

FOCUS ON EXCEPTIONAL CHILDREN

Assessing the Writing of Struggling Learners

Natalie G. Olinghouse and Tanya Santangelo

The importance of writing both in and out of school cannot be overestimated. Academic writing is an essential part of the K–12 curriculum—students write to demonstrate, support, and deepen their knowledge and understanding of themselves, their relationships, and their world (Bangert-Drowns, Hurley, & Wilkinson, 2004; Graham, 2006; Graham & Perin, 2007; Keys, 2000; Sperling & Freedman, 2001). As students move into work settings, writing skills are necessary to gain employment and receive promotions (National Commission on Writing, 2004). In the community, teens and adults use email and other electronic writing for communication, daily social interaction, self-exploration and expression, and reflection on current events (National Commission on Writing, 2008).

Despite the importance of writing, data from the National Assessment of Educational Progress (NAEP) indicate that many students in the United States struggle as writers. In 2007, only 33% of the nation's eighth graders and 24% of the nation's 12th graders scored at or above grade-level proficiency in writing, while in 2002, 28% of fourth graders scored at or above proficiency (Persky, Daane, & Jin, 2003; Salahu-Din, Persky, & Miller, 2008). In 2007, students with disabilities demonstrated even greater difficulty with writing, as less than 7% of eighth graders and less than 5% of 12th graders with LD received scores of proficient or above.

One reason students may have difficulty acquiring and developing good writing abilities is that writing is a complex task requiring the coordination of different skills, processes, and types of knowledge. The complexity of writing means that struggling writers may exhibit difficulties with the process or products of writing or both (De La Paz, Swanson, & Graham, 1998; Graham & Harris, 1996, 2000; Harris & Graham, 1999; Zimmerman & Reisemberg, 1997). Students may have problems in multiple areas of writing or have trouble with only a few areas of writing (Juel, 1988; Whittaker, Berninger, Johnston, & Swanson, 1994). Different student profiles require different approaches to intervention; good assessment practices identify specific areas of concern for each student. While all students benefit from good instructional decisions based on assessment, students who exhibit writing difficulties have special needs that require targeted intervention. A teacher who fully understands a student's specific strengths and needs can design a better instructional program for students using a multitude of assessment tools.

Dr. Olinghouse is an assistant professor in the Department of Educational Psychology at the University of Connecticut. Dr. Santangelo is an associate professor in the Department of Education at Arcadia University.

This article is designed to support teachers in becoming more proficient in writing assessment. First, we discuss the different purposes that guide the assessment process. We then describe the range of skills involved in writing, as well as common difficulties students may have in these areas. The main body of the paper details specific writing assessment methods and tools for a range of writing skills.

WHAT IS ASSESSMENT?

The terms *assessment* and *testing* are often used interchangeably, causing teachers to use abbreviated or ineffective assessment procedures. Assessment is defined as a "process of collecting data for the purpose of making decisions about students or schools" (Salvia, Ysseldyke, & Bolt, 2010, p. 4). It is important to pay attention to a few key terms in this definition. First, assessment is a *process*—it is not a one-time or one-test activity. Second, the assessment process begins with a need for more information to *make a decision* about a student or group of students. School per-

sonnel make decisions about problems: Is there a problem? What is the specific problem? How do we solve the problem? Third, the purpose of the decision guides the whole assessment process, including the *collection of data*. Data are most meaningful and useful when they come from a variety of sources rather than just one source. Data can come from a multitude of tools (what is used to collect data) and methods (how the tool is administered and scored), including tests, observations, judgments, and recollections.

Purposes for Writing Assessment

While educators can choose from an array of writing assessment tools and methods, the purpose of assessment should guide the assessment process. In short, there is no "one size fits all" writing assessment. The best writing evaluation is aligned with the purpose of assessment.

Screening

The purpose of screening is to quickly identify students who are at risk for or are currently experiencing difficulties in a specific academic area. General outcome measures are usually used during screening, which broadly measure the academic area in question. Screening assessments are typically administered three to four times a year and are quick and easy to administer and score. This schedule helps teachers identify students who may be experiencing difficulties at the beginning of the year and then identify students who are not developing as expected throughout the school year. Students who are performing below expectations may need more frequent monitoring or require additional assessment to gain a better understanding of their needs.

Instructional Planning and Modification

Teachers frequently use assessment data to help plan and modify instruction. An assessment plan can tell teachers what needs to be taught, how it can best be taught, and what kind of expectations are realistic for a student or group of students. The data required for this kind of decision making must provide a more in-depth profile of a student's skills and needs. Accordingly, this kind of assessment takes longer to administer and score and often addresses multiple aspects of writing. Results can be used to indicate specific areas to target for instruction, group students with similar instructional needs, or place students in specific remedial instruction.

Summative assessment occurs at the end of a course or instructional unit. Summative assessment measures whether a student has mastered the outcomes of the instruction. While prevalent in schools, these can be least valuable in terms of instructional planning and modification if teachers do not use this information to guide future instruction, set new instruction goals, or provide feedback to the student in a way that can be acted upon.

FOCUS ON EXCEPTIONAL CHILDREN

ISSN 0015-511X

FOCUS ON EXCEPTIONAL CHILDREN (USPS 203-360) is published monthly except June, July, and August as a service to teachers, special educators, curriculum specialists, administrators, and those concerned with the special education of exceptional children. This publication is annotated and indexed by the ERIC Clearinghouse on Handicapped and Gifted Children for publication in the monthly *Current Index to Journals in Education* (CIJE) and the quarterly index, *Exceptional Children Education Resources* (ECER). The full text of *Focus on Exceptional Children* is also available in the electronic versions of the *Education Index*. It is also available in microfilm from Serials Acquisitions, National Archive Publishing Company, P.O. Box 998, Ann Arbor, MI 48106-0998. Subscription rates: individual, \$48 per year; institutions, \$66 per year. Copyright © 2010, Love Publishing Company. All rights reserved. Reproduction in whole or part without written permission is prohibited. Printed in the United States of America. Periodical postage is paid at Denver, Colorado. **POSTMASTER:** Send address changes to:

Love Publishing Company
Executive and Editorial Office
P.O. Box 22353
Denver, Colorado 80222
Telephone (303) 221-7333

CONSULTING EDITORS

Steve Graham
Vanderbilt University

Ron Nelson
University of Nebraska-Lincoln

Eva Horn
University of Kansas

Carrie E. Watterson
Senior Editor

Stanley F. Love
Publisher

Progress Monitoring

Several kinds of progress monitoring can determine whether a student is demonstrating expected progress. State-level assessments monitor progress toward mastery of state standards. Other assessment is used to monitor progress toward a student's IEP goals. Teachers may also conduct short and frequent assessments to monitor progress in a specific curriculum or for specific learning objectives. Formative assessment, a term frequently used in schools, is a form of progress monitoring that uses the results from the assessment to adjust instruction.

The purpose for monitoring progress dictates the writing task and the scoring method. Teachers do not have the responsibility to develop a state-level writing assessment; however, they can develop and use assessments to monitor progress toward IEP goals or to see whether students are making progress in a specific curriculum. Often, these assessments are short and include multiple forms that can be administered frequently. The assessments also need to detect small changes in students' writing ability. If used correctly, they help the educator make changes to an instructional program more quickly.

Eligibility

Assessment data are used to make decisions about eligibility for special education services. In order to qualify for special education services, a student must meet eligibility criteria for a specific disability (which often vary from state to state) and have specific learning needs. A variety of assessment tools and methods are used to determine a student's specific learning needs. In writing, a team often will include norm-referenced tests along with writing samples from the classroom. When making eligibility decisions, it is important to assess multiple aspects of writing to gain an in-depth picture of a student's writing abilities.

WRITING PRODUCTS AND PROCESS

Given the complexity of writing, teachers may feel overwhelmed thinking about how to assess and remediate writing difficulties. Indeed, both products and process are important to developing writing abilities. Given space limitations and the availability of research-based writing assessment information, we have chosen to focus on the products of writing, while acknowledging that process and other areas also are important to assess and teach.

In this article, we use a *levels of language* approach to thinking about writing assessment. Writing, along with the other related language abilities (reading, speaking, and listening), is acquired and developed across multiple levels of language. In writing, a simple way of categorizing these levels of language is thinking about how students develop skills

at the letter, word, sentence, and text level. One might be tempted to think about this development in a linear fashion—for example, expecting students to master letter skills before sentence skills—but many of these can and should be developed simultaneously. Using this framework, we more fully describe each of these levels of language below.

Letter and Word Level

Handwriting, spelling, and vocabulary are important skills to develop in both younger and older writers. The goal of handwriting and spelling instruction is to develop fluency and automaticity; however, some struggling writers may have difficulty with these skills throughout school. The selection of vocabulary while writing plays an important role from beginning writing until adulthood.

Handwriting

Young children begin writing development by playing with scribbling and drawing. As they approach kindergarten, most children begin to form individual letters. This act, commonly known as handwriting, is an important ingredient in writing development. While handwriting legibility is important (if a reader cannot decipher the writing, the message is lost!), it appears that handwriting fluency (automaticity of getting letters down on paper) is more critical. A student may struggle with letter formation, legibility, or speed and automaticity. Difficulties with handwriting can interfere with the writing process, as the demands of letter formation minimize cognitive resources for more important processes such as planning or revising (McCutcheon, 1996). Students who struggle with handwriting may also avoid and dislike writing, slowing overall writing development (Berninger, Mizokawa, & Bragg, 1991).

Unfortunately, handwriting is often overlooked in schools. In a recent national survey, teachers reported that 23% of their students experienced handwriting difficulties (Graham et al., 2008). Students with LD have even greater difficulties with handwriting accuracy and legibility (Graham & Weintraub, 1996; Weintraub & Graham, 1998). Composing using a keyboard rather than by pencil or pen does not necessarily solve handwriting problems. In particular, students through sixth grade write more words, more ideas, and higher quality compositions when handwriting than when keyboarding (e.g., Connelly, Gee, & Walsh, 2007; Hayes & Berninger, 2010). It should be noted that students of any age with poor keyboarding skills will likely have similar difficulties when composing on a keyboard.

Spelling

Spelling development requires an understanding of sound-symbol correspondence and spelling conventions and patterns as well as a memory for regularly and irregularly

spelled words. Many struggling writers are also struggling spellers (e.g., Graham & Weintraub, 1996; Juel, 1988). As with handwriting difficulties, a student who is a poor speller may minimize the writing process or avoid writing altogether. Having to attend to the spelling of individual words can cause the writer to forget the ideas or plans for the composition (Graham, 1990; McCutchen, 1996). Students may choose to include only words they know how to spell, resulting in vocabulary that is simple or imprecise. In addition, poor spelling can cause a reader to react negatively to a composition (Chase, 1986).

Vocabulary

The selection of vocabulary is an important part of the writing process. The best word choice accurately conveys a writer's intention and consideration of the purpose, genre, and audience of the writing. Many different vocabulary measures have been shown to be related to overall writing quality (e.g., Deno, Marston, & Mirkin, 1982; Olinghouse & Leaird, 2009). These include vocabulary diversity (the number of different words used in the composition), vocabulary maturity (more unique or complex words), or word size (longer words). While it is not clear what kind of words is most important to writing quality, it is clear that vocabulary matters.

Research has shown that students with disabilities can have difficulties with vocabulary in written composition (e.g., Dockrell, Lindsay, Connelly, & Mackie, 2007; Houck & Billingsley, 1989; Scott & Windsor, 2000). Students who struggle with writing may use a less rich and simpler vocabulary in their writing. They may use the same word several times, rather than varying their word choice. A written composition may contain words below the writer's grade level, resulting in a paper that lacks precision of ideas and content.

Sentence Level

A variety of skills and knowledge are required at the sentence level. Students first need to understand the basic conventions of a sentence. Young writers, relying on their oral language, will often write long, run-on sentences. As students move into first and second grade, they are expected to understand the concept of a written sentence, along with simple capitalization and punctuation rules. These rules become more complex quickly. For example, later elementary students are expected to properly capitalize and punctuate dialogue. In addition to these writing conventions, students must understand the grammatical or syntactic structure of a sentence. This includes correct word order, subject-verb agreement, the use of clauses and phrases, and parallelism, among others. Older students need to assess the suitability of their sentences within the context of the larger text and as appropriate to the purpose, genre, and audience.

This includes writing with varied sentence lengths or sentence structures.

Struggling writers may have difficulty with basic sentence mechanics, including run-on sentences, or improper or missing capitalization or punctuation (Houck & Billingsley, 1989; Newcomer & Barenbaum, 1991). Their sentences may have several grammatical errors, such as incorrect use of verb tenses or articles (Mackie & Dockrell, 2004; Scott & Windsor, 2000). Students' writing also may contain only simple or repetitive sentence structures.

Text Level

Once students are able to write connected sentences, they are expected to put these sentences together into one paragraph and multiparagraph text. This requires attention to multiple aspects of writing. At the paragraph level, students must understand the conventions of indenting and margins. They also need to construct a paragraph that has a logical organizational structure. This structure is not consistent across genres, and so students learn about different kinds of paragraphs and paragraph structures (e.g., compare/contrast, descriptive, sequence). The concepts of thesis sentences, main ideas, and details become important. One paragraph soon expands into multiparagraph text, where organization occurs both within and across paragraphs. Transitions become necessary between paragraphs and sections.

Given the complexity of writing at this level, a student may struggle with any or all of these aspects. Some struggling writers lack any understanding of a paragraph and write one long paragraph for every composition. When there are paragraphs, they may contain multiple main ideas rather than focus on one main idea at a time. A student experiencing difficulty with writing may have a composition that lacks organization and coherence—instead jumping from topic to topic within the same paragraph. The struggling writer may be unable to change the structure of the paragraphs and larger text depending on the expected genre. For example, a student may turn every composition into a story, even when asked to write in another genre.

