

## DOCUMENT RESUME

ED 372 369

CS 011 789

AUTHOR Baker, Scott K.; Good, Roland  
TITLE Curriculum-Based Measurement Reading with Bilingual Hispanic Students: A Validation Study with Second-Grade Students.  
SPONS AGENCY Department of Education, Washington, DC.  
PUB DATE Apr 94  
CONTRACT 8029D80051-91; H023B00017  
NOTE 35p.; Paper presented at the Annual Meeting of the Council for Exceptional Children/National Training Program for Gifted Education (Denver, CO, April 6-10, 1994).  
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)  
EDRS PRICE MF01/PC02 Plus Postage.  
DESCRIPTORS \*Bilingual Students; Comparative Analysis; \*Curriculum Based Assessment; \*Elementary School Students; Grade 2; \*Hispanic Americans; Primary Education; Reading Achievement; \*Reading Comprehension; \*Reading Diagnosis; Reading Research; Test Reliability; Test Validity

## ABSTRACT

A study investigated the reliability, validity, and sensitivity of Curriculum-Based Measurement (CBM) reading (which relates reading fluency to reading comprehension) with bilingual Hispanic students. For 10 weeks, 50 second-grade bilingual students and 26 students who spoke English only were administered CBM reading measures twice weekly. Reliability of the CBM reading was compared between the groups with respect to the level of student performance and rate of student progress. Evidence for the convergent and discriminant validity of the CBM reading for bilingual Hispanic students was evaluated using criterion measures of English reading and language proficiency. Sensitivity was evaluated by comparing the reading progress of the two groups during the 13-week period. Results indicated that CBM reading was as reliable and valid for bilingual students as for English-only students, and was sensitive to the reading progress of bilingual students. The construct validity of CBM reading as a measure of general reading proficiency in English was supported. (Contains 67 references, 4 tables, and 1 figure of data.) (Author/RS)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

ED 372 369

Curriculum-Based Measurement Reading with Bilingual Hispanic Students:

A Validation Study with Second-Grade Students

Scott K. Baker

Roland Good

University of Oregon

PERMISSION TO REPRODUCE THIS  
MATERIAL HAS BEEN GRANTED BY

S. Baker

TO THE EDUCATIONAL RESOURCES  
INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Running Head: CBM READING WITH BILINGUAL HISPANIC STUDENTS

This research was supported in part by grants No. H023B00017, and 8029D80051-91 from the US Department of Education, Office of Special Education Research. The authors express their appreciation to the students, their parents, and the faculty and staff of the Toppenish School District for their time and effort on this study, and to Mark R. Shinn, Ph.D., for his helpful comments on earlier drafts of this manuscript. Address correspondence regarding this manuscript to Scott K. Baker, Division of Learning and Instructional Leadership, College of Education, University of Oregon, Eugene, OR 97403, (503) 346-3541.

CS011789

## ABSTRACT

This study investigated the reliability, validity, and sensitivity of Curriculum-Based Measurement (CBM) reading with bilingual Hispanic students. For 10 weeks, 50 second-grade bilingual students and 26 students who spoke English only were administered CBM reading measures twice weekly. Reliability of the CBM reading was compared between the groups with respect to the level of student performance and rate of student progress. Evidence for the convergent and discriminant validity of the CBM reading for bilingual Hispanic students was evaluated using criterion measures of English reading and language proficiency. Sensitivity was evaluated by comparing the reading progress of the two groups during the 13-week study. Results indicated that CBM reading was as reliable and valid for bilingual students as for English-only students, and was sensitive to the reading progress of bilingual students. The construct validity of CBM reading as a measure of general reading proficiency in English was supported.

## Curriculum-Based Measurement Reading with Bilingual Hispanic Students:

## A Validation Study with Second-Grade Students

Hispanics currently represent 9% of the total U.S. population. By the year 2020, the Bureau of Census (1992) projects they will surpass African Americans to become the largest minority group in the country (Reddy, 1993). While the white population grew by 6% during the 1980s and the African American population grew by 13%, the Hispanic American population grew by 53%. Low academic achievement, high drop-out rates, and language and cultural differences from other groups, as well as within the Hispanic community itself, challenge public education. So far, public education has experienced only meager success in addressing the educational needs of Hispanic students (Arias, 1986; Casas, Furling, Solberg, & Carranza, 1990; Fradd & Correa, 1989; Millis, Campbell, & Farstrup, 1993; Ortiz, 1986).

Language factors clearly contribute to the challenge. Hispanics represent the largest bilingual group in the country. According to the latest information from the Bureau of Census, 78% of Hispanics speak Spanish in the home. Although many of these individuals also speak English, 39% of Hispanics over the age of 5 reported that they did not speak English very well. Twenty-three percent of Hispanics reported that they lived in linguistically isolated households where no one over the age of 14 spoke English very well. Some do not speak either language well.

For Hispanics perhaps more than any other group, assessment procedures are needed that are sensitive to the local context of the community. Hispanic communities differ dramatically in language dominance, rates of English adoption, public and private language use, socioeconomic status (SES), attitudes toward education, and history of educational opportunities and achievement (Reddy, 1993; Figueroa, 1990; Kanellos, 1993; Laosa, 1975; Luz Reyes, 1992). For example, Mexican Americans, Puerto Ricans, and Cubans make up approximately 77% of Hispanics in the U.S., but additional Hispanic groups, including Nicaraguans, Salvadoreans, Dominicans, Guatemalans, Columbians, and Equadorans (Kanellos, 1993; Reddy, 1993), are numerous

enough to affect the communities in which they live. These Hispanic groups differ in linguistic and cultural ways that are educationally relevant.

### Measuring Reading Proficiency with Bilingual Students

For all students, including Hispanics, the development of English reading skills is a cultural imperative (Deno, 1989), made particularly difficult when English is not a student's first language (Arias, 1986; Millis, et al., 1993; Ortiz, 1986; Willig, 1985). Key to the provision of effective interventions is assessment to (a) identify students experiencing difficulty, (b) determine the most appropriate service to provide, (c) monitor the progress of individuals, and (d) evaluate the effectiveness of interventions (Reschly, 1988). Published norm-referenced tests (PNRTs) of reading frequently are used to assess the reading skills of bilingual Hispanic students (Garcia & Pearson, 1994).

