Top Lang Disorders
Vol. 32, No. 4, pp. 297–318

The Developmental Writing Scale

A New Progress Monitoring Tool for Beginning Writers

Janet M. Sturm, Kathleen Cali, Nickola W. Nelson, and Maureen Staskowski

Developing writers make qualitative changes in their written products as they progress from scribbling and drawing to conventional, paragraph level writing. As yet, a comprehensive measurement tool does not exist that captures the linguistic and communicative changes (not just emergent spelling) in the early stages of this progression. The Developmental Writing Scale (DWS) for beginning writers was developed as a tool that can capture evidence of refined changes in growth over time. This measure is a 14-point ordinal scale that defines qualitative advances in levels of a learning progression for beginning writing from scribbling to cohesive (linguistically connected) and coherent (on an identifiable topic) paragraph-level writing. The measure can be used with young typically developing children and children with disabilities at all ages who are functioning at beginning levels of writing. Limitations of current writing measures, in contrast to the DWS, are described. The development of the DWS and techniques for using the measure are described with regard to construct and content validity. Preliminary research on reliability of DWS scoring and validity for 5 purposes support usefulness of the DWS for educational and research purposes, including monitoring the progress of beginning writers with significant disabilities.
Key words: *beginning writers*, *learning progression*, *writing assessment*, *writing scale*

Author Affiliations: *Central Michigan University, Mount Pleasant, Michigan (Dr Sturm); University of North Carolina at Chapel Hill (Ms Cali); Western Michigan University, Kalamazoo, Michigan (Dr Nelson); and Macomb Intermediate School District, Clinton Township, Michigan (Dr Staskowski).*

Portions of the Developmental Writing Scale were developed when the first author, as PI, was receiving support from NIH Grant R41HD059238-01 to Don Johnston, Inc. for development of software to support early writing development. The authors would like to thank the many teachers, undergraduate, and graduate students for their ongoing support during the development of the Developmental Writing Scale.

The authors disclose that they may receive royalties in the future for products based on the work described in this article.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.topicsinlanguagedisorders.com).

SROD UF DEF POT 1. Hir I whs sentn din rind. SROD UF DEF POT 1. (Grade 5 student)

My role model is my dad because he waled on tools, moshens, and matel baerols. I think it's verey intuorasting to do because I like to fix thing's, and it's fun to do waleding. (Grade 8 student)

Pow Pow Pow, I think it would be a good idea for teachers to have a gun permit. It would reduce violence in schools and outside of schools. It will also protect themselves as well as the students. I think it would be a great idea for it. (Grade 11 student)

Corresponding Author: *Janet M. Sturm, PhD, Central Michigan University, 2167 Health Professions Building, Mount Pleasant, MI 48859 (sturm1j@cmich.edu).*

DOI: 10.1097/TLD.0b013e318272159e

EACH OF THE students whose work is quoted previously (shared with permission from the Tennessee Department of Education, 2010) faces significant challenges for developing sophisticated, conventional writing skills. Although the students are at different grade levels, each presents as a "beginning writer." These students have something else in common—they all received a score of a 1 out of 6 on their most recent end-of-grade writing tests. Despite differences that can be readily observed across the three samples, the holistic scoring method used to evaluate them was not sensitive enough to detect the differences, nor would it be likely to reflect positive changes that would represent progress from one evaluation period to the next.

NEED FOR WRITING ASSESSMENT TOOLS FOR BEGINNING WRITERS

The Developmental Writing Scale (DWS) described in this article was developed in response to the need for better tools to assess beginning writing. Development of the DWS began in response to an immediate need on the part of researchers to track the progress in writing quality of beginning writers with developmental disabilities (DD) who were participating in an investigation of Enriched Writers' Workshop intervention (see Sturm, 2012). The research team needed a tool sensitive enough to detect small advances in writing by students with various diagnoses involving complex cognitive, linguistic, and neuromotor impairments. Each of the students in the target population was a beginning writer, and, subjectively, each student participating in the early research appeared to have made gains in intervention; however, the available writing measures were not sensitive enough to detect the changes. This illuminated the broader need for a writing measure that could be used to capture differences in beginner-level writing samples for children with disabilities across grade levels as well as for typically developing children in the early grades. Such a tool should be sensitive enough to measure advancing emergent writing abilities, descriptive enough to inform intervention planning, stable enough to measure progress reliably, and valid for addressing purposes related both to educational and research concerns.

