamood, & Lindamood, 1973; Juel, 1988; Liberman, Shankweiler, Fischer, & Carter, 1974; Perfetti, Beck, Bell, & Hughes, 1987; Share, Jorm, Maclean, & Matthews, 1984; Stanovich, Cunningham, & Cramer, 1984; Tunmer & Nesdale, 1985; Vellutino & Scanlon, 1987). From a theoretical perspective, this finding is consistent with models of reading acquisition that emphasize the critical role of insight into the alphabetic principle during the initial stages of learning how to read (e.g., Elkonin, 1973; Gough & Hillinger, 1980; Liberman, 1973; Perfetti, 1985; Rozin & Gleitman, 1977). That is, children must realize that letters stand for sounds and when combined, yield words.

Previous research on phonemic awareness has identified tasks that appear to be reliable and valid predictors of reading progress. Yopp (1988), for example, identified two tasks from a battery of 10 phonemic awareness tests that together accounted for 58% of the variance in scores on a learning test designed to simulate the learning-to-read process. The capacity of phonemic awareness tests to predict aspects of reading progress suggests that such tests could be used to identify children who would benefit from instructional intervention (Share et al., 1984).

A possible obstacle to practical application, however, is the unfamiliarity and complexity of many phonemic awareness tasks. Whereas poor performance on a phoneme segmentation task might indicate low phonemic awareness, it might also reflect the child's lack of understanding of task requirements or difficulty in meeting ancillary task demands. For example, some researchers have asked children to count the number of sounds in a spoken word and then to use a pencil or dowel to make a corresponding number of taps (e.g., Liberman et al.,

This article is based in part on a paper presented in April 1990 at the annual meeting of the American Educational Research Association in Boston.

I am indebted to Pamela Schutz for enabling data collection; the principal, teachers, and students of Pendleton Street School for their participation in the study; and Ted Coladarci and the anonymous reviewers for their helpful comments and suggestions.

Correspondence concerning this article should be addressed to Janet E. Spector, College of Education, Shibles Hall, University of Maine, Orono, Maine 04469-0155.

1974). Poor performance might indicate difficulty in segmenting the spoken word, but it might also reflect difficulty in counting the sounds, attending to task instructions, or maintaining the one-to-one correspondence between sounds and taps. Thus, a limitation of conventional tests of phonemic awareness is that they yield too many false negatives, that is, students who are unable to perform the experimental task but who actually possess (or could easily acquire) the ability that the task is designed to measure. Dynamic assessment is an approach that might be useful in eliminating these false negatives.

Dynamic Assessment

Dynamic assessment is a general term used to describe a variety of evaluation approaches that emphasize the processes, in addition to the products, of assessment (see reviews by Campione, 1989; Lidz, 1981, 1987). These approaches include learning potential assessment (e.g., Budoff, 1987a, 1987b; Feuerstein, 1979); testing-the-limits procedures (Carlson & Wiedl, 1978, 1979); mediated assessment (e.g., Bransford, Delclos, Vye, Burns, & Hasselbring, 1987; Burns, 1985); and assisted learning and transfer (e.g., Campione, Brown, Ferrara, Jones, & Steinberg, 1985).

To obtain information about responsiveness to instruction, dynamic approaches require the interaction between tester and student. When a student has difficulty solving a problem or answering a question, the tester attempts to move the student from failure to success by modifying the format, providing additional examples or trials, modeling an appropriate strategy for success, or offering increasingly more direct cues or prompts. The intensity of the dynamic intervention varies across approaches, ranging from brief, standardized prompts (e.g., testing the limits) to complex, individualized probes (e.g., learning potential assessment).

Advocates of dynamic assessment characterize traditional tests as static: A student's failure to solve a problem or answer a question is viewed merely as an indicator of where the student stands on some underlying ability scale. However, although the student's performance may accurately reflect previous achievement, it is not necessarily predictive of how the student will perform in the future (Campione & Brown, 1985; Feuerstein, 1979). That is, although a student might lack the skill in question at the time of testing, with minimal instruction the same student might realize substantial gains. In addition, because static approaches typically provide little feedback or practice prior to testing, failure often reflects children's lack of understanding of the instructions more than it reflects their ability to perform the task.

Most dynamic assessment approaches are linked theoretically to Vygotsky's (1935/1978) views of learning and development. In his work, Vygotsky emphasized the critical role in learning of children's social interaction with adults. The adult acts as an expert model in demonstrating how to solve a problem or perform a particular task. Over time, the child takes on increasing responsibility for task performance, and the adult provides help only when needed. Eventually, the child internalizes the problem-solving routine and is able to perform the task independently.

Traditional tests assess only two states: unaided success and failure. That is, the child either answers a question correctly, without prompts or cues from the examiner, or the child is considered to fail the item. From a Vygotskian perspective, however, the child may be somewhere in between these two states: unable to perform the task independently but able to achieve success with minimal assistance. Vygotsky (1935/1978) considered this in-between state to be the zone of proximal development: "the distance between the actual developmental level as determined by independent problem-solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 85).

Consider, for example, two children who earn the same low score on a traditional phoneme segmentation test that requires students to articulate each of the sounds that they hear in a word. With minimal instruction, the first child experiences significant growth in performance, whereas the second child shows little improvement. Although the two children received the same score initially, different degrees of future success might be predicted for the two children in tasks that presume phonemic awareness.

In the present study, I developed a dynamic measure that assesses the ability of kindergarten children to perform a phonemic awareness task when given supportive prompts and cues. On the basis of previous research on dynamic assessment (e.g., Budoff, 1987a, 1987b; Campione et al., 1985; Carlson & Wiedl, 1979; Embretson, 1987; Ferrara, Brown, & Campione, 1986), I hypothesized that the dynamic measure would more accurately predict future progress in reading than would comparable static measures. Also considered was the ability of the dynamic measure to forecast growth in phonemic awareness. I expected the dynamic measure to be a better predictor of phonemic awareness at the end of the kindergarten year than the traditional static measures of phonemic awareness.

Method

Subjects

In October, parents of students enrolled in a public school kindergarten program received a letter from the school requesting permission for their child to participate in a study of beginning reading. Of 74 students, 52 (70%) returned signed consent forms. The school, located in a small New England town, serves a predominantly middle-class, White population. All students spoke English as a first language. Given the reciprocal relationship between phonemic awareness and reading acquisition, only those children who were unable to read any words on the San Diego Quick Assessment List (LaPray & Ross, 1969) in the fall of the kindergarten year were included in the study. In addition, students who were absent during any of the fall or spring testing sessions were excluded from the analysis. This resulted in a final sample of 38 students. The mean age of children at the beginning of the study was 5 years, 11 months.