### Published Norm-Referenced Tests of Reading

The limitations of PNRTs for educational decision making are well established (Marston, 1989; Fuchs, D., Fuchs, L., Benowitz, & Berringer, 1987; Galagan, 1985; Jenkins & Pany, 1978). In general, PNRTs are able to provide global information on the standing of a student's achievement relative to a national norm. However, because even the best PNRTs are not aligned with any specific curriculum (Good & Salvia, 1988; Jenkins & Pany, 1978; Shapiro & Derr, 1987), and do not facilitate the establishment of local norms, they cannot provide detailed information on how well students are acquiring the particular skills being taught in their classroom, or on how students compare to other students in their class. Consequently, PNRTs cannot differentiate poor performance due to: (a) learning difficulties, (b) poor instruction, or (c) effective instruction on content not on the test.

PNRTs are not intended for measuring individual student progress (Marston, 1989). At best, PNRTs can be used to assess gain over a long period of time for groups of students (e.g., yearly). They generally have only one or at best two forms, and are time intensive to administer, precluding frequent administration. In addition, PNRTs are not sensitivity to actual changes in student learning. PNRTs generally are designed to be used at multiple grades so large changes in problem

difficulty on adjacent items are necessary. As a result, student skill level must improve considerably to improve their performance by one or two items (Marston, 1989).

These problems with PNRTs are exacerbated when they are used to assess the reading skills of bilingual Hispanic students. The reading skills of bilingual Hispanic students may be systematically underestimated when they are administered PNRTs. Garcia (1991) studied the reading performance of 51 bilingual Hispanic students and 53 white students enrolled in the same 5th and 6th grade classrooms. The Hispanic students' reading test scores seriously underestimated their reading comprehension. Their test performance was adversely affected by their limited prior background knowledge of test topics, their unfamiliarity with vocabulary terms, and their tendency to interpret the test literally. When differences in prior knowledge were controlled statistically, the reading performance of the two groups did not differ.

#### Curriculum-Based Measurement with Bilingual Hispanic Students

Curriculum-Based Measurement (CBM) has the potential to address many of the limitations of PNRTs in making educational decisions with bilingual Hispanic students (Shinn, 1989). On CBM reading measures, students read aloud for 1 minute from passages selected randomly from the general education reading curriculum. The number of words students read correctly in 1 minute provides a measure of reading proficiency. The content validity of CBM reading is important because it addresses the uncertain test-curriculum overlap that plagues PNRTs. In addition, CBM reading consists of short duration fluency measures that are sensitive to small changes in performance. Because they are sensitive, brief, and have many alternate forms, they can be used on a repeated basis to evaluate student progress over time.

The validity of CBM reading as a measure of general reading proficiency, including comprehension, is well established (e.g., Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988; Marston, 1989; Shinn, Good, Knutson, Tilly, & Collins, 1992). In general, correlations between CBM reading and published measures of reading comprehension have ranged from .63 to .90, with most coefficients being above .80 (Marston, 1989). CBM reading also has correlated highly with teacher judgments of student reading proficiency (Fuchs & Deno, 1981;

Marston & Deno, 1982). The construct validity of CBM reading has been supported by findings that the measures reliably distinguished students with learning disabilities from students in Chapter 1 and general education and students in Chapter 1 from students in general education (Deno, Marston, Shinn, & Tindal, 1983; Marston, Tindal, & Deno, 1983; Shinn & Marston, 1985). CBM reading also has demonstrated sensitivity to student progress over 10 week, 16 week, and 1-year intervals (Deno, Marston, Mirkin, Lowry, Sindelar, & Jenkins, 1982; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1983; Marston, Fuchs, and Deno, 1986; Marston & Magnusson, 1985).

The use of CBM reading to monitor student progress and guide instructional decisions has resulted in significant increases in student outcomes in both special (Fuchs & Fuchs, 1986; Fuchs & Fuchs, 1987; Fuchs, Deno, & Mirkin, 1984; Fuchs, Fuchs, & Hamlett, 1989) and general education settings (Marston & Magnusson, 1985). CBM reading also has increased student awareness of their own learning (Fuchs et al., 1984). When teachers used CBM reading to monitor progress, they (a) increased the structure of their instruction, (b) were more accurate about student progress, and (c) tried more interventions when a student was not making adequate progress (Fuchs et al., 1984).

#### Concerns Regarding CBM Reading with Bilingual Students

The theoretical basis for the relationship between reading fluency and reading comprehension has been developed and supported empirically only in the context of native English speakers learning to read in English (Laberge & Samuels, 1974; Potter & Wamre, 1990). Basically, the development of native language reading skills is preceded by the development of an extensive spoken language base. As soon as the child learns to produce the sounds of a printed word (i.e., decode), the meaning of the word is accessed with a high degree of automaticity (Adams, 1990; Laberge & Samuels, 1974). Thus, fluent decoding is associated strongly with fluent meaning building.

The linkages between reading fluency and global reading proficiency with students learning to read a second language are not as clear, however. First, the sequence of skill acquisition may be

different. For instance, students may learn to produce the sounds of the printed word or use native language decoding skills to decode second language words before they have acquired the extensive second language base necessary to access meaning automatically (Mace-Matluck, 1979).<sup>7</sup>

Second, second language reading skills develop in the context of pre-existing native language proficiency. In the early stages of second language development, students utilize their native language skills to derive meaning from the spoken and printed words in their second language. In effect, students must first decode printed words, translate them from their second language into their native language referents, and finally build meaning. The intermediate translation step is done less as second language skills develop (Jimenez, Garcia, & Pearson, 1992). Whether the construct validity of CBM reading as a global measure of reading proficiency is affected by the intermediate translation step is unknown. In particular, CBM reading may be more a measure of English language proficiency than of reading proficiency for bilingual students. Although these concerns may limit the validity of CMB reading for bilingual students, no published studies have examined their reliability or validity with bilingual students.

The purpose of this study was to examine the reliability, validity, and sensitivity of CBM reading with Hispanic students who are bilingual in English and Spanish.

### Method

#### Subjects

Subjects in the study were 76 second-grade students selected from two elementary schools in a rural school district in Washington state. Both schools were comprised of primarily Mexican-Americans (over 60% of the students in each school).