Despite an emphasis on improving writing instruction for typically developing beginning writers, students with disabilities continue to lag behind their peers on statewide and national writing assessments. Data from the National Assessment of Educational Progress writing assessment showed nearly half (46%) of eighth-grade students with disabilities scoring below the basic level of proficiency (Salahu-Din, Persky, & Miller, 2008). Children with the most severe and complex disabilities may not even be represented in these national data because they were not considered to be writers. Tools that would be more sensitive to features of early writing than the holistic writing rubrics common to state writing assessments could help change this picture. For students with complex disabilities, who otherwise might never receive a score of more than 1 (below basic) out of 6 (proficient/exemplary) on current summative measures throughout their primary and secondary experience, such tools could improve access to achieving components of core curricular standards.

Quality writing instruction has been shown to improve the writing of students with disabilities (e.g., Graham & Harris, 2005; Joseph & Konrad, 2009), and ongoing progress monitoring plays a critical role in the development of quality writing instruction. Although currently available writing assessments might provide educators with sufficient information for summative assessments, they generally do not provide sufficient or appropriate information to guide the formative assessment process (Heritage, 2010).

An extensive review of the literature did not reveal any existing measures that would be sensitive to fine-grained differences in samples produced by beginning and atypical writers, with or without disabilities. Although several descriptions of beginning writing development were identified for component skills (e.g., Clay, 2006; Sulzby, Barnhart, & Hieshima, 1989), comprehensive scales of developmental writing skills appropriate for

this purpose were not found. As Coker and Ritchie (2010) concluded, "... no measure currently exists to bridge single-letter writing and spelling and beginning composition abilities" (p. 178). Thus, plans were set in motion to develop one. This is a report of the procedures used to develop a DWS for use in assessing samples of students' early writing attempts and the results of preliminary research on the scale's effectiveness.

WHO IS A BEGINNING WRITER?

For the purpose of this article, a beginning writer is one who is learning to use written language to express communicative intent, and beginning writing is defined as starting with emergent writing (drawing, scribbling, and writing letters) and ending with conventional writing abilities, usually acquired by second or third grade for typically developing children. Sulzby and her colleagues (Sulzby et al., 1989; Sulzby & Teale, 1991) defined beginning writers somewhat differently—as individuals who are in the early stages of learning to compose texts that can be read by other literate persons and that can be read conventionally by the writers themselves. This emphasis led Sulzby and her colleagues to focus on stages in the development of conventional spelling as characterizing early written language development. Although intelligible spelling clearly is a component of emergent written language, the boundary between emergent and beginning writing is not clear when spelling metrics are used alone. Consistent with the inclusion of drawing and scribbling in the scale of early writing development by Sulzby et al. (1989), beginning writing involves skills at language levels beyond emergent spelling. In this article, we consider students to be beginning writers who have not yet acquired rudimentary spelling but who have demonstrated other emergent written communication skills. This is consistent with a more inclusive view of candidacy for early literacy instruction proposed by Kaderavek and Rabidoux (2004).

Children as young as 18 months have been noted to make intentional marks on a page (Tolchinsky, 2006). By 3 years of age, children without disabilities may engage in emergent writing activities that include drawing, scribbling, and writing letters. As children continue to develop, they begin to differentiate between drawing and writing (Dyson, 1985, 1986), to form letter shapes, and to develop concepts about print, such as linearity and directionality (Clay, 1975; Ferreiro & Teberosky, 1982). Children then use alphabetic and syllabic principles to match letters to sounds, first randomly, then using invented spelling to represent initial and final sounds in words (Sulzby et al., 1989; Tolchinsky, 2006).

During this period, most children begin to understand that print can be used to communicate a message to an audience that is not present (Clay, 1975; Scott, 2012; Sulzby & Teale, 1991). By the end of second grade, most beginning writers are becoming conventional writers who can compose words and sentences that are intelligible to a reader (Kress, 1982/1994; Tolchinsky, 2006). They are also beginning to produce cohesive, coherent, elaborated texts consisting of multiple sentences (Fitzgerald & Spiegel, 1986, 1990; Halliday & Hasan, 1976; Langer, 1986; McCabe & Bliss, 2003; Newkirk, 1987, 1989; Peterson & McCabe, 1983). Definitions of the key concepts that are associated with learning to write are presented in Table 1.