Procedure

Data for the study were collected in November and May of the kindergarten year. Testing was conducted at the school and was completed in three 15- to 20-min sessions. Children were assessed individually in the fall on letter and word recognition (San Diego Quick Assessment List; LaPray & Ross, 1969), invented spelling (Mann, Tobin, & Wilson, 1987), phoneme segmentation (Yopp, 1988), phoneme deletion (Bruce, 1964), and dynamic phoneme segmentation (with a measure designed for this study). Students were assessed again in the spring on all measures except dynamic phoneme segmentation. The fall tests were administered in two sessions approximately 1 week apart. In Session 1, students completed phoneme segmentation, phoneme deletion, and invented spelling. Session 2 covered dynamic assessment and word recognition. Spring testing required only one session. The order in which the tasks were administered was fixed. In the fall, dynamic assessment was completed last to eliminate the possibility that the instructional prompts and cues would influence performance on the other measures. This design feature, however, precluded investigation of possible order effects, a limitation that must be considered in interpreting the results.

Measures

I used four phonemic awareness tasks: phoneme segmentation, dynamic phoneme segmentation, phoneme deletion, and invented spelling. The Yopp–Singer phoneme segmentation (Yopp, 1988) and Bruce Word Analysis Test (1964) were selected on the basis of Yopp's study of the reliability and validity of 10 phonemic awareness measures. In Yopp's study, the results of a factor analysis indicated that phonemic awareness comprises two factors: simple phonemic awareness and complex phonemic awareness. In that study the Yopp–Singer phoneme segmentation test provided the most reliable measure of simple phonemic awareness, and the Bruce Word Analysis Test provided the most reliable measure of complex phonemic awareness ($rs = .95$ and $.92$, respectively). Together, the two measures accounted for 58% of the variance in the rate at which students learned to decode new words. Invented spelling was included as an additional indicator of phonemic awareness on the basis of previous research linking invented spelling with phonemic awareness (Chomsky, 1971, 1979; Ehri, 1989; Ehri & Wilce, 1980, 1987a, 1987b; Liberman, Rubin, Duques, & Carlisle, 1985; Mann et al., 1987; Morris & Perney, 1984; Read, 1971, 1986). Each of these measures is described below.

Phoneme segmentation. On the Yopp–Singer phoneme segmentation test, subjects were asked to pronounce, in order, each of the sounds in a word. Four examples were given at the beginning of the test to familiarize the student with the task. Feedback was given on the four trial items, but no feedback was given for the remaining 22 words on the test. The present administration procedure differs somewhat from that of Yopp (1988), who provided feedback regarding the correct response after each item. In the present study, feedback was withheld to distinguish more clearly between the static phoneme segmentation test and the dynamic version of the same task.

Using Yopp's (1988) approach, I scored the test for number of words segmented correctly. The maximum score possible was 22. The majority of students ($n = 25$) scored zero in the fall on this measure. To alleviate the substantial floor effect and to make the scaling more similar to that used during dynamic assessment, I awarded points for each sound in a word that was segmented correctly. The maximum score possible for the revised measure was 57, and the correlation between the original and the revised fall measures was .88 ($p < .001$).

Not surprisingly the two scoring procedures yielded equivalent conclusions regarding the benefits of dynamic assessment. Therefore, I chose to report fall phoneme segmentation results only for the revised measure. However, because of the manner in which students' responses were recorded in the spring, it was not possible to score spring data for number of sounds correctly segmented. Spring phoneme segmentation, therefore, is reported as number of words segmented correctly.

Dynamic assessment. The dynamic assessment procedure paralleled the Yopp–Singer task but provided corrective feedback and increasingly supportive prompts and cues when children were unable to segment a word correctly. The tester used the following series of prompts each time a child was unable to segment a word:

- Prompt 1: pronouncing the target word slowly;
- Prompt 2: asking the child to identify the first sound of the word;
- Prompt 3: cuing the child with the first sound;
- Prompt 4: cuing the child with the number of sounds in the word;
- Prompt 5: modeling segmentation using pennies placed in squares to represent the number of sounds in the word;
- Prompt 6: modeling segmentation as above, but working hand over hand with the child while pronouncing the segments;
- Prompt 7: repeating Prompt 6.

Each child attempted a maximum of 12 items: four consonant–vowel (CV) words (i.e., *say, pie, we, two*), four vowel–consonant (VC) words (i.e., *age, eat, egg, if*), and four consonant–vowel–consonant (CVC) words (i.e., *leg, feet, page, rice*). The intent was to select familiar words that included a range of vowel and consonant sounds. One student who, on the first 2 items, was unable to produce a correct segmentation after all seven prompts (including the last three prompts, which required imitation only) was discontinued on the task and received a total score of zero. (The dynamic assessment script is presented in the Appendix.)

Scores on dynamic phoneme segmentation indicate the degree of independence that the child achieved in performing the segmentation task. Each item was scored as follows: 6 = correct response with no prompts required; 5 = correct response after Prompt 1; 4 = correct response after Prompt 2; 3 = correct response after Prompt 3; 2 = correct response after Prompt 4; 1 = correct response after Prompt 5; and 0 = no correct response.

Performance on the last two imitation trials (i.e., Prompts 6 and 7) did not contribute to scores. Initially, I created a second dynamic measure that assigned points for successful second and third trials at imitation. This measure was scored on an 8-point scale, with a score of 8 indicating correct response without prompting and a score of 1 indicating correct response on the third imitation attempt. Because the two measures yielded equivalent results on all analyses reported below, I include here only the results for the 6-point scale. The maximum score possible on the dynamic measure was 72.

The present approach to dynamic assessment has both similarities to and differences from other dynamic approaches. These similarities and differences can be described in terms of three general dimensions identified by Campione (1989): focus, interaction, and target (see Campione, 1989, for examples of studies that differ on these dimensions).

Focus refers to the way in which the researcher operationalizes the processes and effects of dynamic assessment. Whereas most dynamic approaches involve a test–train–test paradigm, investigators differ with respect to the measures they derive. Three types of measures predominate: (a) measures of change from pretest to posttest, (b) measures of posttest performance, and (c) measures of performance during the dynamic assessment itself. This study followed the third approach. The dynamic phoneme segmentation measure indicated the degree of independence the student achieved during dynamic assessment. I also created a second measure that reflected change in performance from the static to the dynamic condition. I obtained this measure by regressing the fall dynamic measure on static phoneme segmentation. The resulting residualized gain reflects variability in dynamic scores that is independent of static performance.