Washington state has a large contingent of Hispanics Americans, of which 73% are Mexican Americans (U.S. Bureau of Census, 1992-b). Most are now permanent members of the communities in which they live. Hispanic communities in rural Washington are similar in background, language, and culture to other Hispanic communities along the West coast corridor from Southern California to Northern Washington. These communities initially were established when large numbers of Hispanics originally from Mexico traveled up and down the West coast

working on agricultural farms. Washington has the 12th largest population of Hispanics in the U.S., and by percentage they are 18th. The community in which the study was conducted has the fourth highest percentage of Hispanics in the state (62.2%) and the highest percentage of residents in the state who speak Spanish in the home (56.7%) (U.S. Bureau of Census, 1992-a). The community also has the highest percentage of residents in the state who do not speak English very well (31.3%). In addition, the area also is beset by one of the highest poverty rates in the state. The per capita income in the community is the third lowest in the state, and 29.1% of the residents individuals live below the official poverty line (U.S. Bureau of Census, 1992-a). Eighty-nine percent of the students in one of the school, and 95% of the students in the other were on free or reduced lunch.

Two samples of subjects were recruited for this study: a core sample and an extended bilingual sample. The core sample was recruited from four of six classrooms selected randomly from one school. All students in the four classrooms were invited in writing to participate. In addition, parents were contacted by phone if district assessment information indicated that they would likely meet the criteria for inclusion in the study. An extended bilingual sample also was recruited to increase the number of students in the study who were bilingual. All students who were identified by the school district as bilingual in the remaining two classrooms in the first school and all five classrooms of a second school were invited to participate. All parents of bilingual students were contacted in writing and by telephone (or home visit if no telephone was available).

Students whose parents provided permission and who met two criteria were included in this study. The first criteria for inclusion was that students must have read an average of 20 words correctly per minute on CBM reading. This criteria was used so that reading growth could be measured accurately during the study (Fuchs L., Fuchs D, Hamlett, Waltz, & German, 1993; Shinn, Gleason, & Tindal, 1989). The second criteria for inclusion was that students must have met the criteria for inclusion in the English-only group or the bilingual group. To be included in the English-only group, students displayed no Spanish language skills and were identified by the district as students that spoke English only. Students in the bilingual group demonstrated at least

minimal Spanish language skills by identifying at least four pictures correctly on a 1-minute Spanish language fluency task used by the school district. Inclusion in the bilingual group based on this criterion for the Spanish language fluency tasks represents a liberal definition of bilingual. The term bilingual has been used to describe a broad range of fluency in two languages (Esquivel, 1988). At one end are "duel-bilinguals" who have native-like fluency in two language; at the other end are "pseudolinguals" who are not fluent in either language but have some degree of proficiency in both languages (Dulay & Burt, 1980). At intermediate points are individuals sometimes referred to as "incipient bilinguals" who have native-like fluency in one language, and some degree of proficiency in the second language (McCollum, 1981). Because many Hispanic students experience difficulty in both languages, students were included in the bilingual group if they displayed even minimal Spanish language skills.

The number of students who were recruited for participation, the number whose parents provided permission, and the number who met the criteria for inclusion are summarized in Table 1. In total, permission was requested from 105 students. Parental permission to participate was obtained for 94 students (90%), of whom 76 (81%) met the criteria for inclusion in this study. The bilingual group was particularly well represented as a result of the subject recruitment procedures. The number of students by gender in the bilingual and English-only groups was approximately equal (50% and 46% of students in bilingual and English-only groups respectively were female).

-----  
Insert Table 1 Here  
-----

All students in the bilingual group also were Hispanic identified as bilingual by the school district and evaluated for ESL services. Thirty-three of the 50 bilingual students were identified as limited English proficient (LEP) and provided with ESL services. The ESL program had two primary components. First, LEP students received direct Spanish instruction for approximately one-half hour per day. Second, a bilingual teaching assistant was assigned to each classroom for

half of each day. The teaching assistants provided the direct Spanish instruction and additional Spanish assistance to students on an as-needed basis. None of the students in the English-only group received ESL services; 10 of the 26 students were Hispanic (38%).

### English Reading Measures

Reading measures included (a) CBM reading, (b) Stanford Diagnostic Reading Test, and (c) Teacher Rating Scale.

CBM reading measures. CBM reading passages were created and administered using standard CBM reading procedures (Shinn, 1989). Twenty passages were selected randomly from the students' curriculum to create 20 alternate-forms. The second book in Reading Mastery Level II (Engelmann & Bruner, 1974) was used because the majority of students who participated in the study received their reading instruction in that book. The number of words read correctly during 1-minute was used for data analysis.

Student performance was summarized by point, level, and slope. A point estimate corresponds to one sample of behavior on 1 day. The first data point was used. Level estimates reflect overall student performance on multiple samples of behavior on multiple days. Level was determined by calculating the mean of all 20, 1-minute reading samples. Slope estimates represent the student's rate of progress over time. Slope was obtained using an ordinary-least-squares regression line fit to the CBM data (Shinn, Good, & Stein, 1989).

Stanford Diagnostic Reading Test (SDRT). The SDRT (Karlsen & Gardner, 1985) is a group-administered, norm- and criterion-referenced measure of reading achievement. The Red Level of the SDRT, described in the manual as appropriate for use with Grade 1, Grade 2, and low-achieving Grade 3 students, was used. The auditory vocabulary, auditory discrimination, phonetic analysis, and reading comprehension subtests were administered at the beginning of the study to investigate the convergent validity of the CBM reading measures. The reading comprehension subtest was administered at the end of the study to examine sensitivity to change. The number of items students answered correctly was used in data analysis. Internal-consistency reliability coefficients for all but 1 subtest exceeded .80, and alternate-form reliability ranged from .75 to .94.

Concurrent validity coefficients between each of the SDRT subtests and the reading subtests of the Stanford Achievement Test ranged from .67 to .88.

**Teacher Rating of Reading.** The Teacher Rating of Reading (TRR) was used to investigate the convergent validity of CBM reading. Teachers rated the reading competence of their students on a 7-point Likert scale, ranging from way below average (1) to way above average (7). To estimate the reliability of the rating scale, five of six teachers rated 36 students at the beginning and end of the study. The stability of the TRR was .83,  $p < .01$ .

### Language Measures

Language measures included: (a) English language fluency, (b) Spanish language fluency, (c) Teacher Rating of English Skill, and (d) the Language Assessment Scales.

**English Language Fluency.** English Language Fluency (ELF) is a curriculum-based measure of English developed for this study to evaluate the discriminant validity of CBM reading. Twenty alternate-forms of the ELF measures were developed. Each form consisted of 48 drawings. Students were instructed in English to identify verbally each of the drawings on the probe. The number of drawings identified correctly in 1 minute was computed for each probe. The sets of 48 drawings were selected randomly without replacement from the 1,019 drawings that met selection criteria in Distar Language II (Engelmann & Osborn, 1977), the language curriculum used by the school district in the second grade. Only drawings of objects that could be identified unambiguously were selected (e.g., flower, glass, and cow). Drawings of actions and concepts were not included.