Process

Although we do not extensively discuss assessment of the writing process in this article, it is important to understand that many written products are composed using a writing process. The writing process consists of engagement in a recursive process of planning, drafting, revising, and editing. This process is dictated by the writing task itself: Some tasks do not require a writing process (e.g., making a grocery list), while others require substantial engagement in the writing process (e.g., writing a research paper). In each component of the writing process, the writer must have strategies for carrying out the process. The writer can use

one or more of the multiple methods of planning before writing—outlining, graphic organizers, or brainstorming, among others—depending on the purpose of the planning and the type of writing task.

The writing process appears to be substantially limited or nonexistent in young and struggling writers (e.g., Englert, Raphael, Fear, & Anderson, 1988; Graham & Harris, 1997). For example, a student may neglect any planning and instead just start drafting. This lack of idea development and organization before writing often leads to disorganized compositions that jump from idea to idea. Revision also appears to be quite difficult for many writers. Struggling writers rarely make substantial revisions, such as clarifying ideas or reorganizing text; instead, they focus on surface level changes that correct spelling errors or illegible handwriting (Graham, 1990; Graham & Harris, 1997; Graham, Harris, MacArthur, & Schwartz, 1991)

The research on validated methods for assessing the writing process is limited. For interested readers, please see the section on portfolio assessment as one method for understanding the writing process of struggling writers.

ASSESSING WRITING SKILLS AT ALL LEVELS OF LANGUAGE

When assessing writing skills, teachers may be concerned with only one area of writing or several aspects of writing. In this section, we describe several methods and tools for assessing writing at different levels of language.

Text

Text level writing is the most challenging for any writer. At the paragraph level and beyond, writers must attend to multiple aspects of writing in order to develop a high quality composition. Teachers can assess students' paragraph and multi-paragraph writing using already constructed quality rubrics or by developing rubrics to assess specific areas of interest.

Quality

Overall assessments of writing quality are used for multiple assessment purposes. Typically, quality judgments use rubrics to assign a score. Holistic, analytic, or primary-trait rubrics are three commonly used types of rubrics.

Holistic. Holistic scoring reflects a rater's overall impression of the writing compared to other writing samples in the group (e.g., considering the aptness of vocabulary, grammar, sentence structure, organization, and content with no factor given undue weight). Holistic scoring can be norm referenced—a score that ranks students within a particular group—or criterion referenced—a score representing predetermined characteristics of writing quality.

Holistic scoring is the most economical and quickest method of scoring direct writing assignments. It is most useful to provide an overall ranking of students in a classroom or school, which is why it is often used in large-scale writing assessments. A school could also use holistic scoring for screening purposes. Students falling below a particular score could be targeted to gather more information about their writing abilities.

Unfortunately, holistic scoring does not provide instructional or intervention guidance for specific areas of concern because only one score is given to represent multiple aspects of writing. Students receiving the same score may have very different writing abilities and, therefore, would require different writing interventions. Holistic scoring also does not work well for progress monitoring purposes. The scales usually represent a significant amount of change from one score to the next, meaning that the scale would not be sensitive to small changes in writing ability.

NAEP provides an example of purpose-specific holistic writing rubrics. Separate rubrics can be found for writing to persuade, writing to explain, and writing to convey experiences, real or imagined. Each rubric includes specific criteria for scoring compositions on a scale of 1 (low) to 6 (high). Raters attend to text structure, idea development, sentence structure, and mechanics. More information can be found on the NAEP website: <http://nces.ed.gov/nationsreportcard/>.

Analytic. In contrast to holistic rubrics, analytic rubrics separately evaluate specific areas of writing to aid instructional decision making. Analytic scales are criterion referenced and assign separate scores in predetermined areas of good writing, such as organization, conventions, or word choice. Analytic scales take longer to score, but they may provide more information that can be helpful to guide instruction. The same rubric is useful across most writing genres.

Analytic rubrics are used frequently in schools and in large-scale writing assessments. One specific analytic rubric, *6 Trait® Writing Scoring Guide* (Northwest Regional Education Laboratory, 2004), has been used extensively in elementary and secondary schools as both an assessment and instructional method. Teachers are cautioned about the use of *6 Traits* for these purposes. First, there is little empirical evidence supporting the use of *6 Traits* as an instructional method. Second, some studies have found that the traits do not represent distinct writing skills (e.g., Gansle, Vanderheyden, Noell, Resetar, & Williams, 2006), indicating that the scores operate more as holistic measures. This happens when scores on one scale are related to scores on another scale. For example, using the *6 Traits* rubric, it would be difficult for a student to score high on Organization and low on Ideas, because good organization requires the development of ideas in writing.

When well-designed analytic rubrics are linked with the goals for instruction, they can provide information about the relative strengths and needs of individual students. The use of a few scales (e.g., Content/Organization and Language Use) rather than several allows for faster scoring and may provide more useful information. Teachers can decide which writing abilities to score based on the goals of instruction. As with holistic rubrics, most analytic rubrics are not sensitive enough to be used for progress monitoring purposes.

Primary Trait. Similar to analytic scales, primary-trait scales also are criterion referenced. They differ from analytic scales in that the scoring guide is developed based on the specific purpose of each writing assignment. Accordingly, each genre of writing or instructional goal would require a different primary trait rubric or scale. Primary trait rubrics can be developed to score any part of a written composition. As with analytic scoring, they are most useful when they are linked to instructional goals. They can be used to assess the primary goal of the writing assignment (i.e., coherence of an argument) or to reflect genre-specific requirements (i.e., plot or character development). Primary-trait scales can be useful in evaluating specific parts of a student's writing and may be sensitive to smaller changes in a student's writing that come from instruction.

As an example, researchers have developed primary trait rubrics based on Stein and Glenn's (1979) story grammar components (e.g., Troia, Lin, Monroe, & Cohen, 2009). These components address the setting (characters, locale, time) and plot elements (initiating event, internal response, goal, attempt, consequence) typically found in stories. Depending on a student's grade, a teacher could look for the presence or absence of these elements or rate each element on a scale of 0 (absent), 1 (present, but not well developed), or 2 (present and well developed).

Paragraph

There are few developed methods for rating and measuring a student's ability to construct paragraphs. Depending on a student's needs and goals, teachers may be concerned with paragraph conventions or paragraph development. Paragraph conventions include aspects such as indenting or skipping a line to indicate paragraph divisions and the proper use of margins, depending on the kind of text written. Paragraph development addresses the overall structure of the paragraph. Paragraph structures vary widely depending on the expected genre, but the inclusion of topic sentences, main ideas, and details is consistent across many paragraph structures. Other important aspects of structure include the placement of a topic sentence, the ordering of sentences, transitions between sentences, and the use of varied sentences lengths.

Paragraph checklists or rubrics can help both teachers and students focus on writing at the paragraph level. Students who are not proficient in paragraph writing will struggle with multiparagraph writing. This is because paragraphs are like mini-compositions: They require understanding of conventions, organization, sequencing, transitions, and variety.

Despite the lack of developed paragraph scoring methods, teachers can use their state's English Language Arts standards to help them identify grade-appropriate paragraph conventions and development. It is often necessary to review the previous grade's standards, as most states assume that subsequent grades will continue to develop these skills. From the standards, teachers can develop a checklist or rubric to score paragraphs. A checklist usually indicates the presence or absence of an element (e.g., yes/no, always/sometimes/never) and works well for younger students. A rubric evaluates the quality of the element (e.g., high/average/low). The scoring guide can be tailored for single paragraphs or text containing multiple paragraphs. The scoring guide can also be developed for different genres of writing or for a specific genre of writing, as needed.

For example, one state's standards indicates that by grade 3 students should be able to write a informational paragraph that is indented, has a variety of sentence structures, and contains a topic sentence and three supporting details. From these standards, a teacher could construct a checklist like the one shown in Figure 1.

Using this checklist, a teacher could easily assess whether a student understands indenting, uses different sentence structures, maintains a consistent topic, and includes a topic sentences and related details. Calculating the accuracy and number of different kinds of sentence structures may be useful to identify a student's sentence patterns. A student may use several correctly written simple sentences but include only one complex sentence or a number of complex sentences that are incorrectly written.

This type of scoring works best for purposes of instructional planning and modification and progress monitoring. From this evaluation, a teacher can identify the primary needs for instruction. A student who writes a paragraph with no topic sentence or focus requires different instruction than a student who includes a topic sentence but does not know how to support it with details. In addition, the scoring guide can be used for progress monitoring purposes. A teacher could provide instruction and periodically evaluate a student's paragraph writing during instruction to ensure that the student is making adequate progress toward the state standards or IEP goals. The scoring guide and the student's writing could also be included as part of a writing portfolio.

One limitation to developing a paragraph checklist or rubric is time. A teacher will need time to develop the scoring guide and to use it with each student's writing. Another

Element		
Is the paragraph indented?	<input type="checkbox"/> yes	<input type="checkbox"/> no
Is there a simple sentence?	<input type="checkbox"/> yes	<input type="checkbox"/> no
___ Number of simple sentences		
___ Number of correct simple sentences		
Is there a compound sentence?	<input type="checkbox"/> yes	<input type="checkbox"/> no
___ Number of compound sentences		
___ Number of correct compound sentences		
Is the paragraph all on one topic?	<input type="checkbox"/> yes	<input type="checkbox"/> no
Is there a topic sentence?	<input type="checkbox"/> yes	<input type="checkbox"/> no
1. Is there a detail?	<input type="checkbox"/> yes	<input type="checkbox"/> no
Does it relate to the topic sentence?	<input type="checkbox"/> yes	<input type="checkbox"/> no
2. Is there a detail?	<input type="checkbox"/> yes	<input type="checkbox"/> no
Does it relate to the topic sentence?	<input type="checkbox"/> yes	<input type="checkbox"/> no
3. Is there a detail?	<input type="checkbox"/> yes	<input type="checkbox"/> no
Does it relate to the topic sentence?	<input type="checkbox"/> yes	<input type="checkbox"/> no

FIGURE 1.
Paragraph Assessment Rubric for Grade 3 Students.

limitation is a lack of guidance from state standards on paragraph writing. Many writing standards are too vague to fully guide a teacher in instruction and assessment. As an example, the Common Core Standards for English Language Arts, national standards that have been adopted by a number of states at this time, provide very little information about paragraph writing for K–12 students.

Sentence

A student's sentence-level skills can be assessed either informally or formally. With informal methods, specific sentence-level criteria are identified and evaluated through observation and analysis of writing samples to learn more about the targeted areas (e.g., variety of sentence styles). Formal assessment most commonly involves writing CBM (W-CBM), which will be discussed later in this paper. W-CBM provides teachers with an indicator of a student's

general writing proficiency (Benson & Campbell, 2009) but also can be used to assess a student's sentence writing skills.

Informal Assessment

Unfortunately, the research offers scant guidance regarding effective ways to informally assess sentence-level knowledge and skills. Although informal procedures can provide teachers with useful information about a targeted area of interest (e.g., syntactical development), the lack of scientifically validated standardized procedures is a threat to the reliability and validity of the data (e.g., Benson & Campbell, 2009). However, the recommendations offered by Gregg and Mather (2002) and Miller (2009) provide practical suggestions that could be applied to a student's existing writing products (i.e., those produced in the context of daily classroom activities and assignments) or to

writing samples gathered through tasks designed specifically for sentence-level assessment. The data gained from this form of informal assessment would be most useful for the purposes of instructional planning and modification and for progress monitoring.

Miller (2009) recommended targeting three specific sentence-level variables; together they are termed writing fluency. They include the following:

1. Word fluency—count the total number of words a student wrote in a particular writing sample and divide by the number of sentences in the piece. Student goal: increase the number of words in each sentence for more sophisticated sentences.
2. Variety of sentence styles—count the number of sentence fragments, simple sentences, compound sentences, complex sentences, and compound-complex sentences. Student goal: reduce the use of sentence fragments and simple sentences while concomitantly increasing the use of compound, complex, and compound-complex sentences
3. Sentence complexity—count the number of declarative, imperative, interrogative, and exclamatory sentences. Student goal: “increase the student’s use of a healthy mix of different sentence types in each piece of writing” (p. 184).

Readers interested in learning more about using this form of informal sentence-level assessment alone or in combination with other informal assessment procedures are encouraged to see Miller’s descriptions and case study.

Gregg and Mather (2002) offered another option for informally assessing a student’s sentence-level knowledge and skills, with specific focus being given to (a) punctuation and capitalization and (b) syntax. Related to *punctuation and capitalization*, Gregg and Mather emphasized that awareness and application of these features is a developmental process that should be guided by analysis of errors, similar to that which is used with spelling. Four recommended questions guide this process and identify salient patterns that should be targeted through instruction (p. 22):

1. Does the student end sentences with correct punctuation?
2. Does the student use internal punctuation correctly?
3. Does the student use capital letters correctly?
4. Does the student indent paragraphs correctly?

Gregg and Mather (2002) recommended a similar error analysis process to assess a student’s knowledge and use of *syntactic structure*. A recommended first step is to compare a student’s oral and written syntax. This can be done

by examining whether he or she does the following when orally retelling text and when writing (pp. 21–22):

- Does the student use correct prefixes and word endings?
- Does the student maintain verb tense?
- Does the student use pronouns correctly?
- Does the student maintain subject–verb agreement?
- Does the student use complete sentences?
- Does the student use sentences with varied lengths?
- Does the student use sentences with varied structures?