Interaction describes the nature of the interaction between the examiner and the student during dynamic assessment. The interaction is either standardized (i.e., all students receive the same prompts or cues) or individualized. The latter condition is clinical; the examiner addresses specific obstacles to success that student responses reveal. This study followed the standardized approach. I used a fixed set of prompts that were administered to all students in the same order. As can be seen above, success on initial prompts reflects need for minimal adult intervention, whereas success on later prompts indicates need for more extensive adult help in performing the task.

In Campione's taxonomy, *target* refers to the nature of the skills that are tested. Early work on dynamic assessment (e.g., Budoff, 1967; Feuerstein, 1979) focused on domain-general skills associated with cognitive ability. Examples of tasks that have been studied include Raven's (1956) Coloured Progressive Matrices, Kohs's (1923) Block Design Task, and the Representational Stencil Design Task (Arthur, 1947). In the past few years, however, researchers have advocated use of academically relevant tasks (Bransford et al., 1987; Campione, 1989; Campione & Brown, 1987). Consistent with this recommendation, the present study has as its target a domain-specific skill that influences reading acquisition.

Phoneme deletion. Bruce's (1964) Word Analysis Test assesses the student's ability to delete phonemes from words. The test comprises 33 words that vary in length (30 are one-syllable words and 3 are multisyllable words). Bruce initially selected the words for their familiarity. The test requires children to say the word that remains after a particular sound is deleted from a stimulus word (e.g., "What word would be left if /j/ were taken away from the beginning of *jam*?"). The positions in the words of the sound to be deleted are equally divided between beginning, middle, and end. In the present study, children completed four sample items prior to the test. I provided corrective feedback on sample items only. Given the length of the test and its difficulty for beginning kindergarteners, I discontinued testing if students were unsuccessful on the first 15 words and recorded a score of zero. Performance on the test was scored as number of words correct.

Invented spelling. Invented spelling requires students to represent in writing the sound structure of words. To assess this ability, I used a measure developed by Mann et al. (1987) in their study of the relationship between invented spelling and phonemic awareness. The test requires students to write the following words: *red, name, bed, lady, fish, men, boat, girl, color, angry, thank you, people, dog,* and *boy*. Mann et al. selected these words for their familiarity to kindergarten students and because each presents the opportunity for students to make "errors" that are characteristic of children who have not yet learned how to read (e.g., "nam" for *name*, "ppl" for *people*). Each item contains one or more of the following elements: (a) a letter name within the word (e.g., "d" in *lady*) or (b) a short vowel, a nasal, a liquid, or a consonant represented by a digraph (e.g., "e" in *bed*, "nk" in *thank you*, "r" and "l" in *girl*, "sh" in *fish*). Following Mann et al.'s instructions, I encouraged students to invent spellings for words that they did not know and to write any of the sounds in a word if they did not know the whole word.

Mann et al. (1987) scored student performance on the spelling task for phonetic equivalence. That is, they rated each response on a scale to reflect its correspondence to the phonetic structure of the word: 4 = correct spelling; 3 = preconventional response that accurately represents the phonological structure of the word; 2 = preconventional response of two or more letters that reflects at least part of the phonological structure of the word; 1 = single-letter response that represents the initial sound of the word; 0.5 = single-letter response that represents a sound in the word other than the initial consonant; and 0 = no response or a response that does not represent any sounds in the word.

Because of the focus in this study on phonemic awareness, I eliminated the distinction between conventionally correct and phonetically equivalent responses. That is, I reduced the scale from 0–4 to 0–3 by awarding 3 points to both correct spellings and preconventional responses that accurately represented the phonemic structure of the word. The highest possible score on this measure was 42. The correlation between scores on the 0–3 scale used in this study and scores on Mann et al.'s (1987) original 0–4 scale was .99. Although Mann et al. did not report information about the reliability of their measure, I assessed the reliability of the scoring system by asking two judges to rate the same 10 protocols. Agreement between the two scorers was 90%.

Word recognition. The San Diego Quick Assessment List (LaPray & Ross, 1969) includes a series of graded word lists beginning at the preprimer level. The words on the lists are typical of words that students encounter in instructional materials at different reading levels. As is the case with most informal reading inventories, the authors provided no information regarding reliability. Evidence of content validity is derived by examining the overlap between words on the test and words typically presented in basal reading material at different levels.

Students in the present study were asked to read word lists arranged in order of increasing difficulty. Testing was discontinued when a student was unable to read any of the words on a particular list. In practice, only the preprimer, primer, and first-grade lists were used because of the limited reading skills of the kindergarten subjects. The test was scored for number correct. The maximum score possible on each word list was 10, and the highest possible score across all three lists was 30.

Prior to reading the graded word lists, students also completed a letter recognition task that required them to state the names of 10 letters. Scores on letter recognition reflected the number of letters that were correctly named.

Peabody Picture Vocabulary Test–Revised (PPVT-R). The students' PPVT-R (Dunn & Dunn, 1981) results were available in school records. The test had been administered by school personnel in the month prior to the reading and phonemic awareness testing. The PPVT-R, a test of receptive vocabulary, has frequently been used in research on phonemic awareness as a measure of verbal ability. I included it in the present study to control for individual differences in verbal ability. At ages 5 years through 6 years, 11 months (the age levels at which subjects in the present study were tested), split-half reliability coefficients ranged from .73 to .84. Alternate-forms reliability is estimated to range from .78 to .80 and from .60 to .67 on immediate test–retest and delayed test–retest, respectively (Dunn & Dunn, 1981).

Results

The data were analyzed to investigate the predictive ability of the dynamic phoneme segmentation measure. I used ordinary least squares regression to identify the fall measure that best predicted spring phonemic awareness and spring reading. This approach required separate analyses, one for each of the four dependent measures administered in the spring: phoneme segmentation, phoneme deletion, invented spelling, and word recognition. The independent variables in each analysis were the following: fall performance on phoneme segmentation (static and dynamic measures), phoneme deletion, invented spelling, and PPVT-R. I also used stepwise regression to identify the contribution of the dynamic measure, beyond that explained by general ability (i.e., PPVT-R) and the static measure of phoneme segmentation. Alpha was set at .05 throughout.

Descriptive Statistics

Means and standard deviations for each measure are presented in Table 1. In the fall, most of the children had considerable difficulty with the phoneme deletion and static phonemic segmentation tasks. Mean scores on phoneme deletion indicated that the average student responded correctly to only 1 or 2 words; many students did not respond correctly to any of the items on the task. Examination of fall scores on the phoneme segmentation task suggests that although the average student correctly segmented only 1 or 2 words on the test, the typical student was able to segment correctly at least one sound for half of the words. Similarly, although most students failed to spell correctly any words on the spelling test in the fall, the majority of students were able to produce letters to represent at least one of the sounds in the words.