Reliability of ELF level was .91,  $p < .01$ , estimated by correlating ELF means for the odd and even numbered data points. The concurrent validity of ELF level with the English LAS and Teacher Rating of English Language were .66,  $p < .01$ , and .66,  $p < .01$ , respectively.

**Spanish Language Fluency.** The curriculum-based measure of Spanish Language Fluency (SLF) was developed for this study to measure the Spanish language skills of the bilingual students. The SLF measures were created by translating the ELF measures described previously into Spanish. The translated measures were evaluated for accuracy and appropriateness by four

bilingual data collectors in the participating schools. When a second answer was suggested (on less than 5% of the drawings), both answers were accepted as correct. Administration and scoring procedures were identical to the ELF, but in Spanish.

The reliability of SLF level was .87,  $p < .01$ , estimated by correlating separate SLF means for the odd and even numbered data points. The concurrent validity of SLF with the Spanish LAS was .32,  $p > .05$ , although it is unclear from this study whether it is the validity of the SLF, the LAS, or both that is limited.

Language Assessment Scales. The Language Assessment Scales (LAS; DeAvila & Duncan, 1977) is a published, norm-referenced test of English and Spanish language proficiency. The LAS short form was administered to examine the discriminant validity of CBM reading. On the Lexical subtest, students are required to produce the correct word for line drawings presented on cards. On the Comprehension subtest, students point to one of three pictures that correctly matches a sentence presented on tape. On the Production subtest, students are required to reproduce verbally a short story presented on tape. The total number of items answered correctly on the three subtests was used in data analysis. The Spanish LAS was identical to the English LAS but administered in Spanish.

The correlation between the English LAS long form and the short form was .98. No reliability data were reported. The concurrent validity of the English LAS short form with the California Test of Basic Skills was .54. Reliability and validity of the Spanish LAS were not available. Although the norms, reliability, and validity of the LAS are limited (Sattler, 1988) the LAS was used because it is one of the better language measures that assesses both English and Spanish skills.

Teacher Rating of English. The Teacher Rating of English (TRE) was used to investigate the discriminant validity of CBM reading. Teachers rated the verbal English language competence of their students on a 7-point Likert scale, ranging from way below average (1) to way above average (7). To estimate reliability, five teachers rated 36 students at the beginning and end of the study. The stability of the TRE was .90,  $p < .01$ .

### Procedures

The LAS was administered by one of the data collectors who was certified in the school district to conduct LAS training. The SDRT was administered by the first author. All CBM data were collected by five trained bilingual individuals from the community. Training took place during two 1-hour sessions. During the first session, the data collectors were taught to administer and score CBM reading with the aid of a video training tape. Also in the first session, data collectors were taught to administer the CBM English and Spanish language fluency measures. The second session focused primarily on continued practice in the administration and scoring of the English and Spanish language measures with students in classroom settings. All data collectors reached a minimum of 90% accuracy on the administration and scoring of the measures before the study.

Data were collected over a 13-week period. The core sample of English-only and bilingual students was administered all of the relevant measures. The extended sample of bilingual students was administered the CBM measures only. All students were administered CBM reading and ELF measures 2 times per week for 10 weeks. The SLF measures were administered 2 times per week for 10 weeks to bilingual students. The SDRT and teacher ratings were completed during week 3 for the core sample. The SDRT reading comprehension subtest was re-administered to the same students during week 13. The LAS was administered during weeks 6 to 10 to students in the core sample.

### Results

The reading and language skills of the English-only students and bilingual students are summarized in Table 2. Differences between the English-only and bilingual groups were not significant on CBM reading measures of point or level of performance, indicating that students in both groups read approximately the same number of words correctly per minute on the first day and throughout the study. In contrast, students in the English-only group scored significantly higher than students in the bilingual group on both the SDRT total score and SDRT reading comprehension subtest. In addition, teachers rated both the reading competence and language competence of English-only students as significantly higher than the bilingual students. Students

in the English-only group were rated slightly above average as a group, and bilingual students were rated slightly below average.

-----  
Insert Table 2 Here  
-----

The rate of reading progress was significantly different for Bilingual students compared to the English-only students. It is noteworthy that the mean slope for English-only students was near 0, ranging from -2.79 to 3.26 words per minute per week. It is unlikely that this low mean slope is an artifact of floor or ceiling measurement effects. Of the 7 students with CBM reading levels below 30 words per minute in this study (i.e., near floor), all but 1 had a positive slope. Similarly, of the 6 students with CBM reading level above 130, all displayed a positive slope. In addition, previous research has found that slope of reading progress can be estimated accurately across a broad range of reading difficulty (Shinn, Gleason, & Tindal, 1989).

On the LAS, the bilingual group's score on the English version ( $M = 70.9$ ) was not significantly different from the Spanish version ( $M = 69.1$ ),  $t(30) = 0.57$ ,  $p > .05$ . However, the bilingual group scored significantly higher on ELF level ( $M = 21.6$ ) than on SLF level ( $M = 13.4$ ),  $t(34) = 7.45$ ,  $p < .05$ . The performance difference on the CBM tasks is not surprising given that the measures were derived directly from the students' English instructional curriculum, although these comparisons should be interpreted cautiously because the test may have varied in difficulty across the two languages.

#### Reliability of CBM Reading Measures

The reliability of CBM reading for English-only and bilingual students is reported in Table 3. For point estimates, the mean 1-week alternate-form reliability was examined. The 20 repeated CBM reading measures yielded 37, 1-week, alternate-form reliability coefficients. The coefficients were summarized by calculating the mean of the Fisher-Z transformed reliability coefficient. The means then were transformed back to reliability coefficients using the inverse Fisher-Z transformation to facilitate interpretation. The coefficients for both groups of students were well

above .80 and indicate strong stability for a 1-minute CBM reading measure. The difference between the reliability coefficients for students in the English-only and bilingual groups was significant, indicating that the reliability of the CBM reading point estimate was higher for bilingual students than for English-only students.

-----  
Insert Table 3 Here  
-----

The reliability of CBM reading for level of student performance was estimated by dividing the repeated CBM reading scores into even and odd data points for each student with 10 to 20 data points. The means of odd and even points then were correlated. The reliability coefficients of .99 for both English-only and bilingual students indicate aggregated CBM reading measures were extremely stable. The coefficients for English-only and bilingual students were not significantly different.