BEGINNING WRITERS WITH DISABILITIES

Among students with disabilities, the writing abilities of students with learning disabilities (LD) have been investigated more extensively than perhaps any other group. Research examining the writing products of students with LD has shown that these students demonstrate difficulty with handwriting, spelling, vocabulary, complex sentence constructions, fluency, text structure, cohesion, and coherence (Ehren, 1994; Graham, MacArthur, Schwartz, & Page-Voth, 1992; MacArthur, Schwartz, & Graham, 1991; Newcomer & Barenbaum, 1991; Scott, 1989).

A reasonable body of research also has highlighted the writing challenges and needs of

Table 1. Definitions of key terms and constructs

Construct	Definition
Drawings	A line drawing or photograph representing an event, object, person, or place
Scribbles	A wavy, circular, or continuous line that may, or may not, show directionality
Letter-like forms	One or more forms representing or resembling printed or cursive alphabetic letters
Words	A group of letters written in a sequence set off by spaces; includes intelligible words not separated by spaces and adjacent to random letters or other intelligible words
Partially formed sentence	At least two words in proximity that appear to be related grammatically as parts of a sentence
Complete sentence	A set of words organized grammatically with a subject and a verb (punctuation not required)
Organized	Discourse that conveys temporal, causal, categorical, or other logical relationships that are consistent with the author's apparent purpose in conveying information, narrating a story, making a persuasive argument, or some other emergent discourse form
Coherence	A central main theme or topic maintained across multiple sentences
Cohesion	Intra- and intersentence language connections made by using cohesive devices (e.g., pronoun or synonym replacement, logical connectors, conclusions that refer to prior content); one test of cohesion is that sentences cannot be reordered without changing meaning

Note. From "Cohesion in English," by M. Halliday and R. Hasan, 1976, London: Longman; Newkirk (1987); "The achievement and antecedents of labeling," by A. Ninio and J. Bruner, 1978, *Journal of Child Language*, 5(1), pp. 1-15. Copyright 2010 by Janet Sturm; and "Forms of Writing and Re-rereading From Writing: A Preliminary Report (Technical Report No. 20)," by E. Sulzby, J. Barnhart, and J. Hieshima, 1989, Berkeley, CA: National Center for the Study of Writing and Literacy. Retrieved November 15, 2010, from http://www.nwp.org/cs/public/print/resource/606
From "Outcome Measures for Beginning Writers With Disabilities," by J. M. Sturm, N. W. Nelson, M. Staskowski, and K. Cali, 2010, November, Philadelphia, PA: Miniseminar presented at the American Speech-Language-Hearing Convention; used with permission of the author.

students with severe speech impairments and physical disabilities who use augmentative and alternative communication (AAC) techniques and supports and who may be described as having complex communication needs (CCN). Slow writing rates (i.e., around 1.5 words per minute) are reported for individuals using AAC, which has a significant impact on writing fluency and makes composing extremely difficult (Koke & Neilson, 1987; Smith, Thurston, Light, Parnes, & O'Keefe, 1989). Many students with CCN also demonstrate delays in phonology, spelling, vocabulary, syntax, and discourse knowledge that impact their writing development (Berninger & Gans, 1986; Harris, 1982; Nelson, 1992; Sturm & Clendon, 2004; Sturm, Erickson, & Yoder, 2003; Udwin & Yule, 1990; Van Balkom & Welle Donker-Gimbrere, 1996; Vandervelden & Siegel, 1999).

Less research is available on the writing abilities and challenges of students with intellectual developmental disabilities (IDD) and social–communicative disabilities, such as autism spectrum disorders (ASD). Disabilities such as IDD and ASD, individually or in combination, present risks to written communication development on multiple levels (cognitive, communicative, and linguistic). Linguistically, students with IDD and ASD can present with a wide range of abilities (Prelock, 2006). Some students with IDD and ASD are unable to produce spoken or written words at all, whereas others demonstrate relative

strengths in these areas. Students with IDD are at risk for demonstrating difficulties across features of writing (e.g., spelling, vocabulary, syntax) (Sturm, Knack, & Hall, 2011). Many students with ASD present with fine motor limitations that impact text production (Broun, 2009), which can contribute to the production of texts that lack fluency and intelligibility. Impairments in social interaction in students with IDD and ASD might also interfere with these students' understandings and production of communicative aspects of written discourse, although limited research is available in this area.

