



## **TECHNICAL REPORT #12:**

### **General Outcome Measures for Students with Significant Cognitive Disabilities: Pilot Study**

*Teri Wallace and Renáta Tichá*

**RIPM Year 3: 2005 – 2006**

**Date of Study: March 2006 – May 2006**

**March 2007**

The College of Education  
& Human Development

---

UNIVERSITY OF MINNESOTA



Produced by the Research Institute on Progress Monitoring (RIPM) (Grant # H324H30003) awarded to the Institute on Community Integration (UCEDD) in collaboration with the Department of Educational Psychology, College of Education and Human Development, at the University of Minnesota, by the Office of Special Education Programs. See [progressmonitoring.net](http://progressmonitoring.net).

Abstract

The goal of this pilot study was to investigate the suitability of the format, administration procedures and duration of newly created general outcome measures (GOMs) in reading for students with significant cognitive disabilities, as well as to find out whether these measures would produce useful and reliable data. It was found that students with significant cognitive disabilities were able to respond to the format, administration directions and timing of the general outcome measures. The results suggest that timed general outcome measures can serve as a useful and efficient assessment tool in reading for students with significant disabilities. Based on the pilot study, more research is needed to establish technical characteristics of the newly created general outcome measures.

### General Outcome Measures (GOMs) for Students with Significant Cognitive Disabilities

Literacy is considered a key factor in future educational and vocational success (Gurry & Larkin, 1999; Kliewer & Landis, 1999). As such, legislation has been established that raises expectations for all students with specific emphasis on improving literacy skills (No Child Left Behind Act of 2001; PL 107-110). This emphasis on *literacy* is intended for all students including students with significant cognitive disabilities. Students with significant cognitive disabilities are defined by NCLB as those who are (1) within one or more of the existing categories of disability under the Individuals with Disabilities Education Act (IDEA); and (2) whose cognitive impairments may prevent them from attaining grade-level achievement standards, even with the very best instruction. While both providing literacy instruction and assessing the performance for students with significant cognitive disabilities are important; they are also challenging, and receiving considerable attention (Browder & Spooner, 2006; Browder, Wallace, Snell, & Klienert, 2005; Downing, 2005) This study examines the development of general outcome measures (GOMs) to assess the reading performance of students with significant cognitive disabilities.

### *Reading Instruction for Students with Significant Cognitive Disabilities*

Reading instruction for students with significant cognitive disabilities has most typically focused on sight words for functional reading. Browder (2001) defined the characteristics of functional reading as 1) the acquisition of specific sight words that have immediate functional use, 2) an alternative way to learn reading skills when literacy is not being achieved, and 3) a way to gain quick success in reading that could promote future reading. While research has found that students with significant cognitive disabilities can learn sight words (Browder & Xin, 1998) and can use them when cooking (Collins, Branson, & Hall, 1995), when reading labels

(Collins & Griffen, 1996), and for self-instruction on the job (Browder & Minarovic, 2000); sight word instruction alone has several limitations. Browder, Courtade-Little, Wakerman, & Rickelman (2006) summarized some of the limitations with supporting research:

- Browder and Xin (1998) found that most studies using sight words have not measured comprehension or functional use;
- Conners (1992) and Katims (2000 b) suggest that reading instruction in general education requires gathering meaning from print rather than simply identifying a word;
- Joseph and Seery (2004) found that sight word instruction focuses on the whole-word recognition in absence of phonetic understanding;
- Groff, Lapp, & Flood (1998) suggests explicit phonetic instruction for those struggling to read; and
- Joseph and Seery (2004) found, in a review of the literature, that students with mental retardation can learn phonics skills but little research has been done.

In addition, in a CEC position paper on issues of assessing students with the most significant cognitive disabilities, Perner (2007) states that it is critical to ensure that functional skills and curriculum are part of the alternative standards and assessments.

Finally, in their chapter, Browder, Courtade-Little, Wakerman, & Rickelman (2006) conclude that it is not necessary to choose between a functional and a literacy-based approach to reading. They suggest that both can benefit students with significant cognitive disabilities in addition to information about literacy concepts, such as: concept of print, words and letters.

Reading is important for all children and while there has been debate over the components of reading and how it is best taught, its value has not been challenged. However, as

noted above, the use of a separate functional curriculum for students with severe disabilities has been prominent since the 1980s (Browder & Spooner, 2006). It wasn't until the requirements of NCLB were established that educators and researchers understood an increased focus on academics would be needed, even within the alternate assessment.

#### *A Requirement for Success and Assessment*

The No Child Left Behind Act (NCLB) of 2001 provides a legal mandate to ensure that all students are learning. This law requires the development of state standards and large scale assessments intended to measure schools' success in achieving established content and achievement standards, including standards in areas related to literacy. This legislation (e.g., Public Law No. 107-110, 115 Stat. 1425, 2002) has mandated that students with significant cognitive disabilities be included in states' accountability systems.

IDEA regulations, published in the Federal Register in December 2003, provide an avenue for students with disabilities to be assessed through one of five options as determined by the child's IEP team, including:

- The regular grade-level State assessment,
- The regular grade-level State assessment with accommodations,
- Alternate assessments aligned with grade-level achievement standards,
- Alternate assessments based on alternate achievement standards, or
- Modified achievement standards.

Alternate assessments, in general, are intended for use with students with disabilities who are unable to participate meaningfully in general state and district assessment systems, even with accommodations (Roach & Elliott, 2006). Alternate assessments based on alternate achievement

standards go further and are for students with significant cognitive disabilities who cannot meet typical grade-level achievement standards.

In 2002, NCLB increased the federal government's emphasis on assessment and accountability systems. Noted by Roach & Elliott (2006), many states have struggled to develop alternate assessments that meet federal mandates for two primary reasons. First, the skills and concepts in the state academic standards were considered inappropriate or irrelevant for students with significant cognitive disabilities, which resulted in alternate assessments that focused on functional domains; and second, the development of alternate assessments was considered a special education function and deemed to be only somewhat connected to states' overall assessment systems. However, the reauthorization of IDEA (2004) and guidelines for using alternate assessment with alternate achievement standards for NCLB (Federal Register, December 9, 2003) both require determining adequate yearly progress for this population using alternate assessments that are linked to the state's academic content standards. States may use alternate achievement standards for up to 1% of students with significant cognitive disabilities and modified achievement standards for up to 2% of students with persistent academic difficulties.

In general, assessing the academic performance and progress of students with significant cognitive disabilities has long been a challenge to the field of education (Perner, 2007). The information gained from using standardized tests with students with significant cognitive disabilities may not provide useful information to teachers to use in educational decision-making. Additional assessment strategies may supplement standardized tests, including: Criterion-referenced tests, observations, fluency measures, portfolios, and others. These may

provide useful data for educational decision-making. Another such assessment strategy is curriculum-based measurement (CBM).