Examination of fall and spring scores reveals that student performance improved on all the phonemic awareness tasks over the course of the year. The increase in accuracy on phoneme segmentation was considerably greater than that for phoneme deletion. On average, students progressed from 12% accuracy in the number of words segmented correctly in the fall to 50% accuracy in the spring on phoneme segmentation, whereas accuracy on phoneme deletion only increased from 5% to 12%.

Scores on word recognition also improved from fall to spring. In the fall, none of the students were able to identify any of the words on the preprimer, primer, or first-grade lists, whereas by spring, only one third of the students still scored zero on word recognition. The average student was able to identify 30% of the words on the preprimer list, and the highest scoring students were able to read all the words on the preprimer list and 1 or 2 words on the primer list.

Correlations Among Measures

Whereas students differed in their ability to perform each of the phonemic awareness tasks, there were significant correlations among all four of the fall phonemic awareness measures (i.e., phoneme deletion, phoneme segmentation, invented spelling, and dynamic phoneme segmentation). The phonemic measures generally correlated more highly with each other than they did with the PPVT-R, a pattern that supports the construct validity of the four phonemic awareness measures (see Table 2).

I also examined the strength of the correlations among the four phonemic awareness tasks. The correlations between the dynamic measure and each of the three static measures were descriptively greater than the correlations among the three static measures themselves. That is, the dynamic measure appears to have more in common with each of the three static measures of phonemic awareness than the three static measures have in common with each other.

Given the focus of the study on prediction, I was particularly interested in examining the simple correlations between the fall and spring measures. There were significant positive correlations between fall and spring phonemic awareness measures and between phonemic awareness scores and spring reading performance (see Table 2). Among all fall measures—including the PPVT-R—dynamic phoneme segmentation was the variable that correlated most highly with spring reading scores and with spring performance on the phonemic awareness measures.

A secondary question regarding fall-to-spring correlations concerned the relationship between benefit from the dynamic assessment condition and growth in word recognition. That is, Was improvement in performance from static to dynamic assessment associated with gains in reading during the second half of kindergarten? I addressed this question using a measure of residualized gain that I created by regressing dynamic phoneme segmentation on static phoneme segmentation scores. The correlation between residualized gain and spring word recognition indicated that gain from static to dynamic assessment was, in fact, positively associated with growth in word recognition ($r = .43$, $p < .01$). In other words, students who showed the most growth in word recognition from fall to spring tended to be those who were helped most by the prompts and cues provided during dynamic assessment. This result is not surprising, given the high correlation between residualized gain and the original dynamic segmentation measure ($r = .78$, $p < .01$).

Among the three static phonemic awareness measures (i.e., phoneme segmentation, phoneme deletion, and invented spelling), spring scores correlated more highly with spring reading than did fall scores. This result might be explained by the temporal factor: One generally expects two measures given at the same time to correlate more highly with each other

Table 1
Mean Scores on Fall and Spring Measures

Measure	Maximum score	Fall *M*	Fall *SD*	Spring *M*	Spring *SD*
PPVT-R	—	100.89	13.81	—	—
Phoneme segmentation					
No. words correct	22	2.68	4.75	11.39	8.18
No. sounds correct	57	13.71	13.28	—	—
Phoneme deletion	33	1.58	2.96	3.55	5.68
Invented spelling	42	10.67	10.92	26.65	11.76
San Diego Quick Assessment List					
Letter recognition	10	8.50	2.47	9.53	1.39
Word recognition	30	0.00	0.00	3.00	3.36
Dynamic phoneme segmentation	56	23.24	19.16	—	—

Note. $N = 38$. PPVT-R = Peabody Picture Vocabulary Test-Revised. Dashes indicate not available.

Table 2
Intercorrelations Among Fall and Spring Measures

Measure	Fall 1	2	3	4	5	Spring 6	7	8	9
Fall									
1. DYN	—	.62**	.43**	.60**	.11	.51**	.66**	.54**	.60**
2. PS		—	.37*	.44**	.09	.31*	.62**	.40*	.38*
3. PD			—	.16	.34*	.34*	.45**	.28	.23
4. IS				—	.10	.41**	.46**	.49**	.55**
5. PPVT-R					—	.51**	−.02	.39*	.27*
Spring									
6. PS						—	.36*	.61**	.46**
7. PD							—	.37*	.67**
8. IS								—	.62**
9. WR									—

Note. $N = 38$. DYN = dynamic phoneme segmentation; PS = phoneme segmentation; PD = phoneme deletion; IS = invented spelling; PPVT-R = Peabody Picture Vocabulary Test–Revised; WR = San Diego Quick Assessment List word recognition.
* $p < .05$, one-tailed. ** $p < .01$, one-tailed.

than measures given at different times. The increase from fall to spring in the magnitude of correlations between phonemic awareness and reading might also result from the reciprocal nature of the relationship between phonemic awareness and reading: Phonemic awareness facilitates reading acquisition, and subsequent progress in reading itself promotes increased phonemic awareness.

However, these simple correlations do not take into account the confounding influences among the phonemic awareness measures. Multiple regression analyses, described below, permitted an estimate of the independent effects of each fall variable, holding constant the effects of all remaining variables.

Multiple Regression

Spring phonemic awareness. To investigate the ability of the dynamic measure to predict phonemic awareness in the spring, each of the spring phonemic awareness variables (i.e., phoneme segmentation, phoneme deletion, and invented spelling) was separately regressed on fall phoneme segmentation (static and dynamic), phoneme deletion, invented spelling, and PPVT-R. I entered fall measures into the regression equation simultaneously. The results of the analysis are listed in Table 3.

As can be seen in Table 3, the dynamic measure accounted for a significant portion of the variance in two of the spring phonemic awareness tests. Along with the PPVT-R, it predicted spring phoneme segmentation. It also forecasted spring phoneme deletion, together with fall static segmentation and deletion. As reflected by the beta values in Table 3, for each standard deviation increase in dynamic segmentation performance, scores on the spring dependent measures increased by approximately one third of a standard deviation.