A similar procedure was used to estimate the reliability of slope of student progress. For each student with 10 to 20 data points, an ordinary-least-squares regression line was calculated for the odd and even data points. The slopes of the regression lines then were correlated to obtain an estimate of the reliability of student progress on CBM reading. The reliability estimates were low (below .50) for both English-only and bilingual students and not significantly different.

#### Convergent Construct Validity

The convergent construct validity of CBM reading for English-only and bilingual students was examined by comparing the correlation coefficients between CBM reading and criterion reading measures. The results are presented in Table 2. The validity coefficients range from .51 to .82. For the bilingual students, three of the four correlations were above .70, indicating a strong relationship between CMB reading and the criterion reading measures. Differences in the correlations of the English-only and bilingual students were tested with the z-test for independent correlations. None of the differences was significant, indicating that the magnitude of the

relationship between CBM reading and criterion reading measures was comparable for both groups of students.

### Discriminant Construct Validity

Discriminant construct validity of CBM reading for bilingual students with respect to English language proficiency was addressed by examining the correlation coefficients between CBM reading and criterion measures of English language. As shown in Table 2, all correlations between CBM reading and criterion measures of English language were significantly different from 0. The magnitude of the correlations range from .44 to .62, indicating a moderate relationship between CBM reading and the criterion language measures. The correlations were not significantly different for English-only and bilingual students.

In addition, the correlations of CBM reading with similar criterion measures of reading and language were compared to examine whether CBM reading was a stronger index of reading than language for bilingual students. The results are presented in Table 4. On three of the four comparisons, CBM reading correlated significantly higher with criterion measures of reading than with criterion measures of language. For example, the correlation between CBM reading and teacher's rating of reading competence was .80, significantly higher than the correlation between CBM reading and teacher's ratings of English language competence,  $r = .62$ ,  $z = 2.18$ ,  $p < .05$ . These findings are consistent with the conclusion that CBM reading is more highly related to reading than to language for bilingual students.

-----  
Insert Table 4 Here  
-----

### Discussion

The generalizability of these results is subject to two limitations. First, because only second-grade students were studied, the extent to which the results will generalize to students in other grades is not clear without additional research. For example, as students move up in grade, the vocabulary used in text becomes more difficult until, around grade 4, when the vocabulary in text

closely approximates basic adult vocabulary (Mason, 1992). At that point, bilingual Hispanic students may experience difficulty understanding words they can decode. For example, Garcia (1991) found that the reading comprehension of bilingual Hispanic students was compromised because they did not understand important vocabulary used in 5th and 6th grade reading texts. Similarly, bilingual Hispanic students may have partial understanding of the meaning of difficult vocabulary but lack the deeper word knowledge needed for sufficient text comprehension (Beck and McKeown, 1991; Baumann & Kameenui, 1990). Thus, in later grades, the relation between CBM reading and text comprehension may be lower for bilingual students than for English-only students, although further research is needed to examine this issue specifically.

A second limitation is that the study was conducted in the rural Pacific Northwest. Generalizability of these findings to students in other parts of the country should be considered critically. It is likely these results will be most representative of communities of Mexican Americans with relatively low SES who have settled permanently along the west coast. Other Hispanic groups, including, for example, Cubans in Florida, Puerto Ricans in New York, and migrant agricultural workers from Central America may differ substantially, although the basic language issues will be the similar. One advantage of CBM reading is that local norms can be created for decision making (Shinn, 1988, 1989), which may address well the problems of disparate language and cultural variables that characterize diverse Hispanic groups.

Subject to these limitations, three conclusions are consistent with the results of the study. First, CBM reading was very reliable in general, and at least as reliable for bilingual students as for English-only students. In terms of educational decision making, Salvia and Ysseldyke (1991) suggest that measures have reliability of at least .80 for screening decisions and .90 for important individual decisions. The reliability of a 1-minute CBM reading measure was clearly sufficient for screening purposes for both English-only and bilingual students (Salvia & Ysseldyke, 1991). When 5 to 10 CBM reading measures were aggregated, the reliability coefficient was sufficient for important individual decisions. Indeed, a reliability of .99 indicates that CBM reading level, based on 5 to 10, 1-minute reading samples is an extremely stable measure for both groups. The

reliability of slope of student progress was low for both groups of students and further investigation in this area is warranted. For example, alternative ways of evaluating the reliability of slope may be appropriate; reliability may be examined in relation to the accuracy of predictions of final student performance (Fuchs & Fuchs, 1992; Good & Shinn, 1990), or in terms of the implications for important educational decisions about students.

Second, the study provided strong support for the validity of CBM reading as a measure of English reading proficiency for bilingual students. Both convergent and discriminant construct validity evidence were examined. Convergent evidence is obtained when the measure (in this case CBM reading) is correlated with other measures of the same construct, preferably obtained from a variety of methods (Messick, 1989). The convergent construct validity of CBM reading was comparable for bilingual and English-only students and was similar to other technical adequacy studies of CBM reading (Marston, 1989; Shinn et al., 1992). For example, Shinn et al. (1992) reported correlations between CBM reading and the reading comprehension subtest of the SDRT of .57 to .60 for 114 third-grade students. The convergent validity evidence for CBM reading for bilingual students also was within the range of commonly reported correlations between published measures of reading achievement. For example, correlations between the SDRT and the reading subtests on the Stanford Achievement Test were reported in the SDRT test manual to range from approximately .60 to .85 (Karlen & Gardner, 1985).

Discriminant construct validity evidence is obtained when the measure is discriminable from (i.e., not highly correlated with) measures of similar but distinctly different constructs (Messick, 1989). The results of this study support the conclusion that CBM reading is (a) strongly related to other measures of reading proficiency, (b) less strongly related to measures of English oral language proficiency, and (c) related to reading proficiency and oral language proficiency in the same way for bilingual students as for English-only students. Discriminant validity evidence was provided by examining the relation of CBM reading measures with language measures. The correlations between CBM reading and criterion reading measures were larger in magnitude than the correlations between CBM reading and criterion language measures for bilingual students. No

differences were found between the bilingual and English-only students in the pattern of correlations. These results suggest that CBM reading is a better measure of English reading proficiency than English oral language proficiency for bilingual students, although reading and oral language proficiency are highly related constructs. Perhaps most importantly, there is no evidence that CBM reading is measuring a different construct for bilingual students than for English-only students.