Curriculum-based measurement has a 30-year research base establishing its reputation as an evidence-based practice in measuring individual performance and progress. CBM's most extensive research history is in reading for elementary-aged students, but there is also research in other instructional areas, such as: writing, spelling, math, science, and more (Allinder & Swain, 1997; Calhoon & Fuch, 2003; Espin & Deno, 1994-95; Espin et al., 2000; Espin et al., 2005; Foegen & Deno, 2001; Fuchs & Fuchs, 2002; Shinn, Deno, & Espin, 2000). While CBM was originally intended for school-age students, it is now used with children across the age spectrum, pre-K through high school. Researchers have expanded the idea of CBM to other systems with similar goals such as assessment processes, e.g. Individual Growth and Development Indicators (IGDIs) and Dynamic Indicators of Basic Literacy Skills (DIBELS) used with children in preschool and daycare settings (Good & Kaminski, 1996; Greenwood, Tapia, Abbott, & Walton, 2003; Hintze, Ryan, & Stoner, 2003; Lembke, Deno, & Hall, 2003).

#### *Extending CBM to Students with Significant Disabilities*

Most recently the question is whether CBM can be effective in measuring the academic performance of students with significant cognitive disabilities (Otaiba & Hosp, 2004; Tindal et al., 2003). Otaiba and Hosp (2004) used CBM as one of their assessment measures to monitor progress while implementing a tutoring model with students with Down syndrome. The researchers also included pre and post measures on specific aspects of the Peabody Picture Vocabulary Test-Revised (PPVT-R) – one-word receptive vocabulary, the Comprehensive Test of Phonological Processing (CTOPP) – phonological processing skills, and the Word Attack and Word Identification subtests of the Woodcock Reading Mastery Test-Revised (WRMT-R).

Otaiba and Hosp state "...we found that CBM was a sensitive, reliable measure for monitoring reading growth for students with Down syndrome. This indicates that teachers can use CBM as a reliable way to monitor students' progress and change instruction accordingly" (p. 33).

It is notable that Otaiba and Hosp (2004) used two CBM probes - letter sounds and passage reading - typically recommended for beginning readers. They added a third - sight words - more specific to the target skills for the participants. One option when applying CBM to students with significant cognitive disabilities is to use beginning reading CBM probes that are readily available.

The present study examines the use of general outcome measurement (GOM) for assessing student performance in the area of reading for students with significant cognitive disabilities. Initially we considered using curriculum-based measurement (CBM) as the approach to measuring performance and progress. However, we later determined CBM had certain limitations that could affect its appropriateness for students who were not capable of a verbal response, who usually had instruction focused on functional sight words rather than academic reading, and for whom consensus regarding progress within the general curriculum had not yet been determined. Therefore we decided to use a measure that provided an indicator of the general outcome area of reading. The intent was to create valid and reliable measures to assess students' performance and progress in academic areas (in this case reading) that align with state standards as well as their individual IEP goals to provide teachers with information they could use to judge students' yearly progress and make meaningful instructional decisions. We chose a reading development model - developed by Chall in 1996 - to guide our GOM development.

*Chall's Reading Development Model*



The model selected as a guiding framework for this study was developed by Chall (1996) and outlines developmental reading stages for preschool through adult readers. According to Chall, there are six stages along a developmental continuum and they are not fixed by grade level. For example, a high school student might be at Stage 1 (Initial Reading/Decoding), which is often associated with typically developing 6 – or 7-year-olds. A 2-year-old might be at Stage 0 (Pre-reading) but so might a student in middle school. This model recognizes that individuals might be at similar development reading levels while at very different ages for whatever reason (i.e., social, cognitive, environmental, experiential).

Table 1: Chall’s stages of reading development and typical associated ages (1996)

Stage 0 - Pre-reading (birth to age 6)

Stage 1 - Initial reading and decoding (ages 6-7)

Stage 2 - Confirmation and fluency (ages 7-8)

Stage 3 - Reading to learn the new (ages 8-14)

Stage 4 - Multiple viewpoints (ages 14-18)

Stage 5 - Construction and reconstruction (ages 18 and older)

Chall’s model (1996) provides a meaningful framework for recognizing the possible usefulness of measures found to be applicable for typically developing emergent and early readers as well as for students with significant disabilities in a similar stage. For example, a Letter Identification measure, associated with Stage 0, might assess performance of 4<sup>th</sup> grade student with significant cognitive disabilities who is at that developmental reading level, and so on.

The first set of research questions addressed the suitability of the format, administration procedures and duration of GOM measures, including the suitability of the criterion measures for

students with significant cognitive disabilities. The second set of questions examined the usefulness and reliability of the data produced using these measures.

## Method

### *Participants*

The participants in the study were 13 students with significant cognitive disabilities from two urban schools in Minnesota. Ten students (77%) were male and three (23%) were female. Students with “significant cognitive disabilities” were defined for the purposes of this study as “students who participate in alternate assessment with alternate achievement standards linked to state grade level content standards” (NCLB, 2005). The 13 students represent a convenience sample. First, two schools with a program for students with developmental cognitive delay were identified by a teacher on special assignment in the school district. Two teachers in each school who teach students with significant cognitive disabilities agreed to participate in the study. Students in the four classrooms whose parents gave permission participated. Students from Kindergarten through grade five were represented. There were seven (54%) African-American, three (23%) Hispanic, two (15%) White, and one (8%) Native American students. Nine out of the 13 students (69%) received free lunch. None of the students were classified as receiving reduced lunch. Two out of the 13 students (15%) were English Language Learners (ELL). Based on their IEP (Individualized Education Program), DCD (developmental cognitive disability) was a primary label for 10 students in the study, one student was labeled SMI (severe multiple impairment), one OHD (other health disability), and one VI (visual impairment). In the case of the students whose primary disability label was not DCD, their secondary or tertiary label suggested this impairment.

Similarly, the demographic composition of students in special education in the school district from which the study sample was obtained was as follows: 67% male and 33% female, 53% African-American, 12% Hispanic, 24% White, and 6% Native American, 73% received free or reduced lunch, and 15% were English Language Learners (ELL). Our sample was therefore a good representation of the district demographics, with the exception of an 11% over-representation of Hispanic and 9% under-representation of White students in our sample.