Letter/Word

Teachers may be distracted by handwriting, spelling, or word selection difficulties when evaluating a student’s writing abilities. While each is important to overall writing development, it is equally important to consider sentence-, paragraph-, and text-level skills. Accordingly, it is helpful for teachers to proactively plan when and how to assess letter and word abilities, thereby ensuring an appropriate amount of attention to these skills within the broader writing continuum.

Spelling

A relatively large body of research and literature relates to assessment of spelling. Among the guiding principles that have emerged from this work are that (1) spelling is a complex, language-based skill that requires the integration and application of several linguistic knowledge sources, such as phonology, orthography, semantics, and morphology (Ehri, 2000; Templeton, Bear, & Madura, 2007); (2) all students’ spelling development follows a relatively predictable sequence (Gregg & Mather, 2002; Moats, 1995; Trieman, 1997); and (3) students’ correct and incorrect spellings can be used to diagnose their level of spelling development, tailor instruction, and monitor progress (Apel, Masterson, & Niessen, 2004; Bear, Ivernizzi, Templeton, & Johnston, 2008; Berninger & Amtmann, 2003). Within that context, we next highlight recommendations for collecting and analyzing data using word inventories and writing samples. Curriculum-based measurement in spelling (S-CBM), an excellent tool for screening and progress monitoring, is discussed later in the article.

Word inventories. A cornerstone for effective assessment of spelling is collecting an adequate sample of a student’s spellings, such that the full domain of spelling patterns within the English language is represented (Apel et al., 2004; Templeton et al., 2007). Selecting and administering a dictated word inventory that has been systematically created and validated based on developmental spelling theory is an excellent way to begin this process (Masterson & Apel,

2000). Word inventories are commonly used for the purposes of screening, instructional planning and modification, and progress monitoring. Two examples of comprehensive, research-based inventories are those created by Bear et al. (2008) and Masterson, Apel, and Wasowicz (2006).

Inventories based on the stage theory of spelling development. Bear et al. (2008) offered three word inventories that, together, are appropriate for students in kindergarten through 12th grade: the Primary Spelling Inventory (for grades K–3), the Elementary Spelling Inventory (for grades 1–6), and the Upper Level Spelling Inventory (for grades 5–12). These inventories are designed to assess student's knowledge of the spelling features that characterize a theoretical model of spelling development that includes five stages: Emergent, Letter Name, Within Word, Syllables & Affixes, and Derivational. Each inventory contains approximately 25 words, requires approximately 20–30 minutes to administer and can be given individually or to a group of students.

The basic steps a teacher would use to administer a Bear et al. (2008) inventory are (1) select an appropriate inventory based on the student's grade and spelling achievement level, (2) administer the inventory by pronouncing each word and (if necessary) using it in a sentence, and (3) score and analyze the results with a feature guide to determine what stage best describes a student's current spelling performance. This third step, which involves carefully examining a student's correct and incorrect spellings, is very important.

Scoring the inventories is more than marking words right or wrong. Instead, each word has a number of orthographic "features" that are scored separately. For example a student who spells *when* as *wen* knows the correct short vowel and consonant and get points for knowing those features even though the spelling is not correct. This analysis provides qualitative information regarding what students know about specific spelling features and what they are ready to study next. (Bear et al., p. 33)

Bear et al. (2008) provided several templates to help teachers organize the data from the spelling inventories and used it to effectively implement the word study curriculum, *Words Their Way*. Readers interested in learning more about the stage-based inventories developed by Bear et al. as well as the error analysis procedures that use the feature guides are encouraged to see Bear et al. (2008), Ivernizzi and Hayes (2004), and Templeton et al. (2007).

Inventories based on the repertoire theory of spelling development. Whereas the Bear et al. (2008) spelling inventories are based on a stage theory of spelling development, the Spelling Performance Evaluation for Language and Literacy (SPELL-2; Masterson et al., 2006) includes word inventories that are based on a repertoire theory of spelling. This theoretical framework suggests students use multiple

strategies and a variety of linguistic knowledge sources throughout the process of learning to spell; what varies is the degree to which each factor is used at various points in development (Apel & Masterson, 2001; Apel et al., 2004). SPELL-2 is designed for students in second grade or older and is administered individually via a multimedia CD-ROM computer program. Assessment begins with a student spelling 40 words included in the Selector Module. His or her performance on this initial corpus of words is used to determine which of four Main Test Modules is most appropriate. Depending on the Main Test Module administered, students spell 82–184 additional words that represent 29–65 orthographic patterns (Apel et al., 2004). Based on the student's performance, the SPELL system also selects and administers other testing modules to provide relevant supplemental data (e.g., phoneme segmentation, morphological awareness).

After all necessary tests are completed, SPELL-2 carefully analyzes how the student spelled each word and determines what spelling patterns he or she has already mastered and which ones are the source of frequent errors (Apel & Masterson, 2001; Kelman & Apel, 2004). This information is then used to generate a formal report that summarizes the student's strengths and needs, identifies appropriate learning objectives, and recommends specific word study lessons in the corresponding curriculum, *SPELL—Links to Reading and Writing*. Readers interested in learning more about the prescriptive assessment process used with SPELL-2 are encouraged to see the descriptions and case studies presented in Apel et al. (2004) and Masterson and Apel (2000, 2010) and to visit the website <http://www/learningbydesign.com>.

Other useful inventories. In addition to administering a comprehensive, theoretically-based word inventory (such as Bear et al., 2008 and Masterson et al., 2006), many educators find it beneficial to use supplemental, specialized inventories that target particular content areas or academic skills. One excellent resource related to writing is the Basic Spelling Vocabulary List (Graham, Harris, & Loynachan, 1993), which can be downloaded for free at <http://www.readingrockets.org/article/22366>. This list consists of 850 high-frequency words that, collectively, account for 80% of the words students use in their writing. The words also occur frequently in other literacy-related tasks, such as content-area reading. Therefore, assessing a student's ability to spell the words on the Basic Spelling Vocabulary List and then targeting instruction at those that are not yet known will be highly beneficial. As presented, the words on the list are organized into recommendations that could be taught in grades 1 through 5. Graham et al. noted, however, that the words could also be grouped in different ways, such as categorically (e.g., colors) or to illustrate specific spelling patterns, to best meet a student's needs.

Writing Samples

Although word inventories represent an effective and efficient strategy to collect a representative sample of spelling patterns, they have the inherent limitation of being decontextualized; that is, they do not represent an authentic composition task where students must attend to the multiple demands of the composing process (Apel et al., 2004; Apel & Masterson, 2001; Moats, 1995). Effective assessment of spelling, therefore, also includes gathering and analyzing writing samples. This assessment practice is particularly beneficial for instructional planning and modification.

Given that spelling accuracy can be influenced by a variety of task-related factors—such as the writing topic, genre, motivation to write accurately, available resources, and response mode—a robust and representative collection of writing products that show a student's spelling attempts should be assembled (Masterson & Apel, 2000; Moats, 1995). This can be done by including both existing writing samples (e.g., daily journal entries, a lab report written in science class, a poem written for the school newspaper) and those produced specifically for spelling assessment purposes (e.g., an opinion essay composed in response to a prompt, a written retelling of a story the examiner constructed specifically to include targeted error patterns) (Apel & Masterson, 2001).

Students' spelling samples gathered in the context of authentic writing should be systematically analyzed to identify error patterns that can inform instruction. This can be done by applying the same framework that was used to analyze a student's spellings of words from a theoretically based word inventory. For example, if a first-grade teacher administered the Bear et al. (2008) Primary Spelling Inventory to a student, the student's writing could be analyzed using the corresponding Qualitative Spelling Checklist (available on the CD-ROM packaged with Bear et al., 2008). This tool includes guiding questions that correspond to the five stages of spelling development. Using the Letter Name–Alphabetic stage as an illustration, the teacher would review the student's spelling samples in each writing product and indicate Yes, Often, or No in response to the following questions:

- Are beginning consonants included? (*b* for *bed*, *s* for *ship*)
- Is there a vowel in each word?
- Are some consonant digraphs and blends spelled correctly? (*ship*, *when*, *float*)
- Are there logical vowel substitutions with a letter name strategy? (FLOT for *float*, BAD for *bed*)
- Are short vowels spelled correctly? (*bed*, *ship*, *when*, *lump*)
- Is the *m* or *n* included in front of other consonants (*lump*, *stand*)

If a teacher administered the SPELL-2 assessment, he or she could use an error analysis procedure based on the repertoire theory of spelling development. Two excellent examples of how to collect writing samples and analyze spelling patterns using this approach can be found in the case studies described in Kelman and Apel (2004) and Apel and Masterson (2001).

Another recommended framework for coding and analyzing a student's spelling errors can be found in Berninger and Amtmann (2003). This model is supported by a substantial body of experimental research and can be applied to spelling samples gathered from authentic writing products or any other source of data (e.g., a student's spellings words from a dictated list given as part of a norm-referenced test). Berninger and Amtmann's scheme focuses on eight common spelling error patterns: phonological processing, phonological-orthographic processing, word-specific orthographic processing, morphological processing, spelling conventions, phonological/orthographic/morphological confusions, prealphabetic principle, and letter production errors (p. 357). For each, a description of the error pattern, diagnostic questions and examples, and intervention recommendations are offered.

A student's writing samples also provide an excellent, authentic resource for assessing his or her ability to recognize correct and incorrect spellings and to self-correct spelling errors (e.g., Kelman & Apel, 2004). To gather this kind of data, the teacher can ask the student to (1) review each of his or her writing products; (2) circle any misspelled words; and (3) offer an alternative, correct spelling for any spelling errors that are identified.

Vocabulary

As with some other areas of writing, research has validated few methods for assessing a student's written vocabulary. Teachers are encouraged to use their state standards to understand the kind of vocabulary students are expected to develop throughout K–12 schooling. Four general categories of vocabulary tend to be prevalent in standards. These include descriptive words, academic words, domain-specific words, and figures of speech.

While word choice in writing is an important aspect of overall writing quality, it is important to understand that the kind of words a writer uses may depend on the genre or type of writing. For example, domain-specific words are likely more prevalent in informational writing than story writing, while story writing may rely on descriptive, sensory words. Therefore, assessing a student's writing vocabulary requires understanding the kinds of words most important in each written genre. However, it does appear that using a variety of different words is important, regardless of the genre (Olinghouse & Leaird, 2009)

Once students are able to type their compositions on the computer, teachers can use an online program called VocabProfiler (Cobb, n.d.; available at <http://www.lexutor.ca/vp/eng/>) to gain an understanding of the kind of vocabulary in a student's writing. Using the student's writing sample, the program provides the percentages of words that are among the 2,000 most frequently used words in English (Brown Corpus; Francis & Kucera, 1982), on the Academic Word List (Coxhead, 2000), and off-list words (which tend to be domain-specific words). Most helpfully, the student's composition is color coded to show which words fall in each category. For example, off-list words are colored red, while academic words are colored yellow. This allows a teacher to quickly see whether students include targeted domain-specific or academic words in their writing.

Handwriting

As noted earlier, handwriting is a critical skill underlying a student's writing development and performance as well as success with other aspects of learning (Graham, Harris, & Fink, 2000). Effective assessment of handwriting is, therefore, tremendously important.

Research spanning several decades documents that the actual practices used in schools often fall short of this goal. If students' handwriting skills are assessed—and that is not always the case—the process is usually informal, imprecise, and based on subjective criteria (Graham, 1986a, 1986b; Graham et al., 2008; Hammerschmidt & Sudsawad, 2004). Therefore, we next provide an overview of how to make handwriting assessment more systematic and comprehensive (e.g., Graham, 1982, 1986a; Graham & Miller, 1980; Graham & Weintraub, 1996; Gregg & Mather, 2002; Tomchek & Schneck, 2006).

Evaluate handwriting legibility using multiple assessment tasks. Because handwriting legibility can be impacted by different task demands and conditions, it is recommended that teachers compile a robust collection of handwriting samples from each student. This often begins by having a student complete a series of developmentally appropriate copying and free-writing assessment tasks. For example, the student could be asked to use his or her best handwriting to:

- reproduce a close model (e.g., a sentence written at the top of his or her paper);
- reproduce a distant model (e.g., a sentence written on the board in the front of the classroom);
- write pre-determined, targeted text from memory (e.g., "Please print the letters of the alphabet and the numbers 0 through 9.") or as it is dictated (e.g., "Please print the letter a. Please print the letter b."); and

- spontaneously generate text in response to a prompt (e.g., My favorite thing to do after school is ___ because ___) or a picture (e.g., "Write at least four sentences that describe what you think is happening in this picture").

It is also beneficial to obtain samples of a student's usual, best, and fastest handwriting to identify whether an how different speed demands impact performance. This can be achieved through the use of a copying activity. First, the teacher gives the student the model text to be copied. The sentence, "The quick brown fox jumps over the lazy dog." is an excellent choice for this activity because it contains all the lower case letters in the alphabet yet is simple enough for most students to learn and spell. Next, the teacher asks the student to copy the sentence a certain number of times, "at the speed you usually use when you write." The designated number of repetitions should ensure the student has ample opportunity to practice the sentence and to write for at least 2 or 3 minutes. Then, the teacher asks the student to copy the sentence one time, "as carefully and neatly as you can." Finally, the teacher asks the student to copy the sentence "as fast and as many times as you can in three minutes." In between each of the three copying activities, the student should be praised for his or her effort and allowed time to rest, as needed.