Students with severe and multiple disabilities risk remaining beginning writers for life. This risk has been compounded historically by a lack of serious effort and expectation on the part of educators to teach such children to write (Koppenhaver & Yoder, 1993). That picture has been changing in recent years (e.g., Bedrosian, Lasker, Speidel, & Politsch, 2003; Koppenhaver & Erickson, 2003), spurred on, in part, by policy-driven expectations for special educators to target achievement of actual academic standards with their students with severe and complex disabilities. This heightens the need for an assessment tool that could quantify the indicators of small units of progress that characterize the early steps in learning to communicate through writing.

CURRENT ASSESSMENT PRACTICES AND LIMITATIONS

As an initial step in the development of the DWS, existing writing measures were reviewed to identify the writing constructs targeted, the developmental range of the measure, and the limitations. Espin, Weissenburger, and Benson (2004) classified classroom writing assessments as holistic, primary trait, and analytic scoring. They contrasted such "typical" assessments with curriculum-based measurement, which they indicated as being "developed with special education in mind" (p. 56).

Holistic scoring criteria are commonly used to score samples produced for end-of-grade writing assessments. Such assessments may be found on Web sites describing statewide assessments (e.g., Massachusetts Department of Education, 2012; Tennessee Department of Education, 2012). They tend to focus on whether or not students have acquired specific writing traits during the course of a school year, using separate rubrics for each grade level. Writers may be scored for levels of proficiency on scales of 1–6, with scores of 1, 2, or 3 denoting lack of proficiency and scores of 4, 5, and 6 denoting proficiency for that grade level.

Primary trait and analytic scoring may also be used in end-of-grade state-level assessments. Especially at the lowest levels, methods of this type may describe traits that are absent, rather than present. An example is the description for kindergarten, Level 1 that indicates, "Writing/drawing shows little or no development of the topic" (Michigan Department of Education, 2000). Some rubrics for beginning writers also seem to equate very different constructs, which may not be interchangeable, such as writing and drawing within the early levels of the Michigan Literacy Progress Profile (Michigan Department of Education, 2000).

Primary trait and analytic scoring may also be used more appropriately for rating samples produced by children functioning beyond the early writing stages. We were seeking appropriate informal assessments of students' original writing samples at the earliest developmental levels. As possible candidates, we identified 2 types of measurement tools that might be suitable—developmental writing continua and curriculum-based measurements.

Developmental writing continua

Developmental writing continua tend to focus on students' acquisition of positive traits in their writing across the primary grades (e.g., Beginning Writer's Continuum; Northwest Regional Educational Laboratory, 2010; North Carolina K-2 Writing Continuum, North Carolina Department of Public Instruction, 2009). Developmental rubrics within such continua tend to follow Sulzby et al's. (1989)

initial stages of writing development, from early emergent to independent/conventional writing and spelling. They describe how young students move from drawing to letter formation to words and sentences. Several developmental writing measures (e.g., the Beginning Writer's Continuum) also measure the development of writing traits (e.g., organization, word choice, and conventions).

Some developmental continua include evidence of student behaviors, such as, "pretends to read own writing," and writing processes, such as, "reads own writing with fluency" (North Carolina Department of Public Instruction, 2009). This is helpful for some purposes, but such measures cannot be used for scoring written products by themselves.

Developmental continua may also lack sufficient refinement to document minimal changes in student writers. For example, progressive levels of the North Carolina K-2 Writing Continuum include descriptors that are nearly identical at the Emergent level, "writes 1 or 2 sentences focused on a topic," the Early Developing level, "writes a few short, patterned, repetitive sentences focused on a topic," and the Developing level, "writes several sentences about a topic." Other problems occur when descriptors are vague, such as, "settles for the word or phrase that will do" and "sections of writing have rhythm and flow" (Northwest Regional Educational Laboratory, 2010), making it difficult for teachers to score reliably. Developmental continua may also be difficult to interpret because they assign single scores to levels on the basis of multiple constructs across a range of potentially unrelated skills. One example is, "uses periods correctly" along with "establishes a relationship between drawing and print" (North Carolina Department of Public Instruction, 2009). In these instances, two writing samples scored at the same level actually may have very different attributes.