It is interesting to note that the English-only students performed significantly better than the bilingual students on the SDRT and on the teacher ratings of reading competence at pretest. However, the groups performed the same on CBM reading point and level. It is unlikely that this pattern is due to a lack of sensitivity of CBM reading, because CBM reading generally has been found to be more sensitive than other measures of reading in documenting differences between groups of students (Shinn, Ysseldyke, Deno, & Tindal, 1986). It also is unlikely that the SDRT and teacher rating differences were spurious because the differences were found on two different measurement procedures. The most plausible hypothesis is that the reading skills of students in both groups were similar when based on material for which explicit instruction was provided, but higher for students who speak English-only when assessed on material for which explicit instruction was not provided. This would be expected if bilingual students learn well what they are taught but are exposed to fewer incidental or informal learning experiences. Of the reading measures, only CBM reading sampled directly and exclusively from the material students were expected to learn. In contrast, the SDRT included many items on which students had not been given explicit instruction. Similarly, teachers likely base their ratings of student reading competence on a mixture of both instructional material and informal reading experiences.

A third conclusion from this study is that CBM reading provides a sensitive measure of reading progress for bilingual students. During the 13-week study, the rate of increase in the number of words read correctly for bilingual students was 1.3 additional words per minute per week. This slope of improvement is similar to the 1.5 to 2.0 weekly growth rates observed by Fuchs et al. (1993) with 76 second-grade general education students, but below weekly growth rates of general

education students (2.7 words per week) and Chapter 1 students (3.9 words per week) reported by Marston and Magnusson (1985). In the present study, the reading improvement of the English-only students in general education was nearly 0 (0.1 words per week) words per minute per week, significantly below the bilingual students and lower than would be expected from Fuchs et al. (1993) and Marston and Magnusson (1985).

An important difference between this study and the Fuchs et al. (1993) and Marston and Magnusson (1985) studies that may have affected the progress of all participants was the time of year the data were collected. In the Fuchs et al. (1993) year-long study, students progressed throughout the academic year, but at a slightly negatively accelerating rate. In other words, student progress tended to level off toward the end of the school year. In this study, data were collected near the end of the school year when student progress may have been slower. The higher rate of progress of the bilingual students versus the English-only students may be related to systematically faster gains in second-language reading. Alternatively, the steeper progress of the bilingual students may result from positive cognitive consequences of bilingualism (Cummins, 1984; Diaz, 1983). Both of these hypotheses warrant further exploration.

### Educational Implications

Reliable, valid, and sensitive reading measures can provide a means to evaluate the effectiveness of reading interventions for individual students. Evaluating reading interventions is crucial for bilingual Hispanic students because of the clear evidence that many of these students experience severe reading problems. Thus, a validated "ruler" or "scale" for monitoring reading performance could increase the effectiveness of reading interventions with bilingual Hispanic students by formatively evaluating interventions and guiding educators' decision making.

CBM reading can be used on a frequent, repeated basis to monitor student acquisition of reading skills, and to provide information on the success of reading interventions on a student-by-student basis. For example, the progress of two bilingual students on CBM reading is presented in Figure 1. The graph indicates that Carmen made adequate progress (2.3 words per week improvement) during the 13-week period. In contrast, Miguel made no progress, and actually had

a decreasing progress line by the end of the study (-0.65 words per week change). If these data were used to make decisions regarding the effectiveness of instruction for these students, the implication is clear: Miguel is not making adequate progress and a change in his program is warranted.

-----  
Insert figure 1 Here  
-----

The example in Figure 1 is especially relevant because Carmen and Miguel began the study reading at about the same level, were in the same classroom, and received their reading instruction from the same teacher under highly similar instructional circumstances. Using CBM reading to determine reading progress, different conclusion about the effectiveness of the reading program would be drawn for each student. This example highlights the need to evaluate program effectiveness at the individual student level. Even generally effective programs are not necessarily effective for all students (Deno, 1990). To meet the needs of all students, program effectiveness must be evaluated on a student-by-student basis.

These findings have important implications for school psychologists who increasingly are asked to assess and guide interventions for bilingual students. These services can best be provided within a decision-making model. Referrals of bilingual Hispanic students frequently include concerns about reading. CBM reading provides a valuable tool to screen for reading problems. More extensive assessment data, including CBM reading across multiple days in material of varying difficulty and an analysis of student reading errors provide important information to determine whether services beyond general education are needed to remediate reading problems.

School psychologists' measurement expertise can play an important role in establishing procedures for the frequent collection of CBM reading data, including the selection of test material, the standardization of test procedures, and the way the test data are recorded. School psychologists also can play a role in interpreting student progress data, and in presenting information to teachers, parents, and students in a clear, understandable manner. Finally, as

consultants, school psychologists can establish procedures for objectively determining the adequacy of student progress, and determining what changes could be made in the reading programs of bilingual students who are not making adequate progress.

## References

- Adams, M. (1990). Beginning to read: Thinking and learning about print. Urbana-Champaign, IL: University of Illinois, Center for the Study of Reading.
- Arias, M. (1986). The context of education for Hispanic students: An overview. American Journal of Education, 95, 26-57.
- Baumann, J. F., & Kameenui, E. J. (1990). Research on vocabulary instruction: Ode to Voltaire. In J. Flood, J. Jensen, D. Lapp, & J. R. Squire (Eds.), Handbook of research on teaching the English language arts (pp. 604-632). New York: Macmillan.
- Beck, I., & McKeown, M. (1991). Conditions of vocabulary acquisition. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), Handbook of reading research (Vol 2, pp. 789-814). New York: Longman.
- Casas, J., Furling, M., Solberg, V., & Carranza, O. (1990). An examination of individual factors associated with the academic success and failure of Mexican-American and Anglo students. In A. Barone & E. Garcia (Eds.), Children at risk: Poverty, minority status, and other issues in educational equity (pp. 103-119). Washington, DC: National Association of School Psychologists.
- Cummins, J. (1984). Bilingualism and special education: Issues in assessment and pedagogy. Clevedon, England: Multilingual Matters.
- DeAvila, E., & Duncan, S. (1977). Language Assessment Scales (2nd ed.). Corte Madera, CA: Linguametrics Group.
- Deno, S. (1990). Individual differences and individual difference: The essential difference of special education. Journal of Special Education, 24, 160-173.
- Deno, S. L. (1989). Curriculum-based measurement and alternative special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), Curriculum-based measurement: Assessing special children (pp. 1-17). NY: Guilford.
- Deno, S., Marston, D., Mirkin, P., Lowry, L., Sindelar, P., & Jenkins, J. (1982). The use of standard tasks to measure achievement in reading, spelling, and written expression: A