### *Materials*

Assessment tools used for this pilot study were six newly developed general outcome measures (GOMs) and three criterion measures. There were three GOM matching measures and three GOM identification measures: picture, letter and word matching; and picture, letter and word identification (see Table 1 and Figures 1 and 2). Each GOM consisted of 30 laminated 8.5x11 inch cards that were numbered from 1 to 30. Card number 1 was a model card, cards with numbers 2 and 3 were practice cards, and cards with numbers 4 – 30 were test cards. In the case of the matching measures, the front of each card had one item boxed using a rectangle with 6-pt thick black lines either on top (pictures and letters) or on the left side (words) and three choices in a row below (pictures and letters) or on the right side (words). The back of the matching cards included a title of each measure, card number, and the project logo. The front side of the identification cards had three choices of items either in a row (pictures and letters) or in a column (words). In addition to the title of the measure, the card number, and the grant symbol, the back of the identification cards also had the correct response spelled out in the middle. The font and size used for letters in both the matching and identification measures was Century Gothic 200, and for words Century Gothic 100. The pictures were black-and-white drawings of a size equivalent to the letters. Each measure was accompanied by a sheet with

detailed administration directions. A scoring sheet was used to record responses every time a GOM was administered (see below). Both the data collector and observer used a small portable tape recorder with an ear piece and a tape with 3, 5, 7 and 10 minute recorded time markers. Both also used a timer.

Table 1

*GOM Measures*

<b>Type</b>	<b>Stimulus</b>	<b>Description</b>
<i>Matching</i>	Pictures	On a card, student points to the picture in a row of 3 that matches the picture in a box.
	Upper Case Letters	On a card, student points to the letter in a row of 3 that matches the letter in a box.
	Sight Words	On a card, student points to the word in a column of 3 that matches the word in a box.
<i>Identification</i>	Pictures	On a card, student points to the picture in a row of 3 that matches the word the researcher says.
	Upper Case Letters	On a card, student points to the letter in a row of 3 that matches the letter the researcher names.
	Sight Words	On a card, student points to the word in a column of 3 that matches the word the researcher says.

The three criterion measures used for this pilot study were the Peabody Picture Vocabulary Test – Third Edition (PPVT-III), receptive portion of the Comprehensive Receptive and Expressive Vocabulary Test – Second Edition (CREVT-2), and the RIPM Early Literacy Knowledge and Reading Readiness Checklist. The PPVT-III is an un-timed and individually administered assessment tool of receptive vocabulary. Assessed individuals are asked to point to one black and white drawing on a page from a choice of four. It is designed for use with participants between the ages two and 90+ years. It has two parallel forms, A and B. The PPVT-III was developed in 1996 using 2,725 participants nationwide between ages two and 90+ years. Eighteen percent of the sample were African American students, 64% White, 13% Hispanic and 5% of students were of other origin. There were participants receiving special education services in the standardization sample: 5.5% were students with learning disabilities, 2.3% students with speech impairment, 2.2% adults with mental retardation, and 1.2%t students with mental retardation. The reliability coefficients reported were above .90. To establish criterion validity of the PPVT-III, the authors used three intelligence tests: the Wechsler Intelligence Scale for Children – Third Edition (corrected correlation coefficients .91 and .92 for the two parallel forms), the Kaufman Adolescent and Adult Intelligence Test (.87 and .91), and the Kaufman Brief Intelligence Test (.82 and .80). The validity of PPVT-III was also examined using the Oral and Written Language Scales, namely Listening Comprehension (.68 and .70) and Oral Expression (.75 and .73).

The second criterion measure we used in the pilot study was the receptive portion of the CREVT-2. In the same way as the PPVT-III, the CREVT-2 is an un-timed, individually administered test. The receptive part of the test was developed for administration to students between four and 90 years of age. Participants are asked to point to one color photograph from a

choice of six on a page. The photographs on each page are grouped by themes, e.g. animals. The CREVT-2 has two forms, A and B. It has been normed on 2,545 individuals. Three percent were Native Americans, 11% Hispanic Americans, 2% Asian Americans, 12% African Americans and 72% other. Six percent of the normative sample had an identified learning disability, 8% had a speech-and-language disorder, 2% mental retardation, and another 2% had another disability. The reliability coefficients reported for the receptive vocabulary subtest were above .90. To establish criterion validity of the receptive portion of the CREVT-2, the authors correlated this subtest with the Peabody Picture Vocabulary Test – Revised (.59 and .61 for form A and B), the Expressive One-Word Picture Vocabulary Test – Revised (.67 and .66), the Wechsler Intelligence Scale for Children – III: Vocabulary (.66 and .71), the Wechsler Intelligence Scale for Children – III: Full Scale (.39 and .44), the Clinical Evaluation of Language Fundamentals Revised (.74 and .74), the Test of Language Development – Primary (.86 and .84), the Comprehensive Test of Nonverbal Intelligence (.56 and .58), and the Gray Oral Reading Test – IV (.71 and .73).

The third criterion measure used in this pilot study was the RIPM Early Literacy Knowledge and Reading Readiness Checklist (Checklist for short) for special education teachers developed for the purposes of this study. The reason for developing this Checklist was to be able to compare student performance on the piloted GOM measures and the special education teachers' view of their students' performance. We based our rationale on the fact that in comparison with a general education setting, special education teachers spend more time with each student in small groups or on an individual basis, and therefore are more likely to have an accurate knowledge of the student's reading or pre-reading performance. When developing the Checklist, we first studied similar existing materials, i.e. the Minneapolis Early Childhood

Special Education Checklist, the Minneapolis Developmental Cognitive Disabilities Checklist, the Minneapolis DCD Scope and Sequence in Reading (Minneapolis Public Schools), and the Checklist for Assessing Early Literacy Development (Katims, 2000 a). The Checklist consists of six subscales: I. Concepts about books, print, letters and words; II. Alphabetic knowledge and beginning decoding skills; III. Phonemic awareness; IV. Sight word vocabulary; V. Beginning comprehension skills; and VI. Daily living reading skills. Each item under each subscale requires a “yes” or “no” response from the special education teacher. The score for each subscale as well as the total score is recorded.

### *Procedures*

*General outcome measures (GOMs) development.* The GOMs were developed based on the principles of Curriculum-Based Measurement (CBM) while considering the skills of students with significant cognitive disabilities. The rationale described in Browder et al. (2005) regarding the need to examine strategies for assessing academic progress of students with significant cognitive disabilities served as the theoretical framework for this research. Specifically, expanding CBM to students with significant disabilities within academic topics seemed plausible. The process of measure development started by studying various documents concerning typically developing students as well as students with significant cognitive disabilities in general, in Minnesota in particular and in other states. Different approaches to curricula for students with significant developmental disabilities were taken into consideration, e.g., developmental or functional. Chall’s (1996) stages of reading development were used to organize potential GOMs into a sequence of reading development of typically developing students. Alternate achievement standards and alternate assessments in reading in Minnesota

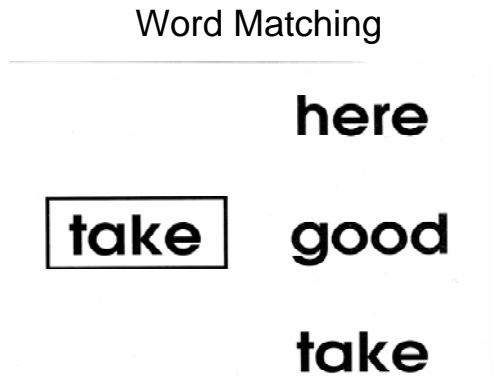


and other states, e.g. Massachusetts, were examined. Progress monitoring measures for students with significant cognitive disabilities in other states, e.g. Oregon, were taken into consideration.