Curriculum-based measures of written expression

Curriculum-based measures of written expression have advantages for older students with disabilities (Espin et al., 2004) and may be useful for beginning writers as well, but they also have limitations. Such measures are designed to monitor students' progress frequently over time by administering timed, on-demand prompts of 3–5 min duration on a weekly or biweekly basis and using count data to document changes. Formats for assessing the writing of beginning writers in kindergarten and first grade include (a) sentence copying, (b) picture-word prompt, (c) story prompt, and (d) sentence writing (Coker & Ritchie, 2010; McMaster et al., 2011). Production-dependent quantitative (count) measures of writing fluency, such as total words written (TWW), and correctness, such as words spelled correctly (WSC) and correct writing sequences (CWS), may be more reliable measures of beginning writing samples than production-independent measures of accuracy (Jewell & Malecki, 2005); however, these quantitative measures often lack face validity with teachers (Gansle, Noell, VanDerHeyden, Naquin, & Slider, 2002) and just adding more of something has limited value for guiding instructional choices.

Curriculum-based measures of written expression measures also are problematic for beginning writers who cannot yet produce more than three intelligible words, as is common at least through the middle of kindergarten. Thus, count measures of TWW and written word accuracy may not be appropriate measures for kindergarten students (Coker & Ritchie, 2010) and other beginning writers. Similarly, timed, on-demand writing prompts may not be appropriate for beginning writers who are still mastering transcription and idea generation skills (Bereiter, 1980) and for students whose disabilities make it difficult to perform motor acts quickly.

In summary, the review of existing options made it clear that a new tool, such as the DWS, was needed to achieve multiple goals for the target population. To be effective with children with more severe and complex disabilities, most of whom make slower gains in writing over time, the new tool needed to accurately represent the fine-grained differences

in early development, spanning from emergent to early conventional writing. It also needed to focus on abilities that are acquired across graduated levels of early writing development and be designed to capture the differences in written products produced by children functioning at different levels within the earliest stages of beginning writing.

PURPOSES FOR THE DWS

To fill the unmet need, we conceptualized the DWS as a new tool that would be valid for measuring the written language of students functioning at the earliest levels of writing development. Because validity must be measured relative to the question, "Valid for what purpose?" we articulated five purposes for the DWS. It should (a) identify small differences in beginning writing skills, (b) offer instructionally relevant information about what to target next, (c) serve as a functional outcome measure for periodic assessment probes and classroom-produced writing artifacts, (d) be easy for educators to learn and use reliably, and (e) quantify evidence of small but significant changes so that educators can celebrate growth with students and their parents. Methods of traditional test theory were used to build in construct and content validity during development and to evaluate the DWS with regard to its reliability and validity for its five purposes. Research questions were posed about (a) how to represent the constructs of early written communication development, (b) whether the new scale would compare favorably with existing measures in terms of its ability to capture evidence of small advances in early writing, (c) whether it would be possible to use reliably, and (d) whether it would meet teachers' needs for a tool they could use to measure progress and guide instruction.

METHODS

Procedures for developing and evaluating the scale

Development of the DWS was conducted in four recursive steps aimed at meeting standards for educational and psychological testing established by the joint committee of the American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (AERA/APA/NCME, 1999). They were to (a) clarify theoretical model of constructs to be measured and the purposes of scale, (b) generate items or scoring criteria consistent with the model and representative of its content, (c) conduct recursive tryouts and modifications until the measure could be scored reliably, and (d) evaluate whether the measure could fulfill its stated purposes.

Step 1: Clarify theoretical constructs and purpose

A developmental progression provided the theoretical framework for the DWS, using constructs described previously in the background section of this article. Key constructs to be represented in the scale included drawing and scribbling, production of print, demonstration of alphabetic and syllabic principles, concepts of words with spaces, formulation of sentences, and production of cohesive and coherent texts. Definitions of these key constructs are provided in Table 1. This step also involved generating the five purposes for the DWS.