Once multiple handwriting samples are obtained, it is recommended that teachers evaluate each sample for global-holistic legibility—that is, the overall readability of the text. Research has shown that scoring for legibility is more reliable when it is guided by a predetermined scoring framework, such as a handwriting evaluation scale. The Zaner-Bloser handwriting curriculum (2010) represents one example. Each grade-level Teachers' Guide includes a handwriting evaluation scale with five grade-appropriate models that are designated as Excellent, Good, Average, Fair, and Poor. In addition to evaluating a student's handwriting samples for global-holistic legibility, teachers can also evaluate legibility analytically. This is done by examining specific traits such as letter formation, uniformity and degree of slant, alignment, line quality, spacing between letters and words, letter size, beginning and ending strokes, and joining of letters.

Evaluate handwriting fluency. Handwriting fluency is the rate at which a student can produce letters or words. To assess fluency, teachers ask students to either copy or free write text for a predetermined period of time. For example, a first-grade student might be given 30 seconds to print as many of the letters of the lowercase alphabet as possible. Similarly, a fourth-grade student might be asked to copy a target sentence as many times as possible within 2 minutes.

A student's handwriting fluency score is typically calculated and reported as the number of letters written per minute or the number of words written per minute. For progress monitoring purposes, a student's fluency score can be graphed and analyzed over time (e.g., once a month). Normative data for handwriting fluency are also available and can be used to compare a student's rate with grade-level peers. For example, Graham, Berninger, Weintraub, and Schafer (1998) reported the average number of letters written per minute by students in first through ninth grade, respectively is 19 (1st), 34 (2nd), 47 (3rd), 63 (4th), 73 (5th), 85 (6th), 100 (7th), 115 (8th), 118 (9th). Teachers interested in using these normative data are encouraged to see the original Graham et al. article for additional information and statistics as well as data related to legibility

Evaluate the handwriting process. A student's handwriting process refers to the way in which he or she approaches and executes writing. Among the factors considered to be most important are posture, writing utensil grip, arm and hand location, and paper positioning. A simple way to gather information about a student's handwriting process is to carefully observe him or her during the series of handwriting assessment activities described above.

Gather additional information from other sources. To supplement the data gathered through the assessment activities described above, teachers are encouraged to seek out other sources of information related to a student's handwriting performance and processes. Examples include compiling and examining samples of written products a student produced under "typical" conditions (e.g., a story written at home, a worksheet completed during a science lab); interviewing a student's parents or former teachers to gain insight about previous handwriting instruction and other relevant skills (e.g., fine motor coordination); and consulting with specialists, such as occupational and physical therapists.

Another way to gather information about a student's handwriting is to select and administer a formal standardized handwriting evaluation tool. Examples that are commonly used and currently available include Children's Handwriting Evaluation Scale, Diagnosis and Remediation of Handwriting Problems, Evaluation Tool of Children's Handwriting, Minnesota Handwriting Test, and Test of Handwriting Skills. These instruments are particularly beneficial for diagnostic purposes, as they target multiple aspects of handwriting and allow a student's handwriting performance to be quantitatively compared to age and grade peers. For detailed discussions and reviews, readers are encouraged to see, for example, Feder and Majnemer (2003); Graham and Weintraub (1996); Rosenblum, Weiss, and Parush (2003); and Tomchek and Schneck (2006).

In summary, the handwriting assessment recommendations offered in this section are best suited for the purposes of instructional planning and modification and for progress monitoring. For example, the initial assessment of a student's handwriting should involve each of the strategies described above. This will offer a comprehensive picture of a student's strengths and needs and, in turn, will help identify appropriate instruction and interventions. Then, to monitor a student's progress, targeted assessment tasks can be administered in regular intervals. In other words, if a teacher discovers that a student has difficulty with handwriting fluency, the student's progress could be monitored by readministering the fluency assessment task (e.g., copying a target sentence for 2 minutes) once a week. These data could be graphed, analyzed, and used to inform subsequent instruction. Although there are limitations to the handwriting assessment plan we outline (e.g., the time required to gather data), it represents a significant improvement over the prevailing method of assessment—informal, imprecise observation (Graham, 1986a, 1986b).

Other Assessments of Writing: Portfolios, Norm-Referenced Tests, and CBM

Several other kinds of writing assessment are available that fall outside of the "levels of language" categorization. Portfolio and norm-referenced testing gather information about multiple areas of writing. W-CBM can be used as a general outcome measure for progress monitoring or to identify specific sentence writing and spelling difficulties. S-CBM, while specific to word spelling, is included in the CBM section to maintain coherence.

Portfolios

At the most basic level, a portfolio can be defined as a systematic and purposeful collection of student work that is intended to document effort, mastery, and/or growth in a particular subject area over a specified period of time (Walther-Thomas & Brownell, 2001; Weigle, 2002). Portfolio assessment is the practice of judging that work (Gearhart, 2009). The portfolios have a long history of use in creative fields, such as art and music because they allow individuals to show the breadth and depth of their accomplishments (Salvia & Ysseldyke, 2004). Today, portfolios are used in nearly all academic areas, with students in preschool through college, for multiple assessment purposes—including eligibility, instructional planning and modification, progress monitoring, and accountability (Gearhart, 2009; Salvia & Ysseldyke, 2004; Tierney et al., 1998; Weigle, 2002).

The use of portfolios with writing saw considerable expansion during the 1980s and 1990s as a result of multiple influences, such as the philosophical and practical implications of the whole language movement, a desire for

authentic assessment, and a greater understanding of the developmental process associated with learning to write. Educators saw portfolio assessment as a learner-centered approach that offered many advantages not associated with traditional assessment methods, such as standardized tests (see Gearhart, 2009, and Salvia & Ysseldyke, 2004, for a detailed discussion of the history and expanded use of portfolios). Examples of frequently-cited benefits associated with portfolio assessment include addressing and adapting to the individual needs of a wide range of learners; richly capturing each student's unique pattern of learning and formative progress; facilitating the practical integration of assessment and instruction; increasing collaboration among students, teachers, and family members; enhancing students' self-reflection and engagement in assessment; and promoting students' writing voices and identities (Boerum, 2000; Gearhart, 2009; Keiffer & Faust, 1996; Salvia & Ysseldyke, 2004; Tierney et al., 1998; Walther-Thomas & Brownell, 2001; Weigle, 2002).

Unfortunately, very little empirical evidence validates the benefits and uses of portfolio assessment (Gearhart, 2009; Salvia & Ysseldyke, 2004). However, based on the available quasi-experimental research and descriptive literature, portfolio assessment with writing is probably best conceptualized as an informal assessment method that—if used in combination with other forms of assessment (e.g., standardized tools and CBM)—is most appropriate for the purposes of eligibility, instructional decision making and modification, and progress monitoring (e.g., Gearhart, 2009; Salend, 1998; Tierney et al., 1998; Wesson & King, 1992).

Salend (1998) described a series of six guidelines to help educators effectively use portfolio assessment in their classrooms. Using this generic framework, we next offer practical recommendations for using portfolios with writing.

Guideline 1: Identify the goals of the writing portfolio. An absolutely essential first step with portfolio assessment in writing is to decide on the underlying purpose and goals for using portfolios (Salend, 1998; Salvia & Ysseldyke, 2004; Tierney et al., 1998). As Gearhart (2009) explained, "portfolios are flexible ways to support and assess students' writing progress, but their breadth and versatility can limit their usefulness if teachers and students use them for unclear or competing purposes" (p. 312). Therefore, before even considering issues such as the writing portfolio content or organization, a teacher must reflectively consider questions such as, "What is it I am hoping to gain from using writing portfolios?" and "How will portfolio assessment align with, complement, and supplement other writing assessment practices?"

The following two recommendations can serve to guide the process of determining clear, appropriate goals for writing portfolio assessment. First, the goals should align with

the general education writing curriculum and content standards as well as the writing-related needs of each individual student (Salend, 1998; Walther-Thomas & Brownell, 2001). For students with disabilities, the latter might include written language goals and objectives indicated on the student's IEP. Second, teachers should create opportunities for students to assume an active role in determining the goals of their personal writing portfolio (Boerum, 2002; Gearhart, 2009; Keifer & Faust, 1996; Tierney et al., 1998; Walther-Thomas & Brownell, 2001). Involving family members in the process of setting goals for the writing portfolio can also be extremely beneficial (e.g., Paratore, Hindin, Krol-Sinclair, & Duran, 1999; Salend, 1998).

Guideline 2: Determine the type of writing portfolio to be used. Once the goals and purposes for using a writing portfolio are clearly specified, the next step is to determine what type of portfolio is most appropriate (Gearhart, 2009; Salend, 1998). Examples of commonly used formats include the following:

- *Working writing portfolio*—contains a student's current compositions in various stages of progress and provides a place for the student to easily locate, review, and revise his or her writing.
- *Process writing portfolio*—contains writing samples and other artifacts to document how a student engages in the multiple phases of writing (e.g., brainstorming, outlining, drafting, revising, editing, publishing). Process portfolios can also target a student's use of writing strategies and self-regulation skills (such as those described in Harris, Graham, Mason, & Friedlander, 2008). This type of portfolio is an excellent method for learning more about a student's writing process.
- *Showcase writing portfolio*—contains a selection of a student's best work and is typically used to provide evidence of a student's overall writing ability, specific accomplishments, or both. For example, to demonstrate writing proficiency across genres, the student might select a final draft of a poem, a narrative story, a persuasive letter, and a report written for social studies.
- *Cumulative writing portfolio*—contains writing samples collected over an extended period of time (e.g., 4 years of high school) to document growth and accomplishments.

Guideline 3: Establish procedures for organizing the writing portfolio. Once an appropriate type of writing portfolio is selected, the next step is to decide how the contents will be organized and stored (Salend, 1998). Examples of possibilities include using accordion-style file folders, three-ring

binders, and boxes with dividers. In addition, several technology and multimedia options facilitate the organization and storage of writing portfolio materials (see, e.g., Stiggins, 2007). Spending time planning how writing portfolios will be organized—prior to their actual use—increases the likelihood that they will be successful vis-à-vis the established goals (Salvia & Ysseldyke, 2004).

Guideline 4: Choose a range of authentic products that relate to the objectives of the writing portfolio. The contents of a writing portfolio must be selected carefully. The following suggestions should help guide the selection process. First, it is essential that each item included in the writing portfolio addresses the clearly specified goals and purposes (Salend, 1998). Second, the writing samples included in the portfolio should be authentic, meaningful works. That is, they should be created for real people and real purposes, such as writing a letter to the editor of a newspaper to express views on a relevant community issue (Gearhart, 2009; Weigle, 2002). Moreover, it is recommended that a student's portfolio include a wide range of writing samples that represent multiple genres as well as varying contexts and levels of support. The advantage of this approach is that it offers a richer and deeper representation of the broader domain of writing. However, it has also been noted that increased diversity in the contents of a portfolio likely compromises a teacher's ability to make reliable judgments about a student's progress over time and to make comparisons between students. As Salvia & Yseeldyke (2004) explained:

It is very difficult for a teacher to judge student progress in writing from diverse products such as a poem, observations from a science walk, a letter, and a story; it is similarly difficult for teachers to compare the progress of two students when one student's portfolio contains persuasive prose and the other student's portfolio contains haiku. The products themselves are not comparable. Generally, the more comparable the products are, the less prone to error are the assessments. (p. 269)

Third, the portfolio should include an adequate number of writing samples to provide both teachers and students with adequate information for decision making. Although no number of samples to be included in a writing portfolio is universally accepted, it is better to err on the side of more rather than less (Salvia & Yseeldyke, 2004; Weigle, 2002). Fourth, unless the goal of the writing portfolio is purely to document summative achievement, it is recommended that the portfolio include writing samples produced at various points in time; this provides documentation of a student's developmental growth as a writer (Salvia & Yseeldyke, 2004; Walther-Thomas & Brownell, 2001; Weigle, 2002). Finally, students should assume an active role in selecting

and organizing the writing samples in their personal portfolio (Gearhart, 2009; Salend, 1998; Walther-Thomas & Brownell, 2001). Family members can also be involved with selecting portfolio contents or with creating items that relate to the use of writing within the family and community (see, e.g., Paratore et al., 1999).

Guideline 5: Record the significance of items included in students' writing portfolios. Each writing sample that is included in a student's portfolio should be accompanied by supplemental documentation (Salend, 1998). The purpose of this supplemental documentation is twofold: (1) to provide information about the context for the writing sample (e.g., date; description or purpose of the task; writing processes and resources used; nature and level of support provided by others such as teachers, peers, family members, etc.) and (2) to promote reflection regarding the significance of each piece vis-à-vis the original goals and purpose of the portfolio. Many students, and particularly those who struggle with writing, benefit significantly when teachers guide and support the process of self-reflection and evaluation that is involved in creating the supplemental documentation (Gearhart, 2009). For example, a student could be asked to respond in writing to a series of prompts that correspond with the necessary information related to both the context and the significance of the piece (e.g., What writing portfolio goal(s) is this piece related to? How does it show progress towards the goal(s)? What writing processes and strategies did you use while creating this piece and why? What did you learn from creating this piece that you will apply to your next composition?). Similar questions could also be used to structure a conversation between the teacher and a student (see Salend, 1998, for additional examples of prompts and summary statements).

In addition to the supplemental documentation that accompanies each writing sample, a student's writing portfolio can include other supplemental materials. For example, a student might create and include a narrative documentary of the myriad ways writing is used in his or her home and community. Similarly, the student could create and include an autobiography of him- or herself as a writer. Supplemental materials can also include feedback and reflections from teachers, peers, and family members (e.g., Paratore et al., 1999). Importantly, the supplemental materials in a portfolio should not be viewed as ends within themselves but rather as rich material that catalyzes collaborative discussions among the teacher, student, and family members and informs subsequent goals and plans (Salend, 1998; Tierney et al., 1998).