In addition to state-level work, materials used in classrooms where students with significant cognitive disabilities are taught were studied, and professionals working with these students were consulted. Several curricula and related materials for students with significant cognitive disabilities were examined in the classroom, e.g. Edmark or Learning Mastery; sight word lists (Dolch List, Fry's 300 Instant Sight Word List). An advisory committee meeting was held with special education teachers, specialists, researchers, and administrators to discuss the context and possibilities for developing GOMs in reading for students with significant cognitive disabilities. Another source of ideas for developing GOMs for this population were progress monitoring measures developed for students in early childhood education, e.g. Individual Growth and Development Indicators (IGDIs) and Dynamic Indicators of Basic Early Literacy Skills (DIBELS).

After the examination of materials and consultation with experts in the field, two sets of GOMs were developed: matching and identification measures. The skill of matching was considered less challenging than the skill of identification. Within the two sets, three types of GOMs were created: pictures, upper-case letters and sight words. Pictures used were original black and white drawings of pictures. Individual cards for the picture measures were created by putting three pictures in a row that had as little in common as possible in terms of visual resemblance, the name of the picture beginning with the same sound, etc. The combinations of three upper-case letters in the letter measures were created randomly with checks for repetitions of the same letters and letters that were visually too similar.

*Figure 1.* Example of a matching item



*Figure 2.* Example of an identification item



After examining five well known lists of sight words to use for the sight-word measures, e.g. Jerry Johns' Revised Dolch List or Thorndike Word Frequency List, the Fry's 300 Instant Sight Words were favored (only the first 100 words were used) for two reasons. First, because it is most explicitly a sight-word list as opposed to lists including mostly high-frequency words, and secondly because the Fry's 300 Instant Sight Words are clearly organized into three sets of words, and consequently it was easy to use only the first part without creating our own guidelines for doing so. The words on the cards were randomly selected from the first 100 Fry's Instant Sight Words with checks for repetitions and visually similar words. The font chosen for

the letter and sight-word measures followed the font most typically used in early elementary grades. The layout of the measures had the following guidelines. It had to fit onto an 8.5 x 11 inch sheet, there needed to be sufficient space around the items to be easily able to discriminate among them, and the layout needed to be as uniform across all the cards and across all the measures in order to keep the visual stimulation for the students in the study as consistent as possible.

*Measure administration and data collection.* All the GOMs were individually administered by a data collector at a desk and in a quiet area, if possible. Data collectors were graduate students in education or psychology who were previously unknown to the students. Each student in the study was given only half of the GOMs. The measures were divided into the following two groups. Set 1 consisted of Picture matching, Letter identification, and Word identification, while set 2 included Picture identification, Letter matching, and Word matching. The data collector placed cards in front of the student one at a time, starting with the model card. Following the administration directions for a particular GOM, the data collector modeled the task. Next, the data collector made sure the student was attentive to the task and was able to point to an item on the card. When shown the matching measures, the student was instructed to match one of the 3 choices with the item in the box by pointing. When administered the identification measures, the student was asked to point to an item on the card that was pronounced out loud by the data collector. For the two practice cards, if the student pointed to a correct item within 5 seconds as measured with a timer, the data collector administered the second practice card. If the student pointed to an incorrect item or did not point to any item, the data collector followed the prompting system described below until the student pointed to the correct practice item. If the student pointed to the correct item on the second practice card within

5 seconds as measured with a timer, the data collector administered the first test item, while starting the tape recorder functioning as a timer (a tape with recorded 3, 5, 7, and 10 minute markers). As the student was responding to the two practice cards and the set of 27 test cards, the data collector recorded the responses on a scoring sheet. All the cards for all GOMs were always administered in the same order. Whenever possible, a second data collector was present to shadow the first data collector in recording the student's responses to check for accuracy.

The CREVT-2 and PPVT-III were individually administered according to the standardized directions of the tests in the same setting as the GOMs by the primary data collector only. These two criterion measures were not timed. The special education teachers in the study were given the Checklist to complete for all their students before data collection ended. Additional data on the students in the study was collected from the district data base in the form of demographic data, such as grade, SES and ELL status, and also IEP goals and objectives in reading.

*Scoring sheet.* Student responses on the GOMs, including basic demographic information were recorded on scoring sheets. Every GOM administration required a separate scoring sheet. The scoring sheet was common across all the GOM measures. It contained a list of GOMs to check for the one administered, space for information about the student, the date of measure administration, and the name of the data collector administering the measure as well as the observer. On the scoring sheet, there was a line for each of the 30 items administered with scoring and prompt level options to circle. On the front page, the data collector circled 0 or 1 for an incorrect or correct response on the two practice items (cards 2 and 3) along with a level of prompt the student needed to make a correct response (0, 1, 2, or 3). On the next two pages, the data collector recorded 0 or 1 for incorrect or correct responses and 0, 1, 2, or 3 for the level of

prompt used for each test item in case the student did not respond to a card. The total number of test cards was 27.

*Prompting system.* In order to ensure that all students in the study were able to respond to the items on the GOMs, a four-level prompting system was developed for this purpose. Prompting systems used in other states, e.g. Colorado and Massachusetts, as well as the prompting system incorporated in the Developmental Assessment for Individuals with Severe Disabilities (DASH-2) were examined before creating the one for this study. The prompting system used in this pilot study consisted of four levels. Level 0 prompt stands for a non-prompted response. If the student responds correctly (in practice items only) or incorrectly to an item, the data collector can present the next card without having to use a prompt. If the student does not respond or responds incorrectly (practice items only) or does not respond (test items), level 1 prompt is used. Level 1 prompt is a verbal prompt where the data collector repeats the instructions already given once. If, at this point, the student has still not responded correctly (practice items only) or has not responded at all to the card, level 2 prompt is implemented. Level 2 prompt has a verbal and a gesture component. The data collector repeats the instruction: “This picture/letter/word is X/Y/Z. Point to the picture/letter/word that says X/Y/Z.” as he or she points to a correct item on the card. If the student is not able to respond correctly (practice items only) or at all to a card, the data collector implements level 3 prompt that consists of a verbal as well as a partial physical component. The data collector repeats the instruction and guides the hand of the student by holding their elbow to point to the correct item.

*Scoring.* During data collection, the data collector marked correct or incorrect responses by circling “1” for a correct response and “0” for an incorrect response. The prompt level was also marked by circling a number between 0 and 3 on each GOM scoring sheet. A line was

drawn on the sheet at 3, 5, 7, and 10 minutes. If the student finished before 10 minutes, the finish time was also recorded. The GOMs were scored by counting and recording the number of correct responses for each time frame on the scoring sheet. The number of each level of prompt (0-3) used was also recorded on the scoring sheet. Most of the time, however, a prompt was not required. The Checklist was scored by counting the number of “yes” and “no” responses for each of the six subscales and in total. The CREVT-2 and PPVT-III criterion measures were scored according to standardized published directions.