I also approached the analysis in a stepwise fashion. The purpose of the second analysis was to examine the increment in explained variance associated with the entry of the dynamic measure when PPVT-R and the static version of phoneme segmentation were already in the equation. I found that the dynamic measure accounted for an additional 12% to 14% of the variance in the spring measures, even after entering PPVT-R and the static measure of phoneme segmentation: for spring phoneme segmentation, $F(1, 34) = 9.27$, $p < .01$; for spring phoneme deletion, $F(1, 34) = 9.23$, $p < .01$; and

Table 3
Multiple Regression With Fall Measures Used to Predict Spring Phonemic Awareness Scores

Fall predictor	IS B	IS β	IS t	PD B	PD β	PD t	PS B	PS β	PS t
DYN	.21	.34	1.68	.10	.33	1.84*	.17	.40	2.09*
PPVT-R	.29	.34	2.45**	−.07	−.17	−1.36	.27	.46	3.41**
IS	.24	.22	1.33	.06	.11	0.45	.11	.15	0.95
PD	−.18	−.05	−0.30	.47	.25	1.75*	.03	.01	0.07
PS	.07	.08	0.49	.12	.29	1.92*	−.03	−.05	−0.29
R^2		.45			.56			.49	

Note. $N = 38$. IS = invented spelling; PD = phoneme deletion. PS = phoneme segmentation; DYN = dynamic phoneme segmentation; PPVT-R = Peabody Picture Vocabulary Test–Revised. For all t tests, $df = 32$.
* $p < .05$, one-tailed. ** $p < .01$, one-tailed.

for spring invented spelling, $F(1, 34) = 6.76$, $p < .05$. In each of the analyses, the increment in explained variance corresponds to roughly one third of the total amount of variance accounted for by all three independent variables combined.

Spring reading performance. Two analyses assessed the ability of the dynamic measure to predict spring reading scores. I used the same approach as was reported above for examining spring phonemic awareness. That is, I first entered fall scores into the equation simultaneously to estimate the contribution of each independent variable, holding constant the effects of the other predictors. In this analysis, independent variables were fall phoneme segmentation (static and dynamic versions), phoneme deletion, invented spelling, and PPVT-R. I then used stepwise regression to investigate the unique contribution of the dynamic measure, above and beyond that of general verbal ability (i.e., PPVT-R) and static segmentation performance. In both analyses, the dependent measure was word recognition.

As can be seen in Table 4, performance on the dynamic version of the phoneme segmentation task was the only significant predictor of word recognition performance at the end of kindergarten. For each standard deviation increase in performance on the dynamic measure, end-of-year reading performance increased by approximately one half standard deviation ($\beta = .47$).

The second analysis found that performance on the dynamic measure accounted for an additional 21% of the variance in word recognition scores after entering PPVT-R and the static phoneme segmentation measure, $F(1, 34) = 11.84$, $p < .01$. The increment in explained variance corresponds to roughly one half of the total variance accounted for by all three independent variables combined.

The dynamic measure in the previous analyses indicates degree of independence achieved during dynamic assessment. To further investigate the relationship between responsiveness to dynamic assessment and growth in reading, I repeated the regression of spring word recognition on fall phoneme awareness variables and the PPVT-R, substituting a measure of residualized gain (described previously) for the original dynamic phoneme segmentation variable.

Table 4
Multiple Regression With Fall Measures Used to Predict End-of-Year Word Recognition Scores

	End-of-year word recognition		
Fall measure	B	β	t
Dynamic phoneme segmentation	.08	.47	2.37*
PPVT-R	.05	.22	1.63
Invented spelling	.08	.27	1.64
Phoneme deletion	−.10	−.09	−0.56
Phoneme segmentation	−.01	−.02	−0.10
R^2		.46	

Note. $N = 38$. PPVT-R = Peabody Picture Vocabulary Test–Revised. Word recognition was measured with the San Diego Quick Assessment List word recognition task. For all t tests, $df = 32$.
* $p < .05$, one-tailed.

The results of this analysis were almost indistinguishable from those depicted in Table 4: Gain from static to dynamic assessment was the best predictor of spring word recognition, $t(32) = 2.38$, $p < .01$. Even after entering static phoneme segmentation and the PPVT-R in a stepwise fashion, dynamic gain accounted for an additional 21% of the variance in spring reading scores, $F(1, 34) = 11.84$, $p < .01$, which is approximately one half the variance accounted for by all three measures combined.

Discussion

The results of the present study support the hypothesis that dynamic assessment enhances the predictive utility of a measure of phonemic awareness. Dynamic phoneme segmentation was a better predictor of kindergarten reading progress than any of the three static measures of phonemic awareness: phoneme segmentation, phoneme deletion, and invented spelling. The dynamic measure was also a better predictor of word recognition than the PPVT-R, a measure that is often used to estimate verbal ability. Similarly, the dynamic measure accounted for more variance in spring phonemic awareness than did any of the other fall phonemic awareness measures. These results demonstrate the applicability of the principles of dynamic assessment to measurement of phonemic awareness and add to the ever-growing corpus of research on the relationship between phonemic awareness and reading acquisition.

Interpretive Considerations

Two interpretive issues are considered: (a) the influence of the distribution of word recognition scores on the results reported above and (b) possible explanations for the effectiveness of the dynamic measure.

Distribution of scores on word recognition. In the spring, 13 students were still unable to read any words on the word recognition test. Conceptually, this result is logical; the students who did not progress in word recognition were also the lowest scorers on dynamic segmentation. Their lack of growth in word recognition thus supports the diagnostic value of the dynamic measure.

Statistically, however, the distribution of word recognition scores raises the possibility of spuriously high correlations between word recognition and dynamic phoneme segmentation. To address this concern, I recomputed the correlation between fall dynamic phoneme segmentation and spring word recognition, eliminating from analysis the 13 students who scored zero on word recognition. Although the correlation dropped from .60 to .43 in the reduced sample ($n = 25$), the magnitude of the latter coefficient remains consistent with my original conclusion regarding the positive relationship between dynamic phoneme segmentation and word recognition.

I also replicated the word-recognition regression analysis reported above, using a combined measure of letter and word recognition. The dependent measure in this analysis reflected number correct across both the letter and word recognition subtests of the San Diego Quick Assessment List (the distribution of scores on this variable was approximately normal). Because the majority of subjects were at ceiling on letter recognition—30 students scored 10 (100%) and 6 subjects scored 9 (90%)—the combined measure served only to spread out the scores of the 13 students who scored zero on word recognition. The regression results for the combined letter-word recognition measure were equivalent to those obtained for word recognition: Dynamic phoneme segmentation was the single best predictor of letter-word recognition, $t(32) = 2.67$, $p < .01$, $\beta = .51$, followed by PPVT-R, $t(32) = 2.17$, $p < .05$, $\beta = .28$, and invented spelling, $t(32) = 1.78$, $p < .05$, one-tailed, $\beta = .28$.

In short, the results reported above do not appear to be an artifact of the distribution of scores on word recognition. Even when problematic cases were removed or when alternate versions of a reading recognition measure were used, the dynamic measure consistently emerged as the best predictor of reading progress.