Purpose 1

The first purpose was to distinguish variations in beginning writing skills. As reviewed previously, a danger is that writing measures at the emergent level may capture only what a child *cannot* do. If assessment tools assign a low score based on a student's limited skills (e.g., "insufficient to score" or "undeveloped"), they may fail to capture emerging indicators of the child's growing awareness that writing allows one to share and communicate ideas through text (e.g., Scott, 2012). Under deficit-oriented systems, students with disabilities who produce writing at beginning levels of development may be viewed as "nonwriters." It is critical that educators and special service providers be able to measure and show what a student *can* do with text.

Traditional writing assessments also may not have enough sensitivity to detect refined changes as a student moves from single letter writing to more conventional skill development. Thus, finer-grained descriptors were added to existing developmental continua consistent with preliminary samples of writing gathered from the target population. Samples were also drawn from typically developing students. Every child is a beginning writer in the early grades. The DWS is designed to be used with a beginning writer of any age, but it is grounded in the writing development of typically developing children between the ages of 3 and 7 years.

Purpose 2

The second purpose was to serve as a formative assessment measure that could support teachers in identifying instructional goals that would help individual students move to the next level in development. Traditional measures might provide a quantitative measure (e.g., TWW) for a particular product; but without indicators of what should come next, teachers might lack clarity about what should be targeted next. The DWS is based on quantitative additions to the written product as well as qualitative ones, such as movement from three *unrelated* words to three *related* words. An educator using this system then could choose instructional goals designed to improve students' writing quality and progress to the next level, such as "The student will connect two to three words to convey sentence-like meanings while writing." Instruction to support achievement of that step could include having the child tell (orally or gesturally) something about a chosen writing topic first. Then, a shared pencil approach or keyboarding with scaffolding could be used to help the child to represent the connected ideas with intelligible words in writing.

Purpose 3

The third purpose was to provide a means of measuring either periodic probes of students' independent writing abilities or naturally occurring writing artifacts that students compose within their classroom writing activities. As a functional outcome measure, the DWS was intended to serve also as a formative assessment tool to assign scores to writing samples produced across time. This would support educators in profiling growth in the work of a student writer. The goal was to design the DWS so that a student participating in a classroom writing program might progress across its levels, such as moving in small steps from single word writing to writing using multiple sentences, over a time frame in which traditional measures would reflect no growth at all. The measure was also designed to be used frequently, on a weekly, monthly, or quarterly basis to monitor progress on existing written samples so that teachers would not have to take time out for "testing."

Purpose 4

The fourth purpose was to provide educators with a measure that is time efficient and easy to use reliably. By designing the DWS to be applicable to existing artifacts with a scoring system that educators could apply quickly, we hoped that time savings would make it possible for teachers to examine many samples for multiple students over time. We reasoned that the instructional relevance of the tool would be enhanced if it were intuitive for educators to learn and to apply quickly to distinguish differences between writing samples.

Purpose 5

The final purpose of the DWS was to make it possible to celebrate students' positive change as writers with them and their parents. This goal stemmed from observations that, far too often, beginning writers who struggle throughout their school-age years with basic writing skills may not be aware of the positive gains they actually are making in learning to communicate through writing. The DWS was developed as a simple ordinal scale so that it could measure writing advances as higher numbers. This would make it possible to show students, through the use of graphs or tables, what they have accomplished with

writing. By plotting data across time, a simple numerical scale could support students in setting personal goals and talking with others (e.g., educators or family members) about their writing strengths and gains they have made.

Step 2: Generate criteria for scoring content

For a norm-referenced test, items must be generated that provide a representative sample of content for each construct being assessed. For a criterion-referenced scale, such as the DWS, the content is generated by students, who produce original writing samples, and the scoring criteria must be generated to capture the features of the content. In our work, this step involved a combination of literature review and consideration of existing empirical data in the form of students' written language samples. An extensive review of the literature was conducted on the development of typically developing writers and a list of initial constructs (e.g., scribbling or random letter patterns) produced by children during text generation was created. The list was used to create a preliminary hierarchy of levels of writing quality. Because a core goal was to create a scale that ranged from emergent to conventional writing skill development, a bank of 460 naturally occurring writing samples produced by typically developing kindergarten and first grade students was used to validate the accuracy of the tool levels and to anchor each level on the DWS. In addition, samples of students with LD between fourth and eighth grades were reviewed to verify the levels on the scale.