*Analysis.* The number of correct responses was used as the unit of analysis for the GOMs. The data was analyzed in two ways. The correct responses given without the implementation of a prompt were used for analysis. The number of “yes” responses was used as a unit analysis for the Checklist. Standard scores were used for analysis for both the CREVT-2 and PPVT-III. The data was analyzed using descriptive and as well as inferential analysis. Means and standard deviations for all time frames of the GOMs, the Checklist, CREVT-2, and PPVT-III were computed. Frequencies were calculated for IEP objectives and Checklist subscale items. Spearman correlations were computed between the GOM measures and all criterion measures to establish criterion validity for the GOM measures.

## Results

This study addressed two sets of research questions. The first set of questions aimed at examining the suitability of the format, administration procedures and duration of GOMs, and also the suitability of the criterion measures for students with significant cognitive disabilities. The second set of questions examined the usefulness and reliability of the data generated with these measures. In addition, the alignment between the GOMs, the Checklist, and the students’ IEP goals and objectives was studied.

The first set of questions is addressed in this section. Based on the data collectors' experience in working with the students as well as on the data produced in the pilot study, the format of the GOMs, i.e. 8.5 x 11 inch laminated cards, was suitable. A similar conclusion can be made about the administration directions for the GOMs. The students were able to follow the administration directions as conveyed by the data collector and to respond to the testing items of the GOMs either with or without the use of the prompting system. In most cases, prompts at levels 1, 2, and 3 were not required for the test items. In total, prompts at levels 1, 2, or 3 were provided 25 times across students and across GOMs by the data collector. Out of 993 opportunities or GOM cards given to students with the possibility of applying a prompt if needed, 25 opportunities that were actually utilized represent approximately three percent. Thus, a prompt was provided only on three percent of the total of GOM cards administered. Level-1 prompt was given 18 times, level-2 prompt was given three times, and level-3 prompt four times. The GOMs on which the students were prompted most often were letter identification and letter matching. The GOMs that did not elicit any or only minimal number of prompts were word identification and picture matching. It needs to be noted, however, that the prompts across GOM measures tended to be given to the same students rather than being distributed across students. As shown in detail in Table 6, the duration of GOMs was a complex issue in the study. Certain durations, i.e. 3 minutes, worked better than others, i.e. 7 and 10 minutes. The format of the criterion measures administered to the students, i.e. CREVT-2 and PPVT-III, was also appropriate. Special education teachers were able to fill in the third criterion measure, i.e. the Checklist, in a meaningful way.

The following section addresses the second set of questions, regarding the aspect of usefulness and reliability of the data produced by the piloted GOMs for the students involved.

Descriptive statistics for the GOMs and criterion measures in the form of means, standard deviations and student sample sizes are reported in Table 2. The unit of analysis in Table 2 is the number of correct responses made without having to apply the prompting system. Due to our design, i.e. half of our sample completed one set of measures and the other half the other set, only the GOMs with a sample size of three or larger are reported in Table 2. The decrease in sample size with an increase in allocated time resulted from some students finishing the measure in a shorter than given time, i.e. ceiling effect. Despite the evidence of a ceiling effect on some of the measures for some students, Picture matching, Letter matching, Picture identification and Word identification measures showed an average increase in the number of correctly matched or identified items with an increase in allocated time. A trend in the spread around the mean (SD) is harder to detect because of the variability in sample size. When the descriptive statistics for GOM scores with and without prompt are compared, the mean scores for scores without prompt tend to decrease, while the spread around the mean tends to increase. All but one student were able to complete the two standardized tests, the Comprehensive Receptive and Expressive Vocabulary Test (CREVT- 2) and the Peabody Picture Vocabulary Test (PPVT-III). The CREVT-2 average standard score was higher with a smaller standard deviation than the average PPVT-III score. The teachers completed the RIPM Early Literacy Knowledge and Reading Readiness Checklist (the Checklist) for all students in the study. The average number of positive answers was approximately 23 out of 55. The spread around the mean for this teacher-completed measure was approximately 11 answers.



Table 2

*Means and Standard Deviations Adjusted for Prompting Level*

Measure	Mean	SD	N
Picture matching 3min correct	18.00	8.54	7
Picture matching 5min correct	19.75	2.06	4
Picture matching 7min correct	24.67	4.04	3
Letter matching 3min correct	15.00	10.22	6
Letter matching 5min correct	15.75	9.91	4
Letter matching 7min correct	17.00	10.54	3
Word matching 3min correct	16.20	10.06	5
Word matching 5min correct	15.33	8.08	3
Word matching 7min correct	19.33	6.81	3
Picture identification 3min correct	15.40	10.31	5
Picture identification 5min correct	18.00	7.39	4
Picture identification 7min correct	20.33	5.86	3
Letter identification 3 min correct	19.00	7.96	7
Letter identification 5 min correct	17.00	10.00	3
Word identification 3min correct	14.67	8.89	6
Word identification 5min correct	17.20	9.09	5
CREVT-2 SS	69.42	7.56	12
PPVT-III SS	54.33	9.12	12
Checklist total "yes"	22.77	10.56	13

*Note:* CREVT- 2 SS = Comprehensive Receptive and Expressive Vocabulary Test standard score; PPVT-III SS = Peabody Picture Vocabulary Test standard score

Inter-rater reliability was calculated between the persons administering the GOM measures and the observers. Sixty percent of observations were checked for reliability. The reliability coefficient was 100%. It needs to be noted, however, that the observers' role was not

solely for reliability purposes but also to check the person administering the measures for administration errors. It was deemed necessary in a pilot study to have the presence of an observer to collect additional data. One of the roles of the observer was to ensure that the data collected was reflective of the students' ability as much as possible and not distorted by their behaviors or errors in recording the data due to such behaviors.

The next paragraph addresses in greater detail the duration of GOMs and the ceiling effect briefly described in the descriptive statistics section. On average, students finished all GOMs before the time limit given for each measure, i.e. 10 minutes. Each measure consisted of 27 timed test cards. Letter identification showed the shortest time of completion, while Picture identification the longest. The average finishing time for the matching and identification measures was similar, i.e. 5 minutes and 22s for matching and 5 minutes and 26 seconds for identification. Word matching and Picture identification measures showed the largest difference between the youngest and oldest students (taking those particular measures). Thus, ceiling effect occurred for all the GOMs at all ages when considering the total administration time of 10 minutes. Picture matching, Word matching, Picture identification, and Letter identification showed a ceiling effect even for the shortest time recorded, i.e. 3minutes, for at least one of the age levels, usually the oldest students in the study.