Guideline 6: Review and evaluate writing portfolios periodically. Periodically, a student's writing portfolio should be

reviewed and evaluated by the teacher as well as the student and family members (Salend, 1998; Salvia & Ysseldyke, 2004). This process is important because it promotes self-reflection among students and provides the teacher with information about the quality of a student's written work. Unfortunately, many of the same characteristics that educators describe as being the advantages of portfolio assessment (e.g., flexibility, diversity, collaborative decisions about contents) inherently compromise the reliability and validity of scoring portfolios (for a detailed critique, see, e.g., Salvia & Ysseldyke, 2004). To mediate these limitations, Gearhart (2009) suggested that evaluative criteria for writing portfolios should be (1) appropriate for the purpose of the assessment, (2) designed to measure students' progress toward important and developmentally appropriate learning goals, (3) meaningful to a range of stakeholders, (4) clear, and (5) fair (p. 206). Moreover, it is essential that the criteria are clearly specified and directly aligned with the purpose of the writing portfolio (Salvia & Ysseldyke, 2004; Weigle, 2002).

One common and practical way to organize the criteria for a writing portfolio is with a rubric (Gearhart, 2009; Salvia & Ysseldyke, 2004). Such rubrics should be simple and contain genre-specific criteria that provide students with guidance and feedback (for examples of rubrics, see Boerum, 2002; Salend, 1998; & Weigle, 2002). Many scholars and practitioners also emphasize the importance of having students assume a meaningful role in developing rubrics, because it increases their sense of ownership in the process and their understanding of how they will be evaluated (Gearhart, 2009; Tierney et al., 1998; Walther-Thomas & Brownell, 2001). Moreover, students should be actively involved in the evaluation process and provided with opportunities to share and discuss their work with others (e.g., peers, teachers, family members) (Paratore et al., 1999; Salend, 1998; Tierney et al., 1998; Walther-Thomas & Brownell, 2001).

Curriculum Based Measurement

CBM is a simple and efficient assessment procedure that involves administering multiple probes to quantify student progress towards long-range instructional goals (Deno, 2003; McMaster & Espin, 2007). CBM is particularly relevant and appropriate for use with students who experience learning difficulties, because it is focused on individual students, is highly sensitive to small increments of progress, and promotes ongoing evaluation and modification of instruction (Espin, Weissenburger, & Benson, 2004). CBM is considered a General Outcome Measure, meaning that it assesses a student's overall skills within the targeted academic domain rather than specific performance on a daily or weekly lesson or unit (Shinn & Shinn, 2002). As Shinn and

Shinn explained, CBM measures "are designed to serve as 'signs' of general achievement. They don't measure everything, but measure important things" (p. 9). A program of research spanning more than 30 years has demonstrated CBM's capacity to provide reliable and valid indicators of performance in multiple academic areas, including spelling (e.g., Deno et al., 1982; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Ritchey, 2008; Ritchey, Coker, & McCraw, 2010; Vaughn, Schumm, & Gordon, 1993).

CBM has been developed for both writing (W-CBM) and spelling (S-CBM). Both work well for screening and progress monitoring purposes. If a teacher is interested in assessing both spelling and overall writing proficiency, S-CBM can be used concurrently with W-CBM.

Writing CBM. A large body of research documents that W-CBM provides teachers with reliable and valid indicators of a student's overall writing proficiency (for detailed reviews of W-CBM research see, e.g., Benson & Campbell, 2009; Espin et al., 2004; Jewell & Malecki, 2005; McMaster & Campbell, 2008; McMaster & Espin, 2007; Powell-Smith & Shinn, 2004; Weissenburger & Espin, 2005). W-CBM is validated as being appropriate for students in first through 12th grade (Fuchs & Fuchs, n.d.; for an example of recent W-CBM research with kindergarteners, see Coker & Ritchey, 2010).

Selecting W-CBM measures. The first step in implementing W-CBM is to decide what type of writing measure will be used throughout the year (Fuchs & Fuchs, n.d.). Factors influencing this decision include the characteristics of the student, the goals for assessment (e.g., screening and/or progress monitoring; within and/or across grade levels), and contextual considerations (e.g., frequency of administration). Research with W-CBM has included and explored the use of varying types of writing measures with different populations of students, including narrative, expository, letter, picture-word, picture-theme, and photo prompts; word and sentence copying; and sentence writing (e.g., Coker & Ritchey, 2010; McMaster & Campbell, 2008; McMaster, Du, & Petursdottir, 2009; McMaster & Espin, 2007). Narrative prompts (also called story starters) that consist of a simple sentence that ends midsentence are used most commonly and have been validated with students across grade levels (Benson & Campbell, 2009; Espin, De La Paz, Scierka, & Roelofs, 2005; Espin et al., 2004; Fuchs & Fuchs, n.d.; Jewell & Malecki, 2003; McMaster & Campbell, 2008; Powell-Smith & Shinn, 2004).

Effective narrative prompts tap age-appropriate background knowledge of the students to whom they will be administered and are representative of a wide range of school experiences, such that they are relevant for students from diverse cultural, socioeconomic, geographic, and linguistic

backgrounds (Fuchs & Fuchs, n.d.). Examples of validated recommended narrative prompts include: "I found a note under my pillow that said ..." "I saw huge footprints in the sand and decided to follow them ..." (see Espin, Wallace, Lembke, Campbell, & Ticha, 2005 and Fuchs & Fuchs, n.d., for others). Two additional resources for prompts are www.progressmonitoring.org/probes/earlywriting.html and www.aimsweb.com/measures-2/written-expression-cbm.

W-CBM administration procedures. The general administration procedures for W-CBM consist of presenting students with the prompt, giving them 30 seconds to think about what they will write, and then having them write for a specified amount of time (Benson & Campbell, 2009; Fuchs & Fuchs, n.d.). Based on existing research, it is recommended that primary-aged students write for 5 minutes (e.g., McMaster et al., 2009), elementary-aged students write for 3 or 5 minutes (e.g., Gansle et al., 2006; Espin et al., 2004; McMaster & Campbell, 2008), and middle- and high-school aged students write for 7 minutes (e.g., Espin et al., 2008; Weissenburger & Espin, 2005). It is important that the same standardized administration procedures (including the amount of time given for writing) be used consistently throughout the school year. Research Institute on Progress Monitoring (2009a) and Powell-Smith and Shinn (2004) both provided examples of standardized administration procedures (i.e., directions and materials).

W-CBM scoring. Once a student completes a W-CBM writing measure, use standardized and validated procedures to score the text (Benson & Campbell, 2009; Fuchs & Fuchs, n.d.). There are four primary W-CBM scoring options:

1. Words Written (WW)—the total number of words written. When determining WW, all spelling, usage, capitalization, and punctuation errors are ignored.
2. Words Spelled Correctly (WSC)—the total number of English words spelled correctly. WSC does not typically require correct usage within the context of the sentence (e.g., *I went to there house for dinner.*), although context has been considered for some research (e.g., McMaster et al., 2009).
3. Correct Word Sequences (CWS)—the total number of adjacent, correctly spelled words that are syntactically and semantically correct within the context of the sample. When scoring for CWS, a vertical line is first placed where a sentence should end and then all incorrect words are underlined. Correct word sequences are marked by putting an upward-facing (blue) carat above the two words. Incorrect word sequences are marked by putting a downward-facing (red) carat below the two words. Carats are also placed at the beginning of each sentence to score for

correct or incorrect capitalization and at the end of each sentence to score for correct or incorrect punctuation. The final count of upward facing (blue) carats is CWS.

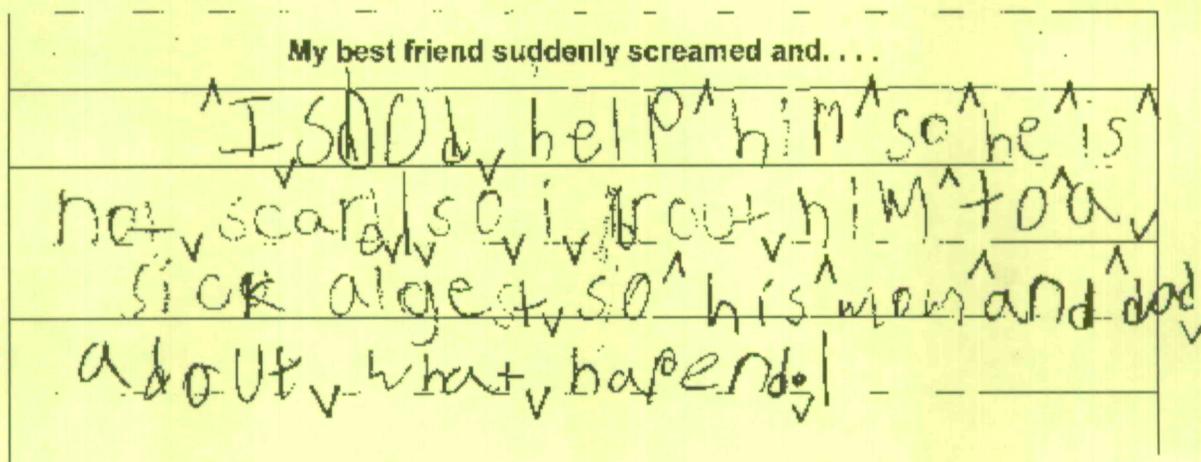
4. Correct Minus Incorrect Word Sequences (CIWS)—the number of incorrect word sequences (final count of downward facing [red] carats), subtracted from the number of CWS.

Figures 2 and 3 provide examples of the four scoring methods using writing samples produced by two students—one in fourth grade and the other in 11th grade. Readers interested in learning more about scoring W-CBM are encouraged to see Espin, Wallace, et al. (2005), Fuchs and Fuchs (n.d.), Powell-Smith and Shinn (2004), and Research Institute on Progress Monitoring (2009b); each resource provides a detailed description of the scoring methods, scoring rules, and multiple examples.

It is important that the same W-CBM scoring method or combination of methods be used consistently through the entire school year (Fuchs & Fuchs, n.d.). To date, research designed to explore the technical adequacy of W-CBM scoring methods suggests that, in general, WW and WSC can provide a basic measure of writing fluency for primary and elementary students, but that more complex scoring procedures, such as CWS and CIWS, produce stronger reliability and validity coefficients (for detailed discussions about the technical adequacy of W-CBM scoring methods see, e.g., Espin et al., 2008; Jewel & Malecki, 2005; McMaster & Campbell, 2008; McMaster et al., 2009; McMaster & Espin, 2007; Weissenburger & Espin, 2005). Fuchs and Fuchs (n.d.) recommended CWS be used for students in grades 1–4 and CIWS be used for students in grades 5–12.

Spelling CBM. Fuchs and Fuchs (n.d.) indicated that S-CBM is an appropriate and validated assessment strategy for students in first through sixth grade. Shinn and Shinn (2002) extended that range to also include students in seventh and eighth grade. S-CBM provides reliable and valid information about a student's overall spelling proficiency.

Creating and administering S-CBM measures. With S-CBM, students attempt to correctly spell a selected set of grade-level words that are dictated to them by the teacher, using a carefully timed, brisk pace (Fuchs & Fuchs, n.d.; Hosp, Hosp, & Howell, 2007; Shinn & Shinn, 2002). In contrast with the ubiquitous "Friday Spelling Tests" that include only words taught during a particular week, S-CBM probes consist of a random sample of all the words taught at a particular grade level over the course of an entire year. Consequently, with every S-CBM probe, students are assessed on a combination of words they have, and have not yet, encountered as part of instruction (Shinn & Shinn, 2002).



Note: Using the four primary W-CBM scoring options, the scores for this probe would be: Words Written (WW) = 24; Words Spelled Correctly (WSC) = 18; Correct Word Sequences (CWS) = 12; Correct Minus Incorrect Word Sequences (CIWS) = -2. The explanation for each Incorrect Word Sequence (IWS) is as follows: (1) incorrect spelling (*should*); (2) incorrect spelling (*scared*)—also missing end punctuation; (3) inappropriate way to begin a sentence (*so*)—also not capitalized; (4) not capitalized (*I*); (5) incorrect spelling (*brought*); (6) incorrect spelling (*psychologist*); (7) incorrect spelling (*about*)—also omitted text (e.g., *would know*); (8) incorrect spelling (*happened*).

FIGURE 2.
Writing Curriculum-Based Measurement (W-CBM) Probe Produced by a Fourth-Grade Student.

The first step in implementing S-CBM, therefore, is to acquire a set of equivalent but different word lists that will be given as probes over the course of the entire school year (Fuchs & Fuchs, n.d.; Hosp et al., 2007). Teachers can create S-CBM probes by randomly selecting words from a master list of graded words, such as those found in a published spelling curriculum or any other reputable source (e.g., the Dolch word list for the appropriate grade level). Teachers who are interested in creating S-CBM probes are strongly encouraged to see Fuchs and Fuchs (n.d.) and Hosp et al. (2007) for detailed information and guidelines related to developing appropriate and effective word lists. An alternate way to obtain S-CBM word lists is to purchase them (see, e.g., *Edformation's Standard Spelling Assessment Lists* as described in Shinn & Shinn, 2002).