Explanations for dynamic effects. What accounts for the predictive superiority of the dynamic phoneme segmentation measure? Embretson (1987) described three possible goals of dynamic assessment: "(a) improving ability estimates; (b) assessing new constructs, such as modifiability of performance; and (c) improving true ability" (p. 167). Embretson's framework is useful in delineating plausible explanations for the effectiveness of the dynamic assessment measure used in this study. The dynamic measure (a) provided a better estimate of phonemic segmentation ability than did the static measures of phonemic awareness; (b) assessed a different ability than did the static measure (i.e., responsiveness to phonemic awareness training); and (c) increased students' phonemic awareness. I examine below the tenability of each possibility.

The first explanation is that the dynamic measure was simply a more sensitive indicator of phonemic awareness than any of the other three phonemic awareness measures. Examination of the distribution of scores on phoneme deletion provides some support for this interpretation. More specifically, a significant floor effect compromised the ability of this measure to predict spring reading performance. In the fall, many students were unable to respond correctly to any of the items on the task. The dynamic measure had the effect of spreading out the scores of those who scored zero on phoneme deletion.

At the same time, it is unlikely that the predictive strength of the dynamic measure was merely an artifact of the expanded scaling technique. Both invented spelling and fall static phoneme segmentation (scored for number of sounds correctly segmented) were nondichotomous scales that awarded credit for partially correct responses. Constrained variance clearly was not an issue for these two measures, yet dynamic phoneme segmentation exceeded them both in forecasting growth in reading.

A more plausible version of the sensitivity account is that the dynamic measure was simply a "cleaner" (Calfee, 1977, p. 297) measure of phonemic awareness. Analysis of the static conditions suggests the possibility that task demands unrelated to phonemic awareness prevented many students from achieving greater success. For example, on both deletion and static segmentation, some students may have had difficulty maintaining the target word in working memory while relating in order the sounds that they heard. In the dynamic condition, on the other hand, the children were asked to identify the initial sound only and were then asked for each succeeding sound, prompting each response by repeating sounds that the child had already given. The dynamic condition clearly placed fewer demands on working memory than did the static segmentation and deletion tests.

The predictive superiority of the dynamic measure over invented spelling might also be explained by demands placed on ancillary skills. Whereas 80% of the students were able in the fall to identify most of the letters on the San Diego Quick Assessment List letter recognition task, they might have been considerably less knowledgeable regarding the relationship between graphemes and phonemes. Also, they might not have known how to form letters correctly. In other words, the dynamic measure, in contrast to the static measures, might have minimized the contribution to successful performance of abilities other than phonemic awareness. Unfortunately, I did not assess students' knowledge of individual letter–sound correspondences or ask them to write the alphabet, so additional research is necessary to test the validity of this explanation.

The second possible interpretation of the present results is that the dynamic measure assessed a new ability: modifiability of performance. As discussed above, the rationale underlying many approaches to dynamic assessment is the Vygotskian notion of zone of proximal development. Vygotsky (1935/1978) used this term to refer to that area just beyond the child's present capabilities on tests requiring independent performance but within the child's capability when given adult assistance or support.

According to Vygotsky (1935/1978), dynamic measures allow observation of

> those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are in the embryonic stage. These functions could be called the "buds" or "flowers" rather than the fruits of development. The actual developmental level characterizes mental development retrospectively, while the zone of proximal development characterizes mental development prospectively. (pp. 86–87)

In the present study, those students who benefited most from the instructional cues and prompts during the dynamic testing session were also the students who experienced the most growth in word recognition during the kindergarten year. From a Vygotskian perspective, the effectiveness of the dynamic measure can be explained in terms of modifiability of performance. That is, the dynamic measure assessed individual differences in responsiveness to segmentation instruction.

The above interpretation, although attractive, requires further investigation. Although the present results are congruent with this account, the study was not designed specifically to investigate why dynamic measures have greater predictive validity than comparable static measures. In subsequent studies researchers might consider use of structural equation modeling, an approach recommended by Embretson (1987) for testing hypotheses about the locus of dynamic effects. Embretson's design specifications include both fall and spring testing on (a) a dynamic measure, (b) a transfer task (e.g., word recognition), and (c) an irrelevant task (i.e., a task that has elements in common with the dynamic task but that does not tap the construct that the dynamic measure is hypothesized to influence). This design would facilitate more complete evaluation of the modifiability hypothesis.

The third possible explanation for the present results is that participation in dynamic assessment actually changed the child's readiness. That is, in teaching nonsegmenters how to segment, the study effected long-term growth in phonemic awareness and reading. I found this explanation attractive initially, given the increase in number of words segmented correctly from fall to spring. Mean performance on phoneme segmentation rose from 12% accuracy in the fall to 50% accuracy in the spring, whereas accuracy on phoneme deletion only increased from 5% to 12%. On the other hand, the relative improvement in phoneme segmentation over phoneme deletion is explained equally well by the demands of the phoneme deletion and segmentation tasks themselves. As other investigators have reported, phoneme deletion is a more complex skill and one that develops later than phoneme segmentation (e.g., Perfetti et al., 1987; Yopp, 1988). In other words, the greater degree of improvement in segmentation ability can be easily explained by developmental factors. The results do not necessarily imply that participation in dynamic assessment effected long-term improvement in phoneme segmentation ability.

The training effects hypothesis is also called into question by the results of experimental studies that have attempted to teach segmentation skills. First, previous training studies suggest that segmentation training is most effective when it is accompanied by training in blending or spelling (Bradley & Bryant, 1983; Fox & Routh, 1984; Williams, 1980; also see discussions by Adams, 1990, pp. 328–332; Ehri, 1989; and Wagner & Torgesen, 1987). Therefore, the present condition was not optimal for inducing long-term change. Second, studies that have yielded successful results for phonemic awareness training have taken place over longer periods of time with more intensive treatment. Bradley and Bryant (1983), for example, provided 40 training sessions over a 2-year period (once or twice per week). Lundberg, Frost, and Petersen (1988) carried out their phonemic awareness training program over the entire kindergarten year, whereas Fox and Routh (1984) conducted training sessions 4 to 5 days a week (15 to 20 min) for 5 weeks. In contrast, a study that failed to yield a main effect of segmentation-and-blending training was conducted by Treiman and Baron (1983). The training took place over a 4-day period, an amount of time that is more similar to (but still four times greater than) the intervention in this study.