The relationship between the piloted GOMs and the three criterion measures, i.e. PPVT-III, CREVT-2 and the Checklist, in this pilot study was explored using nonparametric statistics. More specifically, the relation between the number of correct items on the GOMs and criterion measures was examined. Another factor that encourages caution when interpreting the results representing the relationship between the GOM and criterion measures is the presence of a ceiling effect on all the GOMs in the case of 10-min administration and some at the other

intervals. Consequently, only the nonparametric (Spearman) correlation coefficients for the GOM 3-min administration are reported in Table 3. The reported correlation coefficients ought to be treated only as indicators of the possible relation between measures needing further exploration. Results in Table 4 were calculated using correct scores without having to use a prompting system. Keeping in mind the small size of the sample and thus potential for sampling error, the most stable correlations occurred between the GOMs and the Checklist and the least stable between the GOMs and the CREVT-2. The correlations between the GOMs and the Checklist ranged from .64 for Letter matching 3min and 1.00 for Picture identification 3min. The correlations between Word matching 3min (.46), Letter identification 3min (.65) and Word identification 3min (.61) and the PPVT-III in Table 4 suggest a potential relationship between PPVT-III and some aspects of reading or pre-reading as assessed by the GOMs.

Table 3

*Correlations of GOM Measures 3 min with CREVT-2, PPVT-III and the Checklist Adjusted for Prompting Level*

GOM measure	CREVT-2	PPVT-III	Checklist	N
Picture matching 3min correct	.02	-.06	.74	7
Letter matching 3min correct	-.79	.03	.64	6
Word matching 3min correct	-.98	.46	.90	5
Picture identification 3min correct	-.82	.21	1.00	5
Letter identification 3 min correct	.00	.65	.92	7
Word identification 3min correct	-.46	.61	.84	6

In order to put the results into a broader context of reading instruction for the students in the pilot study, their Individualized Education Plans (IEP) goals and objectives were reviewed.

It was found that goals and objectives tended to vary across teachers but were similar across students within a teacher. The objectives were described in more detail in the students' IEPs than goals, and consequently the objectives became a richer source of information. Table 4 lists the IEP objectives as written by the teachers, along with the frequencies with which the objectives were written, as well as their alignment with the related Checklist items positively answered by the teachers, and also GOMs. Eleven objectives related to reading or pre-reading skills were identified by four special education teachers. All but one objective, i.e. Identify sight words from a reading program, matched the items on the Checklist. At the same time, not all items on the Checklist were covered by the IEP objectives. The relation between the two variables as computed using the Spearman correlation coefficient was .48 ( $p < .17$ ). However, when the item concerning matching letters was removed, the correlation coefficient increased to .91 ( $p < .001$ ). On all but the "matching letters" item, the number of students for whom the special education teachers identified reading IEP objectives was very similar to the number of students for whom the teachers positively answered the related item on the Checklist. In contrast, the "matching letters" item as an IEP objective was recorded for only 1 student, while it was recorded on the Checklist that 12 students mastered this skill. Because of the discrepancy between IEP objectives written and positive answers on the Checklist for the item "matching letters", this item was removed to recalculate the relationship between IEP objective and positive answers on the teachers Checklist. One explanation for why teachers did not write objectives for "matching same beginning letters in words" but answered that students have mastered the skill of matching upper-case letters is that they perceived these two tasks as different. "Matching identical upper-case letters" could have been considered as something relatively easy for the students, while "matching same beginning letters in words" could be a more complex task and

might not be a teaching priority for students with significant cognitive disabilities, similarly to writing objectives for “identifying beginning sounds”. The information in Table 6 also shows that pictures were not included in the IEP objectives. However, there were items on the Checklist that are not included in the table containing information about pictures, e.g. “If given a picture book and asked to read the book, can refer to pictures for clues”. Consequently, the alignment between the two GOMs, Picture matching and Picture identification, was missing because of the lack of IEP objectives involving pictures.

In summary of the results, the data produced by relating the GOMs and criterion measures in this pilot study can inform the assessment process for students with significant cognitive disabilities despite its preliminary nature. The format and administration procedures worked well for the pilot study sample. The first and most obvious limitation of the data produced in this pilot study was the ceiling effect on the scores resulting from a small number of test cards (27) in all of the GOMs, in combination with too much time (duration) allocated to the administration of these measures (7 and 10 minutes). The second limitation is demonstrated by the results produced by the receptive subtest of CREVT-2. In comparison to the other two criterion measures, the Checklist and the PPVT-III, the CREVT-2 produced the most unstable results that were difficult to interpret. In spite of these limitations, the results suggested that the GOMs examined in this pilot study could be potentially used to assess the academic performance of students with mild to moderate cognitive disabilities. First, the students in the study were able to respond to the GOMs without or with the prompting system. Second, the evidence suggests not only a strong positive relationship between the GOMs and the Checklist, but also between the IEP objectives and the related Checklist items (less one item). This indirect link between the IEP objectives and the GOMs implies that the content of the GOMs corresponds to the academic

skills in reading and pre-reading skills perceived as important by, at least some, special education teachers when planning instruction for their students.

Table 4

*Alignment of IEP Objectives, GOM Measures, and the Checklist Subscale Items*

IEP Objectives	Frequency	Checklist Subscale Items	Frequency	GOM Measure
Identify upper and lower case letters	7	Can identify upper and lower case letters	9	Identification: Upper case letters
Identify consonant letter sounds	5	Can identify letter sounds	6	Identification: Sight words
Identify sight words	11	Can identify words by sight	9	
Identify sight words from a reading program	2			
Read sight words in sentence, passage	7	Can read a passage containing sight words	7	
Identify beginning sounds	1	Can identify same beginning sounds	4	
Identify middle, ending sounds	2	Can identify same ending sounds	3	
Answer questions about story	3	Can answer questions about a passage	6	
Match same beginning letters in words	1	Can match identical upper case letters	12	Matching: Upper case letters
Decode short vowel words	2	Can decode short vowel words	2	
Decode long vowel words	1	Can decode long vowel words	0	

## Discussion

The importance of reading and assessing reading performance and progress for all students is clear. However, strategies for assessing students with significant cognitive disabilities as long been a challenge as has determining the appropriateness of certain curriculum approaches - functional or academic. The standards movement has had an impact in this area: All students must be included in the accountability system and all students must meet standards that align with grade level standards in reading but how?

Chall (1996) provides a framework for thinking about reading development that allows for such development to be decoupled from age, which suggests that even an older student may be at a very early stage of reading development. This perspective supports the approach we have examined in our research – that general outcome measures (GOMs) can be used to measure *early* literacy development with *older* students who have significant cognitive disabilities. The hypothesis was that such measures could be created to measure students' performance in academic areas aligned with state standards as well as IEP goals; ultimately, providing teachers with a tool to measure individual annual growth. First, some primary practical and technical characteristics of such measures needed to be established. The research questions posed for this pilot study were focused on two areas: The first set of research questions addressed the suitability of the format, administration procedures and duration of GOM measures, including the suitability of the criterion measures for students with significant cognitive disabilities, while the second set of questions examined the usefulness and reliability of the data produced using these measures.