S-CBM probes are administered using standardized procedures and directions (Fuchs & Fuchs, n.d.; Hosp et al., 2007; Shinn & Shinn, 2002). Administration can be done individually or with a group of students. Fuchs and Fuchs recommended that each S-CBM probe include 12 words and be administered in the following manner (for an example of slightly different administration procedures, see Shinn & Shinn, 2002). First, the teacher gives each student a piece of lined paper that is numbered from 1–12 and provides a

space for students to write each of the 12 words that they will attempt to spell correctly. Next, the teacher says the first spelling word (“Number 1—horse”), uses it in a sentence (“The horse is brown.”), and repeats the word (“Horse.”). After the word is presented, students are given 10 seconds to spell the word. After allowing for 10 seconds of response time—even if students are not finished spelling the word—the teacher presents the next spelling word in the same manner as the first (“Number 2—carefully. She carefully put the dish on the shelf. Carefully.”). The teacher repeats this process until all 12 words have been presented and students have been provided with 10 seconds of response time for each.

Scoring S-CBM. S-CBM is scored using a standardized and validated procedure that involves counting the total number of correct letter sequences (CLS) within each word (Fuchs & Fuchs, n.d.; Hosp et al., 2007; Shinn & Shinn, 2002). A CLS is a pair of letters that are placed together correctly. Each CLS is marked using an upward-facing carat. If the first letter of the word is correct, an upward-facing carat is placed at the beginning of the word. If the last letter of the word is correct, an upward-facing carat is placed at the end of the word. A student's CLS score is calculated by summing the carats. The maximum CLS score for any word is

One day my friend and I went to the zoo. . . .

^and^all^we^wanted^to^see^was^the^polar^bears.^We^got^to^
 the^entrance^of^the^zoo^purchased^our^tickets^and^went^
 quickly^inside.^We^looked^for^a^place^where^we^could^find^
 a^map^to^find^out^where^the^polar^bears^would^be.^We^
 found^one^and^quickly^went^in^the^direction^in^which^the^
 polar^bears^were.^We^ran^as^fast^as^we^could^to^get^
 there^first^so^we^could^spend^as^much^time^with^them^
 as^we^could.^~~We^went^to^the^zoo^and^~~Finally^we^got^to^the^
 and^the^polar^bears^were^playing^in^their^pool^with^
 a^big^red^ball.^They^looked^like^they^were^having^a^
 good^time.^After^we^spent^a^long^time^with^the^polar^
 bears^we^slowly^walked^away.^

Note: Using the four primary W-CBM scoring options, the scores for this probe would be: Words Written (WW) = 123; Words Spelled Correctly (WSC) = 122; Correct Word Sequences (CWS) = 126; Correct Minus Incorrect Word Sequences (CIWS) = 121. The explanation for each Incorrect Word Sequence (IWS) is as follows: (1) incorrect spelling (*entrance*); (2) missing end punctuation; (3) not capitalized (*They*).

FIGURE 3.

Writing Curriculum-Based Measurement (W-CBM) Probe Produced by an 11th-Grade Student.

always the total number of letters in the word plus 1. Here is an illustration of CLS scoring using the word "HORSE":

- Spelled correctly, there are 6 correct letter sequences:
^H^O^R^A^S^E^
- Spelled this way, there are 4 correct letter sequences:
^H^O^R^A^S
- Spelled this way, there is 1 correct letter sequence: ^H
R Z

Figure 4 provides an example of CLS scoring using a S-CBM probe produced by a fourth-grade student named Jon Mark.

With CLS, correct spelling of the complete word is not required to earn points; partial credit is given for correctly placed pairs of letters (Shinn & Shinn, 2002). Consequently, the developmental progression of spelling is captured and closer approximations of correct spellings are represented in a student's score. As Hosp et al. (2007) explained, this is

	Dictated Word	Possible CLS	Earned CLS
1	poplar	8	4
2	lovable	8	4
3	danger	7	7
4	scratch	8	6
5	tomorrow	9	4
6	wind	5	5
7	beautiful	10	3
8	different	10	6
9	front	6	6
10	pounce	7	4

Note: A correct letter sequence (CLS) is a pair of letters that are placed together correctly; each is marked with an upward-facing carat.

FIGURE 4.
Spelling Curriculum-Based Measurement (S-CBM) Probe Produced by a Fourth-Grade Student Named Jon Mark.

advantageous because it “gives the teacher, student, and parent a better idea of how close the student is to learning spelling patterns and getting the correct answer” (p. 77). Rescoring Jon Mark’s S-CBM probe (shown in Figure 4) dichotomously for words that are correct or incorrect provides an illustration. Jon Mark spelled 3 out of 10 words completely correct (i.e., *danger*, *wind*, and *front*), which results in a score of 30%. Alternately, he had 49 out of 78 letter sequences correct, which results in a score of 63%. The latter scoring method using CLS more accurately captures his spelling attempts for the seven words that he did not spell entirely correct.

Although CLS scoring requires a bit more time and effort than scoring a word simply as correct or incorrect, the benefits of using CLS are well documented: It is a better indicator of a student’s overall spelling proficiency, it is more sensitive to small changes in growth, and it is able to provide

diagnostic feedback (e.g., Fuchs et al., 1993; Ritchey et al., 2010; Shinn & Shinn, 2002). Readers interested in learning more about CLS scoring procedures and guidelines are encouraged to see Fuchs and Fuchs (n.d.), Hosp et al. (2007), and Shinn and Shinn (2002).

To complement the quantitative method of scoring S-CBM probes for CLS, many teachers find it beneficial to also analyze a student’s responses qualitatively to look for patterns of mastery and errors (Fuchs & Fuchs, n.d.). This can be accomplished by applying the same frameworks and procedures described previously (e.g., those developed by Apel & Masterson, 2001; Bear et al., 2008; Berninger & Amtmann, 2003) to examine a student’s S-CBM spelling samples. Another option is to use the qualitative error analysis framework found in Fuchs and Fuchs (n.d.). This model provides teachers with guiding questions related to five salient areas (p. 39):

1. Simple consonants and vowels (e.g., Does the student consistently identify consonants at the beginning and end of words?)
2. Pairs of consonants (e.g., Does the student recognize and write both the sounded consonants in words like “fast” and “hold”?)
3. Silent letters (e.g., Does the student write silent letters correctly, in such words as “know” and “climb”?)
4. Long vowels (e.g., Does the student correctly write long vowel sounds in a variety of ways such as words with a final e [such as “late”]?)
5. Multisyllabic words, including words with suffixes (e.g., Does the student include a vowel in each syllable of multisyllabic words?)

Given the limited number of words on each S-CBM probe, it is often beneficial to detect and analyze patterns using the data from multiple probes. Teachers find qualitative analysis particularly beneficial for identifying a student’s specific strengths and weaknesses and using those patterns to plan and modify instruction.

Charting and graphing W-CBM and S-CBM. The final step for using W-CBM and S-CBM is to graph a student’s score. Fuchs and Fuchs (n.d.) emphasized that graphing is a vital aspect of CBM because “the graphs give teachers a straightforward way of reviewing a student’s progress, monitoring the appropriateness of the student’s goals, judging the adequacy of the student’s progress, and comparing and contrasting successful and unsuccessful instructional aspects of the student’s program” (p. 22).

For both writing and spelling, it is recommended that three baseline probes be administered and progress be monitored weekly or monthly (Benson & Campbell, 2009; Hosp et al., 2007). Four excellent resources to learn more about

graphing scores, setting ambitious goals, and using decision rules to revise instruction and goals include Benson and Campbell (2009), Fuchs and Fuchs (n.d.), Hosp et al. (2007), and www.iris.peabody.vanderbilt.edu/resources.html.

Norm-Referenced Tests

Norm-referenced tests allow educators to compare a student's level of writing proficiency to a normative sample (Berninger & Amtmann, 2003; Masterson & Apel, 2000). Consequently, this type of assessment is appropriate to help answer questions such as "Is the student's present level of writing achievement within the expected range for his or her current age and/or grade-level?" "How does the student's writing proficiency compare with his or her achievement in other academic areas (e.g., reading, math)?" and "How do the student's various writing skills compare with each other (e.g., spelling, grammar)?" Within that context, norm-referenced tests have been an integral—and often mandated—part of the special education eligibility and placement decision-making process (Salvia & Ysseldyke, 2004).

In this section, we first offer a brief description of three norm-referenced tests that educators frequently use to assess a student's writing skills: The *Woodcock-Johnson Tests of Achievement—Third Edition* (WJ-III; Woodcock, McGrew, & Mather, 2001), the *Wechsler Individual Achievement Test—Third Edition* (WIAT-III; The Psychological Corporation, 2009), and the *Test of Written Language—Fourth Edition* (TOWL-4; Hammill & Larsen, 2009) (for detailed reviews, see, e.g., Salvia & Ysseldyke, 2004). We then highlight three important points educators should consider when selecting and interpreting the results from norm-referenced tests.

Three commonly used norm-referenced tests. The WJ-III is a comprehensive, norm-referenced test that is designed to measure a student's levels of achievement in multiple academic domains, such as reading, math, and written language. The WJ-III includes two parallel forms (A and B), each of which is divided into two batteries—Standard and Extended. The updated normative sample for the WJ-III included individuals between the ages of 1 and 80+, which makes the test appropriate for all ages. As shown in Table 1, the WJ-III includes six subtests that target written language: Spelling, Writing Fluency, Writing Samples, Editing, Sounds of Spelling, and Punctuation and Capitalization. A student's scores from these subtests can be combined into three written language cluster scores: Broad Written Language (includes Spelling, Writing Fluency, and Writing Samples subtests), Basic Writing Skills (includes Spelling and Editing subtests), and Written Expression (includes Writing Fluency and Writing Samples subtests). The WJ-III also includes a Handwriting Legibility Scale that can be used to analyze a student's handwriting samples produced in

conjunction with the Writing Samples subtest or any other written product.

Like the WJ-III, the WIAT-III is a comprehensive norm-referenced test that can be used to measure achievement in multiple academic domains. It does not, however, have multiple batteries or include parallel forms. The WIAT-III is appropriate for individuals between the ages of 4 and 50 and includes four subtests related to written language: Spelling, Alphabet Writing Fluency, Sentence Composition, and Essay Composition (see Table 1). Three of these subtests are administered only to students in certain grade levels: Alphabet Writing Fluency (for grades PK–3), Sentence Composition (for grades 1–12+), and Essay Composition (for grades 3–12+). A student's scores from the writing-related subtests on the WIAT-III can be combined into one composite score for 'Written Language.'

The TOWL-4 is a norm-referenced test that was designed specifically for the purpose of assessing multiple aspects of written language. It includes two parallel forms (A and B) and is appropriate for students who are 9.0 to 17.11 years old. A companion test to the TOWL-4, the *Test of Early Written Language—Second Edition* (Hresko, Herron, & Peak, 1996), is designed to be used with younger students. The TOWL-4 includes seven subtests: Vocabulary, Spelling, Punctuation, Logical Sentences, Sentence Combining, Contextual Conventions, and Story Composition (see Table 1). Scores from these subtests can be combined into three composite scores: Contrived Writing (includes Vocabulary, Spelling, Punctuation, Logical Sentences, and Sentence Combining), Spontaneous Writing (includes Contextual Conventions and Story Composition), and Overall Writing (includes all 7 subtests).

Important points for consideration. Many educators perceive norm-referenced tests to be advantageous, because the relatively short subtests and standardized procedures help simplify the otherwise complex and challenging process of assessing a student's writing skills (Penner-Williams, Smith, & Gartin, 2009). Another frequently cited benefit of norm-referenced tests is their technical adequacy. Although norm-referenced tests such as the WJ-III, WIAT-III, and TOWL-4 are used quite often to assess writing, research and our experience suggests that educators may not have a robust understanding of what information is—and is not—provided from the results. Therefore, we next offer a brief summary of three important considerations related to selecting and interpreting norm-referenced tests.

Are all levels of language and relevant skills assessed? First, it is important for educators to consider whether a norm-referenced test has adequate breadth and depth to provide comprehensive information about a student's writing skills. One question that can help guide this inquiry is: Does

TABLE 1
Summary of Writing-Related Subtests on Three Commonly-Used Norm-Referenced Tests

Test Name	Subtest Descriptions
Wechsler Individual Achievement Test—3rd Edition (WIAT-III)	<p>Spelling: The student attempts to correctly write/spell dictated letters, letter blends, and words. The student's responses are scored for accuracy.</p> <p>Alphabet Writing Fluency (administered to students in grades PK–3): The student writes down as many letters of the alphabet as he or she can within a 30 second time limit. The student's responses are scored for number of correctly written letters.</p> <p>Sentence Composition (administered to students in grades 1–12+): The student (1) combines 2 or more separate sentences and writes one sentence that means the same thing (<i>Sentence Combining</i>) and (2) writes sentences that use a target word in an appropriate context (<i>Sentence Building</i>). The student's sentences are scored for syntax, semantics, and (for Sentence Combining) joining the original text.</p> <p>* + Essay Composition (administered to students in grades 3–12+): The student is given 10 minutes to write an essay about his or her favorite game. The student's essay is scored for (1) total number of words written, (2) theme development and text organization (i.e., quality of the introduction, conclusion, paragraphs, transitions, persuasive reasons, and elaborations), and (3) grammar and mechanics using the CIWS scoring method (a description of CIWS can be found in the section devoted to W-CBM).</p>
Woodcock-Johnson Tests of Achievement—3rd Edition (WJ-III)	<p>Spelling: The student attempts to reproduce prewriting marks (e.g., lines, scribbles) and to correctly write/spell dictated letters and words. The student's responses are scored for accuracy.</p> <p>Writing Fluency: The student writes as many sentences that include three target words and relate to a stimulus picture stimulus as he or she can within a 7-minute time limit. The student's sentences are scored for appropriate syntactic use of the targeted words.</p> <p>Writing Samples: The student writes sentences in response to pictures, topic prompts, and incomplete paragraphs. Some require the use of specialized vocabulary. The student's sentences are scored for quality of expression.</p> <p>Editing: The student verbally identifies and attempts to correct errors in capitalization, punctuation, and grammar that are featured in written sentences and paragraphs. The student's responses are scored for accuracy.</p> <p>Sounds of Spelling: The student listens to audio-taped nonsense sounds and words and writes how the word would likely be spelled if it were a real English word. The student's spellings are scored for accuracy.</p> <p>Punctuation and Capitalization: The student is asked to print specified upper and lower case letters, make designated punctuation marks, and apply punctuation and capitalization rules to sentences and phrases. The responses are scored for accuracy.</p>

(continued)

TABLE 1. (continued)

Test Name	Subtest Descriptions
Test of Written Language—4th Edition (TOWL-4)	<p>Vocabulary: The student writes a sentence that appropriately incorporates a targeted word. The sentences are scored for correct word usage (i.e., parts of speech and meaning).</p> <p>Spelling: The student writes sentences from dictation. The sentences are scored for spelling accuracy.</p> <p>Punctuation: The student writes sentences from dictation (same as those used for Spelling). The sentences are scored for proper use of punctuation and capitalization.</p> <p>Logical Sentences: The student revises illogical sentences so that they make sense. The sentences are scored for meaning and grammatical correctness.</p> <p>Sentence Combining: The student combines 2 or more separate sentences into one sentence that means the same thing and is grammatically correct. The sentences are scored for meaning, cohesion and fluency, and grammatical correctness.</p> <p>* + Contextual Conventions: The student writes a story in response to a stimulus picture. He or she has 5 minutes to plan the composition and 15 minutes to write. The student's story is scored for proper use of orthographic and grammatical conventions (e.g., spelling, punctuation, sentence construction, noun-verb agreement).</p> <p>* + Story Composition: The story written for the Contextual Conventions subtest is rescored for holistic quality (e.g., plot, prose, character development, interest to readers, vocabulary).</p>

Note. * subtest uses a spontaneous writing format; + subtest targets text-level of language.

the test assess all three levels of language? As emphasized previously, each level of language plays an important role in a student's overall writing proficiency; thus, they should all be targeted for assessment.