I conclude, therefore, that the sensitivity and modifiability explanations are more tenable than the training effects hypothesis. The sensitivity account suggests that the dynamic measure was a better predictor because it is a cleaner (or truer) measure of phonemic awareness. The modifiability account, which is the Vygotskian interpretation, proposes that, in assessing responsiveness to instruction, the dynamic measure is better able to forecast future growth in segmentation ability. Thus, a goal for subsequent research on dynamic assessment is to disentangle what appear to be equally plausible interpretations of its effectiveness.

Implications for Research and Practice

Clearly, there are unresolved issues surrounding dynamic assessment. For example, How does one assess the reliability of a dynamic measure? By its nature, dynamic assessment induces change in the individual. Such change, in turn, reduces reliability as conceptualized by classical test theory (i.e., consistency or stability of measurement over time and items). In her review of the psychometric challenges posed by dynamic assessment, Embretson (1987) provided some suggestions for assessing the reliability and validity of dynamic measures. More research is needed to document the effectiveness of the recommended procedures and to establish standards for evaluating the technical adequacy of dynamic measures.

With respect to the dynamic measure developed for use in the present study, further investigation is necessary to rule out the possibility that the results were influenced by the order in which the tasks were administered. Subsequent research should also consider repeating the dynamic procedure in the spring to permit more complete analysis of spring-to-spring correlations.

Although the dynamic measure developed for use in this study requires further scrutiny, the study suggests a number of practical applications. One potential application is development of a screening device to identify children in need of instructional intervention to facilitate reading acquisition. Although no test identified in this or any previous study has accounted for enough variance in reading achievement to stand alone as a screening measure, the dynamic measure of phonemic awareness developed for purposes of this study has advantages over previous measures. First, it is a better predictor of reading progress. Second, in allowing the examiner to observe the child's responsiveness to instructional prompts and cues, the measure might be useful in identifying an appropriate instructional approach to develop phonemic awareness in each child. Finally, the present procedure does not require extensive training on the part of the examiner and can be administered in one 15- to 20-min session. For similar reasons, the measure would also be valuable in research on reading and phonemic awareness in which predictive strength and ease of administration are critical.

The results of this study also suggest the benefits of developing dynamic measures for other phonemic awareness tasks, such as invented spelling, or even for the criterion measure, word recognition. Although the present findings support the predictive validity of the dynamic phoneme segmentation measure, it is possible that a dynamic measure of one of the other phonemic awareness measures (e.g., invented spelling) would prove to be an even better predictor of early reading progress.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Arthur, G. (1947). *A point scale of performance tests.* New York: Psychological Corporation.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read—A causal connection. *Nature, 301,* 419–421.

Bransford, J. C., Delclos, V. R., Vye, N. J., Burns, M. S., & Hasselbring, T. S. (1987). State of the art and future directions. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 479–496). New York: Guilford Press.

Bruce, D. J. (1964). An analysis of word sounds by young children. *British Journal of Educational Psychology, 34,* 158–170.

Budoff, M. (1967). Learning potential among institutionalized young adult retardates. *American Journal of Mental Deficiency, 72,* 404–411.

Budoff, M. (1987a). Measures for assessing learning potential. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluation learning potential* (pp. 173–195). New York: Guilford Press.

Budoff, M. (1987b). The validity of learning potential assessment. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 52–81). New York: Guilford Press.

Burns, M. S. (1985). *Comparison of "graduated prompt" and "mediational" dynamic assessment and static assessment with young children* (Tech. Rep. No. 2). Nashville, TN: Vanderbilt University, John F. Kennedy Center for Research on Human Development.

Calfee, R. C. (1977). Assessment of independent reading skills: Basic research and practical applications. In A. S. Reber & D. L. Scarborough (Eds.), *Toward a psychology of reading: Proceedings of the City University of New York conferences* (pp. 289–323). Hillsdale, NJ: Erlbaum.

Calfee, R. C., Chapman, R. S., & Venezky, R. L. (1972). How a child needs to think to learn how to read. In L. Gregg (Ed.), *Cognition in learning and memory* (pp. 139–182). New York: Wiley.

Calfee, R. C., Lindamood, P., & Lindamood, C. (1973). Acoustic-phonetic skills and reading—Kindergarten through twelfth grade. *Journal of Educational Psychology, 64,* 293–298.

Campione, J. C. (1989). Assisted assessment: A taxonomy of approaches and an outline of strengths and weaknesses. *Journal of Learning Disabilities, 22,* 151–165.

Campione, J. C., & Brown, A. L. (1985). *Dynamic assessment: One approach and some initial data* (Tech. Rep. No. 361). Urbana, IL: Center for the Study of Reading. (ERIC Document Reproduction Service No. ED 269 735)

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–115). New York: Guilford Press.

Campione, J. C., Brown, A. L., Ferrara, R. A., Jones, R., & Steinberg, E. (1985). Breakdowns in flexible use of information: Intelligence-related differences in transfer following equivalent learning performance. *Intelligence, 9,* 297–315.

Carlson, J. S., & Wiedl, K. H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency, 82,* 529–564.

Carlson, J. S., & Wiedl, K. H. (1979). Toward a differential testing approach: Testing-the-limits employing the Raven matrices. *Intelligence, 3,* 323–344.

Chomsky, C. (1971). Write first, read later. *Childhood Education, 47,* 296–299.

Chomsky, C. (1979). Approaching reading through invented spelling. In L. B. Resnick & P. A. Weaver (Eds.), *Theory and practice of early reading* (Vol 2, pp. 43–65). Hillsdale, NJ: Erlbaum.

Dunn, L., & Dunn, L. (1981). *Peabody Picture Vocabulary Test-Revised.* Circle Pines, MN: American Guidance Service.

Ehri, L. C. (1979). Linguistic insight: Threshold of reading acquisition. In T. G. Waller & G. E. Mackinnon (Eds.), *Reading research: Advances in theory and practice* (Vol. 1, pp. 63–114). San Diego, CA: Academic Press.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading disability. *Journal of Learning Disabilities, 22,* 356–365.

Ehri, L. C., & Wilce, L. S. (1980). The influence of orthography on readers' conceptualizations of the phonetic structure of words. *Applied Psycholinguistics, 1,* 371–385.

Ehri, L. C., & Wilce, L. S. (1987a). Cipher versus cue reading: An experiment in decoding acquisition. *Journal of Educational Psychology, 79,* 3–13.

Ehri, L. C., & Wilce, L. S. (1987b). Does learning to spell help beginners learn to read words? *Reading Research Quarterly, 18,* 47–65.

Elkonin, D. B. (1973). U.S.S.R. In J. Downing (Ed.), *Comparative reading* (pp. 551–579). New York: Macmillan.

Embretson, S. E. (1987). Toward development of a psychometric approach. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 141–170). New York: Guilford Press.