The present research was conducted as a pilot study intended to determine if using newly developed general outcome measurement could potentially work to assess students' performance



in reading or early literacy. While the sample was small and findings must be interpreted carefully, the results are positive. Examination of measure format and administration, criterion validity, and reliability provided enough support for researchers to suggest the need for further technical adequacy and progress studies. The results suggest GOMs may be an appropriate way to measure the performance of students with significant cognitive disabilities in an academic area, such as reading.

#### *Measure Development and Administration*

Curriculum, assessment and expected progress in academic content for students with significant cognitive disabilities are areas that have largely remained unexamined. Recent federal and state requirements to ensure all students are progressing and meeting state standards draws attention to students with such challenges. While portfolios and mastery monitoring strategies, e.g. a checklist, have been used as alternate assessments in some states, e.g. Massachusetts and Nebraska respectively, each experiences its challenges. In the case of portfolios, two of the biggest drawbacks are the time spent creating a portfolio and using a portfolio for measuring student progress. The main challenge of using a mastery monitoring approach is again measuring student progress that goes beyond a single skill. The goal of this pilot study was to develop general outcome measures for students with significant cognitive disabilities that were time efficient, were reliable and valid, and had the potential of measuring student progress across time. Using Chall's model of reading development and previous experience with curriculum-based measurement (CBM), it was anticipated that general outcome measures (GOMs) could be developed and used with students with significant cognitive disabilities. There were many things to consider, such as timing of measures, mode of response,

student's verbal ability, etc. The pilot study provided empirical support for using these newly created GOM measures.

Specifically, the administration of the measures using laminated cards that required only a "pointing" response worked well for students to engage in the task. While we started with 10-minute measurement intervals, it was clear that students could generate responses within 3 minutes. In fact, some students obtained a ceiling effect given too much time and not enough cards. Therefore, analyses were conducted using the 3-minute samples. Up to a point (given the ceiling effect), the more time given to complete the task, the greater number of correct responses students gave, which suggests times measures can be used with students with significant cognitive disabilities. The students required demonstration, training and practice in order to appropriately respond to the measures; however, very little formal prompting was needed with this sample of students.

It can be concluded that the format of measures used in this pilot study included appropriate stimulus material and procedures for students with significant cognitive disabilities. In fact, in follow-up meetings with teachers, they were eager to begin using the measures. They expressed gratitude and excitement at the potential of the measures. Early on, the teachers were not certain such an "academic" measure would work with the students. At the same time, the students' performance on the GOMs had a strong relationship with how their teachers assessed the students' knowledge and skills, using the Checklist (.64 – 1.0). This was an interesting discrepancy: Initially, the teachers did not have high hopes for the students' performance on the GOMs, yet their assessment of the students ultimately matched well with the students' actual performance on the measures. The teachers began to imagine a way, using GOMs, to measure

students' academic (reading) performance that was data-based and objective rather than perceptual and subjective.

### *Criterion Validity and Reliability*

Preliminary reliability and validity results from this pilot study indicate that it is worth putting more effort into examining GOM measures for students with significant cognitive disabilities. Inter-rater reliability was 100 percent, indicating raters are able to use the measures with the same results, keeping in mind they consulted each other on occasion. Other forms of reliability need to be assessed in the future, such as test-retest, etc. The relation between the GOMs and teacher-completed Checklist as well as the PPVT-III to an extent (Letter identification at 3 minutes, .65; and Word identification at 3 minutes .61, suggests that the newly created GOMs relate well to teacher judgment and also to a standardized measure of vocabulary/language development necessary for beginning reading). Furthermore, it was found that teacher written IEP objectives for the students in the study related highly with the teacher filled Checklist at the end of the year, suggesting that the items included in the Checklist were relevant to what teachers focus on with their students. While additional work needs to be done with a larger number of students to ensure the technical adequacy of the GOMs, these findings lead us to believe further work is worth pursuing.

### *Limitations*

Perhaps the greatest limitation to the generalization of our results is the size of the study sample. The resource needs required for the initial study were prohibitive but findings have helped to narrow the requirements. For example, 10 minutes is not needed to get an adequate response rate; therefore, we have reduced the time in a follow-up to 5 minutes and may reduce that further. This example also illustrates an additional limitation, which is the ceiling effect

obtained in the study that impacts interpretation of the results. One of our criterion measures did not seem appropriate for our student sample. The CREVT-2 did not produce interpretable results and limits our ability to examine a potential relationship.

#### *Further Research Needed*

Even with its limitations, this study does serve its purpose as a pilot study. It provides us with enough information to suggest there may be a relationship between the GOMs and the criterion measures, whether they are teachers' perspectives or a standardized assessment typically used in assessing aspects of literacy and reading ability. The results are encouraging and suggest a need for further research. Additional research is needed to examine the technical adequacy of these and other potential measures (reading, writing, math); to determine how these measures work with typically developing students facing early literacy; to study how these measures might work with students who have severe cognitive disabilities; to examine if the measures can be used to assess progress over time; and to identify components of teacher use of GOM measures for these students, to name just a few.

General outcome measures (GOMs) in literacy can help students with significant cognitive disabilities, and the teachers who are assessing their performance, demonstrate their knowledge and skill in literacy and the academic area of reading in a way that is fast, objective and data-based. Further studies are needed to establish the technical characteristics of these measures. Once GOMs for students with significant cognitive disabilities are established as reliable and valid, further research is needed into whether these measures are sensitive to growth and progress over time. With sound technical properties and being time efficient, these measures have a great potential to give special education teachers an indication throughout the year what

their students' performance in reading might look like at the end of the year as shown by alternate assessments.

Author Note

Teri Wallace and Renáta Tichá are at the Institute on Community Integration, University of Minnesota. We wish to thank DCD teachers in Minneapolis Public Schools, MN.

Address correspondence to Teri Wallace, 111A Pattee Hall, 150 Pillsbury Dr. SE, Minneapolis, MN, 55455, walla001@umn.edu.

The Research Institute on Progress Monitoring at the University of Minnesota is funded by the U.S. Department of Education, Office of Special Education Programs (Award H324H030003) and supported the completion of this work.