Most norm-referenced tests include items that are related to the first two levels of language. As shown in Table 1, this is the case for the WJ-III, WIAT-III, and TOWL-4. Unfortunately, there is less consistency with regard to the third level of language; some norm-referenced tests assess text-level skills, whereas others do not. For example, the WIAT-III Essay Composition subtest provides educators with information about a student's text-level skills, as does the TOWL-4 Story Composition subtest. The WJ-III, in contrast, does not include a text-level measure; students are never required to put sentences together to form paragraphs or to write multi-paragraph connected text.

In addition to being cognizant of what levels of language are assessed by a norm-referenced test, it is important for educators to consider whether all the skills that comprise

each level are targeted. Here again, educators will find inconsistencies; some norm-referenced tests provide an assessment of all the relevant skills within a particular level of language, but others do not. For example, handwriting is an important skill related to the first level of language. However, neither the WIAT-III nor the TOWL-4 assesses a student's handwriting legibility. The WJ-III provides a Handwriting Legibility scale in the Examiner's Manual, but it is not formatted as a distinct subtest, it is not included in any of the written language composite scores, and it does not yield derived scores (e.g., standard score, percentile, age-/grade-equivalency). Vocabulary is another critical skill. However, only the TOWL-4 provides an educator with specific information regarding a student's vocabulary skills. In both the WIAT-III and the WJ-III, vocabulary is conflated with other writing skills (e.g., fluency, syntax, overall quality).

Are both contrived and spontaneous writing formats used? A second important point for consideration relates to

whether a norm-referenced test utilizes both contrived and spontaneous writing formats to gather information about a student's writing skills. Contrived formats are used to evaluate discrete components of written language (e.g., spelling, punctuation, grammar) in isolation. A student might, for example, read a printed sentence and insert the correct punctuation marks. Alternately, the student might spell a series of dictated words. Contrived formats are beneficial for gathering information about a student's strengths and weaknesses as they pertain to specific components of writing. They do not, however, measure a student's ability to apply the targeted skill or skills within the context of authentic writing tasks. That information can only be obtained through the use of spontaneous writing formats. Spontaneous writing tasks require students to integrate and apply the multiple components of writing to efficiently and effectively produce an actual written product. For example, a student might write a persuasive essay in response to a prompt. The composition would then be evaluated for overall quality and, in some cases, also for specific written language components (e.g., use of vocabulary, transitions, spelling).

Given the different, yet complimentary, purposes of contrived and spontaneous writing formats, comprehensive assessment of writing necessitates the use of both methods (Hammill & Larsen, 2009). Some, but not all, norm-referenced tests meet this criterion. As shown in Table 1, the WIAT-III and TOWL-4 gather data using both contrived and spontaneous writing formats. In contrast, all six subtests included in the WJ-III are contrived; students are never required to produce an authentic writing sample wherein they integrate and apply specific writing skills.

Is the full domain sampled? A third consideration involves examining whether a norm-referenced test adequately samples the domain for each written language component. In other words, does it provide enough text questions to comprehensively understand a student's strengths and weaknesses vis-à-vis particular writing skills? Because norm-referenced subtests include a limited number of items, the answer is invariably no. One illustration can be seen by contrasting the number of sample spellings a student produces when assessed using a comprehensive word inventory (such as SPELL-2, which was described previously) and a Spelling subtest from a norm-referenced instrument. With SPELL-2, a student is asked to spell between 122 and 224 words (depending on which Main Test Module is administered); this large corpus of words ensures the full range of English orthographic patterns is represented. In contrast, the TOWL-4 assessment requires a student to spell at most 26 words and in many cases far fewer because of ceiling guidelines. To their credit, the authors of the TOWL-4 explicitly highlight the fact that because norm-referenced tests restrict the domain of what is assessed, the results are impotent for

the purpose of instructional planning and modification (Hammill & Larson, 2009):

Although the TOWL-4 results may contribute to the selection of long-term educational goals, they should not be used as the basis for planning day-to-day instructional programs for individual students. For example, an examiner might suggest that a student who did poorly on the Spelling subtest needed instruction in spelling (a long-term goal). To write an actual instructional plan, however, would require the examiner know which words and spelling rules the student needs to learn. The TOWL-4 subtest deals with relatively few words, too few to warrant using the subtest items for instructional planning. Interpreting the TOWL-4 results could contribute to a comprehensive evaluation of a student's problem in writing, along with other sources such as clinical teaching, criterion-referenced testing, and informal assessment procedures. (p. 54)

Another example of restricted domains can be seen by examining the spontaneous writing tasks included in the WIAT-III and the TOWL-4. With each test, students are asked to write in only one genre; persuasive with the WIAT-III and narrative with the TOWL-4. A student's authentic writing skills in other genres are not assessed. This is important to consider because research suggests that a student's writing skills are not necessarily consistent across genres (Olinghouse & Santangelo, in preparation; Schoonen, 2005).

In summary, norm-referenced tests, such as the WJ-III, WIAT-III, and TOWL-4, can help educators understand how a particular student's writing skills compare with those of the norm group. Therefore, as Masterson and Apel (2000) explained, "clinicians or teachers who must have a standard score for justification of placement in special education services may benefit from administration and scoring of [norm-referenced] tests or subtests" (p. 51). However, for other assessment purposes, such as progress monitoring and instructional planning and modification, other forms of gathering information are more appropriate and beneficial.

As a final note, we would like to make readers aware of one norm-referenced test, the *Process Assessment of the Learner- Second Edition for Reading and Writing* (PAL-II RW; Berninger, 2007) that differs from those we highlighted. Although not yet widely used, the PAL-II RW is comprehensive and specifically designed to be used as part of a three-tier model of intervention and assessment with students in kindergarten through 6th grade (Peterson, Martinez, & Turner, 2010). The PAL-II measures handwriting, spelling, narrative composition fluency, and expository note taking and report writing. Readers interested in learning more about a norm-referenced instrument that can be used for the purposes of screening, eligibility, and instructional/intervention design are encouraged to consider the PAL-II.

CONCLUSION

A good assessment process requires a great amount of thought and planning. So why assess? First and foremost, teachers who assess thoughtfully and frequently are better at setting instructional goals, which results in higher student achievement. Assessments also are necessary to identify students who need more support, including special education services. Finally, assessment can have a direct influence on students. Communication of assessment results helps students understand and set short-term learning goals; clarify the types of skills, processes, or knowledge to be learned; and gain feedback about their progress.

Teachers can use a variety of writing assessment methods and tools in the classroom. First, it is important to determine the purpose of assessment, as this will drive the rest of the assessment process. Teachers can assess one or more levels of language depending on their assessment purpose, grade or ability of students, or instructional focus. Generally, multiple writing samples across a number of genres provide more accurate information than using only one writing sample.

To get started, teachers are encouraged to think about their year-long writing goals for each student or group of students. For teachers who are less familiar with writing assessment, it is best to start with a small, manageable assessment plan for instructional planning and modification purposes. As a teacher becomes more familiar and confident in assessing writing, he or she can then add more components to the assessment plan, such as screening or progress monitoring tools.

REFERENCES

- Apel, K., & Masterson, J. J. (2001). Theory-guided spelling assessment and intervention: A case study. *Language, Speech, and Hearing Services in Schools, 32*, 182–195.
- Apel, K., Masterson, J. J., & Niessen, N. L. (2004). Spelling assessment frameworks. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 644–660). New York: Guilford Press.
- Bangert-Drowns, R. L., Hurley, M., & Wilkinson, M. M., (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*, 29–58.
- Bear, D. R., Ivernizzi, M., Templeton, S., & Johnston, F. (2008). *Words their way. Word study for phonics, vocabulary, and spelling instruction* (4th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- Benson, B. J., & Campbell, H. M. (2009). Assessment of student writing with curriculum-based measurement. In G. Troia (Ed.), *Instruction and assessment for struggling writers* (pp. 337–357). New York: Guilford Press.
- Berninger, V. (2007). *Process Assessment of the Learner, 2nd Edition. Diagnostic for Reading and Writing (PAL-II RW)*. San Antonio, TX: The Psychological Corporation.
- Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 345–363). New York: Guilford Press.
- Berninger, V. W., Mizokawa, D. T., & Bragg, R. (1991). Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychology, 29*, 57–79.
- Boerum, L. J. (2000). Developing portfolios with learning disabled students. *Reading & Writing Quarterly, 16*, 211–238.
- Chase, C. (1986). Essay test scoring: Interaction of relevant variables. *Journal of Educational Measurement, 23*, 33–41.
- Cobb, T. (n.d.). *Web Vocabprofile* [an adaptation of Heatley & Nation's (1994) *Range*]. Retrieved from <http://www.lexutor.ca/vp/>
- Coker, D. L., & Ritchey, K. D. (2010). Curriculum-based measurement of writing in kindergarten and first grade: An investigation of production and qualitative scores. *Exceptional Children, 76*, 175–193.
- Connelly, V., Gee, D., & Walsh, E. (2007). A comparison of keyboarded and handwritten compositions and the relationship with transcription speed. *British Journal of Educational Psychology, 74*, 551–564.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213–238.
- De La Paz, S., Swanson, H. L., & Graham, S. (1998). The contribution of executive control to the revising by students with writing and learning difficulties. *Journal of Educational Psychology, 90*, 448–460.
- Deno, S. L. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention, 28*, 3–12.
- Deno, S. L., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children, 48*, 368–371.
- 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* (Vol. IRLD-RR-87). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.
- Dockrell, J. E., Lindsay, G., Connelly, V., & Mackie, C. (2007). Constraints in the production of written text in children with specific language impairments. *Exceptional Children, 73*(2), 147–164.
- Ehri, L. L. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*, 19–36.
- Englert, S., Raphael, T., Fear, K., & Anderson, L. (1988). Students' metacognitive knowledge about how to write informational texts. *Learning Disability Quarterly, 11*, 18–46.
- Espin, C. A., De La Paz, S., Scierka, B. J., & Roelofs, L. (2005). The relationship between curriculum-based measures in written expression and quality and completeness of expository writing for middle school students. *The Journal of Special Education, 38*, 208–217.
- Espin, C. A., Weissenburger, J. W., & Benson, B. J. (2004). Assessing the writing performance of students in special education. *Exceptionality, 12*, 55–66.
- Espin, C., Wallace, R., Lembke, E., Campbell, H., & Ticha, R. (2005, April). *Curriculum-based measurement (CBM): Reading and writing measures in secondary education*. Presentation made at the Council for Exceptional Children's Convention and Expo. Retrieved from <http://progressmonitoring.org/RIPMProducts2.html#scoring>
- Espin, C., Wallace, T., Campbell, H., Lembke, E. S., Long, J. D., & Ticha, R. (2008). Curriculum-based measurement in writing:

- Predicting the success of high-school students on state standards test. *Exceptional Children*, 74, 174–193.
- Feder, K. P., & Majnemer, A. (2003). Children's handwriting evaluation tools and their psychometric properties. *Physical and Occupational Therapy in Pediatrics*, 23, 65–84.
- Francis, W.N. and Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin.
- Fuchs, L. S., & Fuchs, D. (n.d.). *Using CBM for progress monitoring in written expression and spelling*. Retrieved from National Center on Student Progress Monitoring website: http://www.studentprogress.org/summer_institute/2007/Written/Writing_Manual_2007.pdf
- 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, 27–48.
- Gansle, K. A., VanDerHeyden, A. M., Noell, G. H., Resetar, J. L., & Williams, K. L. (2006). The technical adequacy of curriculum-based and rating-based measures of written expression for elementary school students. *School Psychology Review*, 35, 435–450.
- Gearhart, M. (2009). Classroom portfolio assessment for writing. In G. Troia (Ed.), *Instruction and assessment for struggling writers* (pp. 311–336). New York: Guilford Press.
- Graham, S. (1982). Measurement of handwriting skills: A critical review. *Diagnostique*, 8, 32–42.
- Graham, S. (1986a). A review of handwriting scales and factors that contribute to variability in handwriting scores. *Journal of School Psychology*, 24, 63–71.
- Graham, S. (1986b). The reliability, validity, and utility of three handwriting measurement procedures. *Journal of Educational Research*, 79, 373–380.
- Graham, S. (1990). The role of production factors in learning disabled students' compositions. *Journal of Educational Psychology*, 82, 781–791.
- Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187–207). New York: Guilford.
- Graham, S., Berninger, V., Weintraub, N., & Schafer, W. (1998). Development of handwriting speed and legibility in grades 1–9. *Journal of Educational Research*, 92, 42–52.
- Graham, S., & Harris, K. R. (1996). Self-regulation and strategy instruction for students who find writing and learning challenging. In M. Levy and S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 347–360). Mahwah, NJ: Erlbaum.
- Graham, S., & Harris, K. R. (1997). It can be taught, but it does not develop naturally: Myths and realities in writing instruction. *School Psychology Review*, 26, 414–424.
- Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist*, 35, 3–12.
- Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology*, 92, 620–633.
- Graham, S., Harris, K. R., & Loynachan, C. (1993). The basic spelling vocabulary list. *Journal of Educational Research*, 86, 363–368.
- Graham, S., Harris, K. R., MacArthur, C. A., & Schwartz, S. (1991). Writing and writing instruction for students with learning disabilities: Review of a research program. *Learning Disability Quarterly*, 14, 89–114.
- Graham, S., Harris, K. R., Mason, L., Fink-Chorzempa, B., Moran, S., & Saddler, B. (2008). How do primary grade teachers teach handwriting? A national survey. *Reading and Writing*, 21, 49–69.
- Graham, S., & Miller, L. (1980). Handwriting research and practice: A unified approach. *Focus on Exceptional Children*, 13(2), 1–14.
- Graham, S., & Perin, D. (2007). What we know, what we need to know: Teaching adolescents to write. *Scientific Studies of Reading*, 11, 313–335.
- Graham, S. & Weintraub, N. (1996). Review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review*, 8, 7–87.
- Gregg, N., & Mather, N. (2002). School is fun at recess: Informal analysis of written language for students with learning disabilities. *Journal of Learning Disabilities*, 35, 7–22.
- Hammerschmidt, S. L., & Sudsawad, P. (2004). Teachers' survey on problems with handwriting: Referral, evaluation, and outcomes. *American Journal of Occupational Therapy*, 58, 185–191.
- Hammill, D. D., & Larsen, S. C. (2009). *Test of Written Language-4*. Austin, TX: Pro-Ed.
- Harris, K. R., & Graham, S. (1999). Problematic intervention research: Illustrations from the evolution of self-regulated strategy development. *Learning Disability Quarterly*, 22, 251–262.
- Harris, K. R., Graham, S., Mason, L. H., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore: Brookes.
- Hayes, J. R., & Berninger, V. W. (2010). Relationships between idea generation and transcription: How the act of writing shapes what children write. In C. Bazerman, R. Krut, K. Lunsford, S. Mcleod, S. Null, P. Rogers, et al. (Eds.), *Traditions of Writing Research* (pp. 166–180). New York: Routledge.
- Hosp, M. K., Hosp, J. L., & Howell, K. W., (2007). *The ABCs of CBM: A practical guide to curriculum-based measurement*. New York: Guilford Press.
- Houck, C. K., & Billingsley, B. S. (1989). Written expression of students with and without disabilities: Differences across the grades. *Journal of Learning Disabilities*, 22(9), 561–568.
- Hresko, W., Herron, S., & Peak, P. (1996). *Test of Early Written Language* (2nd ed.). Austin, TX: Pro-Ed.
- Ivernizzi, M., & Hayes, L. (2004). Developmental-spelling research: A systematic imperative. *Reading Research Quarterly*, 39, 216–228.
- Jewell, J. & Malecki, C. K. (2005). The utility of CBM written language indices: An investigation of production-dependent, production-independent, and accurate-production scores. *School Psychology Review*, 34, 27–44.
- Juel, C. (1988). Learning to read and writing: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 80, 437–447.
- Keiffer, R. D., & Faust, M. A. (1996). Portfolio purposes: Teachers exploring the relationship between evaluation and learning. *Assessing Writing*, 3, 149–172.
- Kelman, M. E., & Apel, K. (2004). Effects of a multiple linguistic and prescriptive approach to spelling instruction: A case study. *Communication Disorders Quarterly*, 25, 56–66.
- Keys, C.W. (2000). Investigating the thinking processes of eighth grade writers during the composition of a scientific laboratory report. *Journal of Research in Science Teaching*, 37, 676–690.
- Mackie, C., & Dockrell, J. E. (2004). Nature of written language deficits in children with SLI. *Journal of Speech, Language, and Hearing Research*, 47, 1469–1483.

- Masterson, J. J., & Apel, K. (2000). Spelling assessment: Charting a path to optimal assessment. *Topics in Language Disorders, 20*(3), 50–63.
- Masterson, J. J., & Apel, K. (2010). Linking characteristics discovered in spelling assessment to intervention goals and methods. *Learning Disability Quarterly, 33*, 185–198.
- Masterson, J. J., Apel, K., & Wasowicz, J. (2006). SPELL Spelling Performance Evaluation for Language and Literacy (2nd Ed.) [Computer software]. Evanston, IL: Learning by Design.
- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8*, 299–325.
- McMaster, K. L., & Campbell, H. (2008). New and existing curriculum-based writing measures: Technical features within and across grades. *School Psychology Review, 37*, 550–566.
- McMaster, K. L., Du, X., & Petursdottir, L. (2009). Technical features of curriculum-based measures for beginning writers. *Journal of Learning Disabilities, 42*, 41–60.
- McMaster, K., & Espin, C. (2007). Technical features in curriculum-based measurement in writing. *The Journal of Special Education, 41*, 68–84.
- Miller, L. (2009). Informal and qualitative assessment of writing skills in students with disabilities. *Assessment for Effective Intervention, 34*, 178–191.
- Moats, L. C. (1995). *Spelling: Development, disability, and instruction*. Baltimore: York Press.
- National Commission on Writing for America's Families, Schools, and Colleges. (2004, September). *Writing: A ticket to work...or a ticket out. A survey of business leaders*. New York: College Entrance Examination Board. Retrieved from http://www.writingcommission.org/prod_downloads/writingcom/writing-ticket-to-work.pdf
- National Commission on Writing for America's Families, Schools, and Colleges. (2008, April). *Writing, technology and teens*. Pew Internet and American Life Project. New York: College Entrance Examination Board. Retrieved from <http://www.writingcommission.org>
- Newcomer, P. L., & Barenbaum, E. M. (1991). The written composing ability of children with learning disabilities: A review of the literature from 1980–1990. *Journal of Learning Disabilities, 24*(10), 578–593.
- Northwest Regional Educational Laboratory (2004). *An introduction to the 6 + 1 trait writing assessment model*. Portland, OR: Author.
- Olinghouse, N. G., & Leaird, J. T. (2009). The relationship between measures of vocabulary and narrative writing quality in second- and fourth-grade students. *Reading and Writing: An Interdisciplinary Journal, 22*, 545–565.
- Olinghouse, N. G., & Santangelo, T. (in preparation). Generalizability of writing scores across genres.
- Paratore, J. R., Hindin, A., Krol-Sinclair, B., & Duran, P. (1999). Discourse between teachers and Latino parents during conferences based on home literacy portfolios. *Education and Urban Society, 32*, 58–82.
- Penner-Williams, J., Smith, T. E. C., & Gartin, B. C. (2009). Written language expression: Assessment instruments and teacher tools. *Assessment for Effective Intervention, 34*, 162–169.
- Persky, H. R., Daane, M. C., & Jin, Y. (2003). *The nation's report card: Writing 2002*. U. S. Department of Education, Institute of Education Sciences. Washington, DC: National Center for Education Statistics.
- Peterson, L. S., Martinez, A., & Turner, T. L. (2010). Test review: Process Assessment of the Learner-Second Edition. *Journal of Psychoeducational Assessment, 28*, 80–86.
- Powell-Smith, K. A., & Shinn, M. R. (2004). *Administration and scoring of written expression curriculum-based measurement (WE-CBM) for use in general outcome measurement*. Retrieved from AimsWeb website: <http://www.aimsweb.com/measures-2/written-expression-cbm>
- The Psychological Corporation. (2009). *Wechsler Individual Achievement Test-III*. San Antonio, TX: Author.
- Research Institute on Progress Monitoring (2009a). *Administration directions*. Retrieved from <http://progressmonitoring.org/probes/earlywriting.html>
- Research Institute on Progress Monitoring (2009b). *Scoring protocol: CBM for beginning writers*. Retrieved from <http://progressmonitoring.org/probes/earlywriting.html>
- Ritche, K. D. (2008). The building blocks of writing: Learning to write letters and spell words. *Reading and Writing: An Interdisciplinary Journal, 21*, 27–47.
- Ritche, K. D., Coker, D. L., & McCraw, S. B. (2010). A comparison of metrics for scoring beginning spelling. *Assessment for Effective Intervention, 35*, 78–88.
- Rosenblum, S., Weiss, P. L., & Parush, S. (2003). Product and process evaluation of handwriting difficulties. *Educational Psychology Review, 15*, 41–81.
- Salend, S. J. (1998). Using portfolios to assess student performance. *Teaching Exceptional Children, 31*(2), 36–43.
- Salahu-Din, D., Persky, H., & Miller, J. (2008). *The Nation's Report Card: Writing 2007* (NCES 2008–468). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C.
- Salvia, J., & Ysseldyke, J. E. (2004). *Assessment in special and inclusive education* (9th ed.). Boston, MA: Houghton Mifflin.
- Salvia, J., Ysseldyke, J. E., & Bolt, S. (2010). *Assessment in special and inclusive education* (11th ed.). Boston: Houghton Mifflin.
- Schoonen, R. (2005). Generalizability of writing scores. An application of structural equation modeling. *Language Testing, 22*, 1–30.
- Scott, C. M., & Windsor, J. (2000). General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech, Language, and Hearing Research, 43*, 324–339.
- Shinn, M. R., & Shinn, M. M. (2002). *Administration and scoring of spelling curriculum-based measurement (S-CBM) for use in general outcome measurement*. Retrieved from AimsWeb website: <http://www.aimsweb.com/measures-2/spelling-cbm/>
- Sperling, M., & Freedman, S.W. (2001). Research on writing. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 370–389). American Educational Research Association.
- Stein, N. & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In R. D. Freedle (Ed.), *Advances in discourse processes: Vol. 2. New directions in discourse processing* (pp. 53–119). Norwood, NJ: Albex.
- Stiggins, R. J. (2007). *Introduction to student-involved assessment FOR learning*. Upper Saddle River, NJ: Prentice Hall.
- Templeton, S., Bear, D. R., & Madura, S. (2007). Assessing students' spelling knowledge: Relationships to reading and writing. In J. R. Paratore and R. L. McCormack (Eds.), *Classroom literacy assessment: Making sense of what students know and do* (pp. 113–131). New York: Guilford Press.
- Tierney, R. J., Clark, C., Fenner, L., Herter, R. J., Simpson, C. S., & Wisner, B. (1998). Portfolios: Assumptions, tensions, and possibilities. *Reading Research Quarterly, 33*, 474–486.

- Tomchek, S. D., & Schneck, C. M. (2006). Evaluation of handwriting. In A. Henderson & C. Pehoski, C. (Eds.), *Hand function in the child: Foundations for remediation* (pp. 291–309). St. Louis, MO: Elsevier.
- Troia, G., Lin, S. C., Monroe, B. W., & Cohen, S. (2009). The effects of Writing Workshop instruction on the performance and motivation of good and poor writers. In G. A. Troia (Ed.) *Instruction and assessment for struggling writers: Evidence-based practices* (pp. 77–112). New York: Guilford Press.
- Trieman, R. (1997). Spelling in normal children and dyslexics. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 191–218). Mahwah, NJ: Erlbaum.
- Vaughn, S., Schumm, J. S., & Gordon, J. (1993). Which motoric condition is most effective for teaching spelling to students with and without learning disabilities? *Journal of Learning Disabilities*, 26, 191–198.
- Walther-Thomas, C., & Brownell, M. T. (2001). An interview with Bonnie Jones: Using student portfolios effectively. *Intervention in School and Clinic*, 36, 225–229.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge, UK: Cambridge University Press.
- Weintraub, N., & Graham, S. (1998). Writing legibly and quickly: A study of children's ability to adjust their handwriting to meet common classroom demands. *Learning Disabilities Research & Practice*, 13, 146–152.
- Weissenburger, J. W., & Espin, C. A. (2005). Curriculum-based measures of writing across grade levels. *Journal of School Psychology*, 43, 153–169.
- Wesson, C. L., & King, R. P. (1992). The role of curriculum-based measurement in portfolio assessment. *Diagnostique*, 18, 27–37.
- Whittaker, D., Berninger, V., Johnston, J., & Swanson, H. (1994). Intra-individual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences*, 6, 107–130.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson-III*. Itasca, IL: Riverside.
- Zaner-Bloser, Inc. (2010). *Zaner-Bloser Handwriting: Teacher Edition*. Columbus, OH: Author.
- Zimmerman, B., & Reiserberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology*, 22, 73–101.