- normative and developmental study (Research Report. No. 87). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- Deno, S., Marston, D., Shinn, M., & Tindal, G. (1983). Oral reading fluency: A simple datum for scaling reading disability. Topics in Learning and Learning Disability, 2, 53-59.
- Deno, S., Mirkin, P., & Chlang, B. (1982). Identifying valid measures of reading. Exceptional Children, 49, 36-45.
- Diaz, R. M. (1983). Thought and two languages: The impact of bilingualism on cognitive development. In E. W. Gordon (Eds.), Review of Research in Education (pp. 23-53). Washington, DC: American Educational Research Association.
- Dulay, H., & Burt, M. (1980). The relative proficiency of limited English proficient students. NABE Journal, 4(3), pp. 1-23.
- Engelmann, S., & Bruner, E. (1974). Distar reading II. Chicago, IL: Science Research Associates.
- Engelmann, S., & Osborn, J. (1977b). Distar language II. Chicago, IL: Science Research Associates.
- Esquivel, G. (1988). Best practices in the assessment of limited English and bilingual children. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology (pp. 113-125). Washington, DC: National Association of School Psychologists.
- Figueroa, R. (1990). Best practices in the assessment of bilingual children. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology (pp. 113-125). Washington, DC: National Association of School Psychologists.
- Fradd, S., & Correa, V. (1989). Hispanic students at risk: Do we abdicate or advocate? Exceptional Children, 56, 105-110.
- Fuchs, L. S. & Fuchs, D. (1987). The relation between methods of graphing student performance data and achievement: A meta-analysis. Journal of Special Education Technology, 8, 5-13.

- Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation on student achievement: A meta-analysis. Exceptional Children, 53, 199-208.
- Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. School Psychology Review, 21, 45-58.
- Fuchs, L. S., Deno, S. L., & Mirkin, P. (1984). The effects of frequent curriculum based measurement and evaluation on pedagogy, student achievement and student awareness of learning. American Educational Research Journal, 21, 449-460.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Computers and curriculum-based measurement: Effects of teacher feedback systems. School Psychology Review, 18, 112-125.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? School Psychology Review, 22(1), 27-48.
- Fuchs, L., & Deno, S. (1981). A comparison of reading placements based on teacher judgment, standardized testing and curriculum-based assessment (Research Rep. No. 56). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.
- Fuchs, L., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. Remedial and Special Education, 9, 20-28.
- Galagan, J. E. (1985). Psychoeducational testing: Turn out the lights, the party's over. Exceptional Children, 52, 288-298.
- Garcia, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. Reading Research Quarterly, 26, 371-392.
- Garcia, G. E., & Pearson, P. D. (1994). Assessment and diversity. In L. Darling-Hammond (Ed.), Review of research in education (pp. 337-391). Washington DC: American Educational Research Association.
- Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Boston, MA: Allyn and Bacon.

- Good, R. H., & Salvia, J. (1988). Curriculum bias in published, norm-referenced reading tests: Demonstrable effects. School Psychology Review, 17, 51-60.
- Good, R. H., & Shinn, M. R. (1990). Forecasting accuracy of slope estimates for reading curriculum-based measurement: Empirical evidence. Behavioral Assessment, 12, 179-193.
- Jenkins, J. R., & Pany, D. (1978). Standardized achievement tests: How useful for special education? Journal of Special Education, 44, 448-453.
- Jimenez, R. T., Garcia, G. E., & Pearson, P. D. (1992). Opportunities and obstacles in bilingual reading : The reading strategies of bilingual students. Unpublished Manuscript.
- Kanellos, N. (1993). The Hispanic-American almanac. Detroit, MI: Gale Research Inc.
- Karlsen, B., & Gardner, E. (1985). Stanford Diagnostic Reading Test (3rd ed.). San Antonio, TX: The Psychological Corporation.
- Laberge, D., & Samuels, S. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.
- Laosa, L. M. (1975). Bilingualism in three United States Hispanic groups: Contextual use of language by children and adults in their families. Journal of Educational Psychology, 67, 617-627.
- Luz Reyes, M. (1992). Challenging venerable assumptions: Literacy instruction for linguistically different children. Harvard Educational Review, 62, 427-446.
- Mace-Matluck, B. (1979). Order of acquisition: Same or different in first- and second-language learning. The Reading Teacher, 25, 7-14.
- Marston, D. (1989). Curriculum-based measurement: What is it and why do it? In M. Shinn (Ed.), Curriculum-based measurement: Assessing special children (pp. 18-78). NY: Guilford Press.
- Marston, D., & Deno, S. (1982). Implementation of direct and repeated measurement in the school setting (Research Report No. 106). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.

- Marston, D., & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. Exceptional Children, 52, 266-276.
- Marston, D., Fuchs, L., & Deno, S. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. Diagnostic, 11, 77-90.
- Marston, D., Tindal, G., & Deno, S. (1983). A comparison of standardized achievement tests and direct measurement techniques in measuring student progress (Research Report No. 126). Minneapolis: University of Minnesota Institute for Research on Learning Disabilities.
- Mason, J. A. (1992). Reading stories to preliterate children: A proposed connection to reading. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), Reading acquisition (pp. 215-242). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McCollum, P. A. (1981). Concepts in bilingualism and the relationship to language assessment. In J. G. Erickson & D. R. Omark (Eds.), Communication assessment of the bilingual bicultural child (pp. 25-41). Baltimore, MD: Guilford Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New York: Macmillan.
- Millis, I. V. S., Campbell, J., & Farstrup, A. (1993). National assessment of educational progress 1992 reading report card for the nation and the states: Data from the national and trial states assessment. Washinton, DC: U.S. Department of Education, National Center for Education Statistics.
- Nunnally, J. (1978). Psychometric theory. New York: McGraw-Hill.
- Ortiz, V. (1986). Reading activities and reading proficiency among Hispanic, Black, and White students. American Journal of Education, 95, 59-75.
- Potter, M. L., & Wamre, H. M. (1990). Curriculum-based measurement and developmental reading models: Opportunities for cross-validation. Exceptional Children, 57, 16-25.
- Reddy, M. (Ed.). (1993). Statistical record of Hispanic Americans. Detroit, MI: Gale Research Inc.