## References

- Allinder, R. M., & Swain, K.D. (1997). An Exploration of the Use of Curriculum-Based Measurement by Elementary Special Educators. *Diagnostique*, 23(2), 87-104.
- Browder, D.M. (2001). *Curriculum and assessment for students with moderate and severe disabilities*. New York: Guilford Press.
- Browder, D. M., Courtade-Little, G., Wakerman, S., & Rickelman, R.J. (2006). From sight words to emerging literacy. In: D. M. Browder & F. Spooner (2006). *Teaching Language Arts, Math, & Science to Students with Significant Cognitive Disabilities*. Baltimore: Paul H. Brookes.
- Browder, D. M., & Minarovic, T. (2000). Utilizing sight words in self-instruction training for employees with moderate mental retardation in competitive jobs. *Education and Training in Mental Retardation and Developmental Disabilities*, 35, 78-89.
- Browder, D. M. & Spooner, F. (2006). *Teaching Language Arts, Math, & Science to Students with Significant Cognitive Disabilities*. Baltimore: Paul H. Brookes.
- Browder, D. M., Wallace, T., Snell, M., & Klienert, H. (2005). *A White Paper: Progress Monitoring for Students with Significant Cognitive Disabilities*. DC: AIR.
- Browder, D. M., & Xin, P. Y. (1998). A meta-analysis and review of sight word research and its implications for teaching functional reading to individuals with moderate and severe disabilities. *Journal of Special Education*, 32, 130-153.
- Calhoun, M.B., & Fuchs, L.S. (2003). The Effects of Peer-Assisted Learning Strategies and Curriculum-Based Measurement on the Mathematics Performance of Secondary Students with Disabilities. *Remedial and Special Education*, 24, 235-245.

- Chall, J.S. (1996). *Stages of reading development* (2nd Ed.). Fort Worth, Texas: Harcourt Brace.
- Collins, B. C., Branson, T. A., & Hall, M. (1995). Teaching generalized reading of cooking product labels to adolescents with mental disabilities through the use of key words taught by peer tutors. *Education and Training in Mental Retardation and Developmental Disabilities, 30*, 65-75.
- Collins, B.C., & Griffen, A.K. (1996). Teaching students with moderate disabilities to make safe responses to product warning labels. *Education and Treatment of Children, 19*, 30-45.
- Connors, R. A. (1992). Reading instruction for students with moderate mental retardation: Review and analysis of research. *American Journal of Mental Retardation, 103*, 1-11.
- Downing, J.E. (2005). *Teaching Literacy to Students with Significant Disabilities*. Thousand Oaks, CA: Corwin Press.
- Espin, C.A. & Deno, S.L. (1994-95). Curriculum-based measures for secondary students: Utility and task specificity of text-based reading and vocabulary measures for predicting performance on content-area tasks. *Diagnostique, 20*, 121-142.
- Espin, C., Shin, J., Deno, S. L., Skare, S., Robinson, S., & Benner, B. (2000). Identifying Indicators of Written Expression Proficiency for Middle School Students. *Journal of Special Education, 34*, 140-153.
- Espin, C.A., Wallace, T., Campbell, H., Lembke, E.S., Long, J.D., & Ticha, R. (2005). Predicting the success of secondary-school students on state standards tests: Validity and reliability of Curriculum-Based Measures in written expression.
- Foegen, A., & Deno, S. L. (2001). Identifying growth indicators for low-achieving students in middle school mathematics. *Journal of Special Education, 35*(1), 4-16.



Fuchs, L. S., & Fuchs, D. (2002). Curriculum-Based Measurement: Describing Competence, Enhancing Outcomes, Evaluating Treatment Effects, and Identifying Treatment Nonresponders, *Peabody Journal of Education*, 77, 64-84.

Good, R. H. I., & Kaminiski, R. A. (1996). Assessment of instructional decisions: Toward a proactive/prevention model of decision-making for early literacy skills. *School Psychology Quarterly*, 11, 326-336.

Greenwood, C. R., Tapia, Y., Abbott, M., & Walton, C. (2003). A building-based case study of evidence-based literacy practices: Implementation, reading behavior, and growth in reading fluency, K-4. *Journal of Special Education*, 37, 95-110.

Groff, P., Papp, D., & Flood, J. (1998). Where is the phonics? Making a case for its direct and systematic instruction. *The Reading Teacher*, 52, 138-141.

Gurry, S. E., & Larkin, A. S. (1999). Literacy learning abilities of children with developmental disabilities: *What do we know? Currents in literacy*.

Hintze, J.M., Ryan, A. L., & Stoner, G. (2003). Concurrent Validity and Diagnostic Accuracy of the Dynamic Indicators of Basic Early Literacy Skills and the Comprehensive Test of Phonological Processing. *School Psychology Review*, 32, 541-556.

Individuals with Disabilities Education Improvement Act of 2004, P. L. No. 108-446, 20 U.S.C. section 611-614.

- Joseph, L. M., & Seery, M. E. (2004). Where is the phonics? A review of the literature on the use of phonetic analysis with students with mental retardation. *Remedial and Special Education, 25*, 88-94.
- Katims, D. S. (2000 a). *The Quest for Literacy: Curriculum and Instructional Procedures for Teaching Reading and Writing to Students with Mental Retardation and Developmental Disabilities*. MRDD Prism Series, Volume 2.
- Katims, D.S. (2000 b). Literacy instruction for people with mental retardation: Historical highlights and contemporary analysis. *Education and Training in Mental Retardation and Developmental Disabilities, 36*, 363-372.
- Kliewer, C. & Landis, D. (1999). Individualizing literacy instruction for young children with moderate to severe disabilities. *Exceptional Children, 66*, 85-100.
- Lembke, E., Deno, S. L., & Hall, K. (2003). Identifying an Indicator of Growth in Early Writing Proficiency for Elementary School Students. *Assessment for Effective Intervention, 28*(3-4), 23-35.
- No Child Left Behind Act of 2001. (2002). Pub. L. No. 107-110, 115 Stat. 1425
- Otaiba, S.A., & Hosp, M. (2004). Providing effective literacy instruction to students with Down syndrome. *TEACHING Exceptional Children, 36*, 28-35.
- Roach, A.T., & Elliott, S. N. (2006). The influence of access to general education curriculum on alternate assessment performance of students with significant cognitive disabilities. *Educational Evaluation and Policy Analysis, 28* (2), 181-194.

Shinn, J., Deno, S. L., & Espin, C. (2000). Technical Adequacy of the Maze Task for Curriculum-based Measurement of Reading Growth. *Journal of Special Education, 34*, 164-172.

Tindal, G., McDonald, M., Tedesco, M., Glasgow, A., Almond, P., & Crawford, L. (2003). Alternate assessments in reading and math: Development and validation for students with significant disabilities. *Exceptional Children, 69*, 481-494.

U.S. Department of Education. *No Child Left Behind*.

<http://www.ed.gov/nclb/>

U.S. Department of Education (2006). The 26th Report to Congress on the Implementation of the Individuals with Disabilities Education Act. Washington, D.C.: Author.

Wallace, G., & Hammill, D.D. (2002). *Comprehensive Receptive and Expressive Vocabulary Test: Examiner's Manual* (2<sup>nd</sup> ed.). Austin, TX: Pro-ed.

Williams, K.T., & Wang, J. (1997). *Technical References to the Peabody Picture Vocabulary Test – Third Edition (PPVT-III)*. Circle Pines, MN: AGS.