Ferrara, R. A, Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies in proximal development. *Child Development, 57,* 1087–1099.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The Learning Potential Assessment Device, theory, instruments, and techniques.* Baltimore, MD: University Park Press.

Fox, B., & Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. *Journal of Psycholinguistic Research, 4,* 331–342.

Fox, B., & Routh, D. K. (1984). Phonemic analysis and synthesis as word attack skills: Revisited. *Journal of Educational Psychology, 76,* 1059–1064.

Goldstein, D. M. (1976). Cognitive-linguistic functioning and learning to read in preschoolers. *Journal of Educational Psychology, 68,* 680–688.

Golinkoff, R. M. (1978). Critique: Phonemic awareness skills and reading achievement. In F. B. Murray & J. J. Pikulski (Eds.), *The acquisition of reading: Cognitive, linguistic, and perceptual prerequisites* (pp. 23–41). Baltimore, MD: University Park Press.

Gough, P., & Hillinger, M. L. (1980). Learning to read: An unnatural act. *Bulletin of the Orton Society, 30,* 179–196.

Helfgott, J. A. (1976). Phonemic segmentation and blending skills of kindergarten children: Implications for beginning reading acquisition. *Contemporary Educational Psychology, 1,* 157–169.

Jorm, A. F., & Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics, 4,* 103–147.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437–447.

Kohs, S. C. (1923). *Intelligence measurement.* New York: Macmillan.

LaPray, M., & Ross, R. (1969). The graded word list: Quick gauge of reading ability. *Journal of Reading, 12,* 305–307.

Lewkowicz, N. K. (1980). Phonemic awareness training: What to teach and how to teach it. *Journal of Educational Psychology, 72,* 686–700.

Liberman, I. Y. (1973). Segmentation of the spoken word and reading acquisition. *Bulletin of the Orton Society, 23,* 65–77.

Liberman, I. Y., Rubin, H., Duques, S., & Carlisle, J. (1985). Linguistic abilities and spelling proficiency in kindergarteners and adult poor spellers. In D. B. Gray & J. F. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 163–176). Parkton, MD:

York Press.
Liberman, I. Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education, 6,* 8–17.
Liberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology, 18,* 201–212.
Lidz, C. S. (1981). *Improving assessment of schoolchildren.* San Francisco: Jossey-Bass.
Lidz, C. S. (Ed.). (1987). *Dynamic assessment: An interactional approach to evaluating learning potential.* New York: Guilford Press.
Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23,* 263–284.
Mann, V. A., Tobin, P., & Wilson, R. (1987). Measuring phonological awareness through the invented spellings of kindergarten children. *Merrill-Palmer Quarterly, 33,* 365–391.
Morris, D., & Perney, J. (1984). Developmental spelling as a predictor of first-grade achievement. *The Elementary School Journal, 84,* 441–457.
Perfetti, C. A. (1985). *Reading ability.* New York: Oxford University Press.
Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first-grade children. *Merrill-Palmer Quarterly, 33,* 283–319.
Raven, J. C. (1956). *Coloured progressive matrices.* New York: Psychological Corporation.
Read, C. (1971). Pre-school children's knowledge of English phonology. *Harvard Educational Review, 41,* 1–34.
Read, C. (1986). *Children's creative spellings.* New York: Routledge, Chapman & Hall.
Rosner, J., & Simon, D. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities, 4,* 384–392.
Rozin, P., & Gleitman, L. (1977). The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principle. In A. Reber & D. Scarborough (Eds.), *Toward a psychology of reading* (pp. 55–141). Hillsdale, NJ: Erlbaum.
Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology, 76,* 1309–1324.
Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology, 29,* 175–190.
Treiman, R., & Baron, J. (1983). Phonemic-analysis training helps children benefit from spelling-sound rules. *Memory & Cognition, 11,* 382–389.
Tunmer, W. E., & Nesdale, A. R. (1985). Phonemic segmentation skill and beginning reading. *Journal of Educational Psychology, 77,* 417–427.
Vellutino, F. R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly, 33,* 321–363.
Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds. and Trans.), *Mind in society: The development of higher psychological processes* (pp. 79–91). Cambridge, MA: Harvard University Press. (Original work published 1935).
Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 10,* 192–212.
Williams, J. P. (1980). Teaching decoding with an emphasis on phoneme analysis and phoneme blending. *Journal of Educational Psychology, 72,* 1–15.
Williams, J. P. (1984). Phonemic analysis and how it relates to reading. *Journal of Learning Disabilities, 17,* 240–245.
Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly, 23,* 159–177.

Appendix

Dynamic Assessment Task

Introductory Instructions to Students

"Do you remember the game we played last week with sounds? I said a word and asked you to break the word apart. You had to tell me each sound that you heard in the word. When I said 'old,' you had to say 'o-l-d.' Let's try a few more."

General Procedures

Begin with Item No. 1 ("say") and continue through Item No. 12. Whenever a child is unable to segment a word correctly, move the child through the prompts (in order) until he or she produces the correct segmentation. For example, if on Item No. 4, the child achieves success on Prompt 2, move next to Item No. 5. If the child cannot segment Item No. 5 on the first attempt, move the child through the prompts again, beginning with Prompt 1. Discontinue testing after the first 2 items if a child does not achieve success in segmenting on either Item No. 1 or Item No. 2. Remember, a child is considered successful if he or she responds correctly to any of the prompts (including the 3 prompts that require imitation only).

Prompts

Prompt 1: "Listen while I say the word very slowly." Model slow pronunciation.
"Now can you tell me each sound?"

Prompt 2: "What's the first sound you hear in ________?"
If first sound is correct: "Now can you tell me each of the sounds?"
If incorrect or no response: "Try to tell me just a little bit of the word."
If child still does not isolate first sound, skip Prompts 3 and 4. Go to Prompt 5.

Prompt 3: If child correctly identified first sound but not next sound(s):
"____is the first sound in ________."
"What sound comes next?"
"Now can you tell me each sound?"

Prompt 4: "There are 2 [or 3] sounds in ________.
What are they?"

Prompt 5: "Watch me." Model segmentation of word: Place a token in a square as each sound is spoken, then repeat word as a whole. After demo say the following: "Try to do what I just did."
Score response as correct if child can imitate correct segmentation.

Prompt 6: "Let's try together." Model segmentation of word with child. Work hand-over-hand with child and ask child to pronounce segments along with you.
"Now try to do it yourself. Do what we just did together."

Prompt 7: Model again with child. "Now try again to do it yourself."

Received March 4, 1991
Revision received January 8, 1992
Accepted January 27, 1992 ■