- Reschly, D. (1988). Special education reform: School psychology revolution. School Psychology Review, 17, 459-475.
- Salvia, J., & Ysseldyke, J. (1991). Assessment (5th ed.). Boston: Houghton-Mifflin.
- Sattler, J. M. (1988). Assessment of children (3rd ed.). San Diego, CA: Jerome Sattler.
- Shapiro, E. S., & Derr, T. F. (1987). An examination of the overlap between reading curricula and standardized achievements tests. School Psychology Review, 21, 59-67.
- Shinn, M. (1988). Development of curriculum-based local norms for use in special education decision making. School Psychology Review, 17, 61-80.
- Shinn, M. (Ed.). (1989). Curriculum-based measurement: Assessing special children. New York: Guilford.
- Shinn, M. R., Gleason, M. M., & Tindal, G. (1989). Varying the difficulty of testing materials: Implications for Curriculum-Based Measurement. Journal of Special Education, 23, 223-233.
- Shinn, M. R., Good, R. H., & Stein, S. (1989). Summarizing trend in student achievement: A comparison of evaluative models. School Psychology Review, 18, 356-370.
- Shinn, M., & Marston, D. (1985). Differentiating mildly handicapped, low-achieving and regular education students: A curriculum-based approach. Remedial and Special Education, 6, 31-45.
- Shinn, M., Good, R., Knutson, N., Tilly, D., & Collins, V. (1992). Curriculum-based measurement of oral reading fluency: A confirmatory analysis of its relation to reading. School Psychology Review, 21, 459-480.
- U.S. Bureau of the Census (1992-a). Population projections of the United States, by age, sex, race, and Hispanic origin: 1992 to 2050 (No. P25-1092). Washington, DC: U.S. Government Printing Office.
- U.S. Bureau of the Census (1992-b). Statistical abstract of the United States (112th edition). Washington, DC: U.S. Government Printing Office.
- Willig, A. (1985). A meta-analysis of selected studies on the effectiveness of bilingual education. Review of Educational Research, 55, 269-317.

Table 1. Number of Students in the Study from Recruitment to Inclusion by Group

	Number Recruited	Received Permission	Met Criteria
Core Sample			
Bilingual Students	35	32	25
English-only Students	34	28	26
Extended Bilingual Sample	36	34	25

Table 2. Mean English Reading and Language Scores by Group.

Variables	English-Only Students			Bilingual Students			t(df)
	n	Mean	SD	n	Mean	SD	
<u>Reading</u>							
Curriculum-Based Measurement--Point	26	86.7	37.1	50	68.5	41.7	1.9(56.2)
Curriculum-Based Measurement--Level	26	76.7	34.6	50	69.5	38.9	0.8(56.3)
Curriculum-Based Measurement--Slope	26	0.1	1.5	50	1.3	1.2	3.5(40.5)**
Stanford Diagnostic Reading Test Total Score--Pretest	24	129.2	22.9	19	108.0	16.7	4.4(35.7)**
Stanford Reading Comprehension Subtest--Pretest	25	42.7	4.4	21	34.0	8.3	4.3(29.2)**
Stanford Reading Comprehension Subtest--Posttest	--	--	--	25	38.4	6.2	--
Teacher Rating of Reading	26	4.4	1.3	25	3.4	1.8	2.2(43.0)*
<u>Language</u>							
CBM English Language Fluency--Level	26	25.1	5.4	50	22.2	6.3	2.1(57.9)*
Language Assessment Scales -- English	--	--	--	31	70.9	20.3	--
Teacher Rating of English	26	4.5	1.2	25	3.3	1.4	3.3(47.7)**

Note. Curriculum-based measure of slope is the predicted increase in the number of words read correctly per week. Test statistic is the Behrens-Fisher t-test for independent groups with unequal sample sizes using Welch's solution for df.

\*p < .05. \*\*p < .01.

Table 3. Reliability and Validity of CBM Reading by Group

Criterion	English-Only Students	Bilingual Students	Test Difference <sup>a</sup>
Reliability			
Point <sup>a</sup>	.87** (26)	.92** (50)	4.17*
Level <sup>b</sup>	.99** (26)	.99** (50)	0.00
Slope <sup>bc</sup>	.39 (26)	.49** (50)	1.14
Convergent Construct Validity			
Stanford Diagnostic Reading Test--Total Score <sup>b</sup>	.51* (24)	.53* (19)	0.08
Stanford Reading Comprehension Subtest--Pretest <sup>b</sup>	.56** (25)	.73** (21)	0.93
Stanford Reading Comprehension Subtest--Posttest	--	.76** (25)	--
Teacher Rating of Reading <sup>b</sup>	.82** (26)	.80** (25)	0.20
Discriminant Construct Validity			
CBM English Language Fluency--Level <sup>b</sup>	.54** (26)	.44** (50)	0.52
Teacher Rating of English <sup>b</sup>	.62** (26)	.62** (25)	0.00
Language Assessment Scales--English	--	.47** (31)	--

**Note.** The number of subjects is reported in parentheses.

<sup>a</sup>Test statistic for reliability coefficient differences is the t-test for independent sample means. The means were derived using Fisher Z transformation.

<sup>b</sup>Test statistic for differences between correlations for CBM reading and criterion reading and language measures is the z-test for independent correlations (Glass and Hopkins, 1984, p. 307).

<sup>c</sup>Reliability is based on Spearman-Brown Prophecy formula (Nunnally, 1978).

\* $p < .05$ . \*\* $p < .01$ .

Table 4. Differences in Correlations between CBM Reading and Criterion Measures of Reading and Language for Bilingual Students

Reading Measure	Correlation with		Correlation with CBM Reading	z-test Difference <sup>a</sup>
	CBM Reading	Language Measure		
Teacher Rating of Reading	.80	Teacher Rating of English	.62	2.18* (25)
Stanford Diagnostic Reading Test Total--Pretest	.59	Language Assessment Scales -- English	.26	1.38 (11)
Stanford Reading Comprehension Subtest--Pretest	.70	Language Assessment Scales -- English	.27	2.18* (12)
Stanford Reading Comprehension Subtest--Posttest	.88	Language Assessment Scales -- English	.40	2.24* (14)

Note. The number of subjects is reported in parentheses. Correlations between CBM Reading and the LAS vary somewhat because the number of students in each analysis differs because of missing data.

<sup>a</sup>Test statistic is the z-test for dependent correlations, Glass & Hopkins (1984), p. 310.

\* $p < .05$ .

Figure Caption

Figure 1. Progress of two bilingual students on CBM reading.