Step 3: Conduct recursive tryouts and modify criteria as needed

The initial scoring criteria were revised multiple times after reviewing the scoring of writing samples that had been produced by students with DD. During this process, the research team revised descriptions of linguistic features and skills linked to student writing samples. This initial scale was then used by the authors to code, as a group and then independently, a wide range of beginning writing samples. The first author also assembled a cadre of undergraduate and graduate students to collect and code samples to allow comparison of scores for samples produced sequentially.

In the process of development, samples coded with evolving versions of the tool included more than 1500 samples produced by students with DD (ages 5–25 years) and more than 200 samples of typically developing students. During this coding process, outlier samples were identified, reviewed by the authors as a group, and additional coding rules were created. If a particular sample was not currently represented on the scale, a new level was created. If a sample could not be reliably coded, we revised the descriptors for a given level. The process of creating the DWS, thus, was recursive in nature, and refinements were made to resolve difficulties in coming to initial agreement.

A substudy was also conducted on interrater reliability in which the DWS was used to score 285 samples produced by 11 students with disabilities. All samples were scored initially by one graduate student, a research assistant who was trained and experienced in the use of the scale. Then, 20% of the samples were rescored by a second graduate student, who was trained to use the scale as part of a graduate course. Percent agreement and Cohen's κ were calculated to determine the level of agreement between scorers beyond chance agreement (Cohen, 1960; Hayes & Hatch, 1999).

Step 4: Evaluate the tool's validity for its primary purposes

Construct and content validity were partially ensured by the steps used to generate the components of the tool. Evaluation of the tool's validity for achieving its five purposes is ongoing. This includes a large-scale validation study (Cali, manuscript in preparation), in which writing samples of typically developing kindergarten and first-grade students are being used to test the sensitivity of the DWS to advances in children's writing (Purposes

1–3). The first author is also using the scale with children with various disabilities in an evaluation of an Enriched Writers' Workshop approach to understand better how it works for guiding instruction and documenting progress (Purposes 2, 3, and 5) (see Sturm, 2012).

Preliminary evaluation of the validity of the DWS for meeting its purposes was also conducted by comparing the DWS with existing developmental scales from Clay (2006) and Sulzby et al. (1989). A goal of DWS development was to build on existing tools and also to improve on them by capturing developmental linguistic changes in writing quality; thus, the constructs measured should be related, but not exactly the same. Unlike traditional evaluation of concurrent validity, in which one looks for commonalities between the new tool and existing tools, we also were looking for distinctions. To test further for distinctions, existing writing samples were coded using the DWS and the two additional scales (i.e., Clay, 2006; Sulzby et al., 1989), and comparisons were made between results.

Finally, a small pilot survey study was conducted to investigate the validity of the DWS for its purpose of being easy for educators to learn and useful for them to apply. Using a protocol approved by a Human Subjects Institutional Review Board, professionals who had been helping to pilot test the scale were invited to participate in the survey. The 8-item examiner-created Likert-style questionnaire used the choices, *very much agree, agree, undecided, disagree, and very much disagree*. Open spaces were provided following each item to allow for written comments. Item 8 on the questionnaire used a categorical response format (*daily, weekly, biweekly, monthly, and quarterly*) to obtain the estimated frequency of use of the DWS when examining student writers.

RESULTS

The developmental writing scale

The process of recursive tryouts and modifications resulted in the current version of the DWS, with 14 writing levels and scoring criteria as outlined in Table 2. This table presents the ordinal 14-point developmental scale, standardized scoring criteria, examples, and accommodations related to each level. Using this scale, examiners assign level "scores" to samples of original written products produced by beginning writers with or without disabilities.

An important feature of the DWS is that it is not genre specific. Thus, it can be used to score and measure any genre (e.g., narrative or expository) composed by a beginning writer. These should be samples of original text production, not immediately influenced by teacher or clinician scaffolding. During their production, students should be allowed access to any accommodations (e.g., bins of pictures that might stimulate topic selection and/or access to a keyboard to compose text) while producing the samples to be scored. Examples of how accommodations might be accounted for in scoring appear in Table 2.

Comparisons of the DWS with other tools

Several comparisons were made to illuminate similarities and differences with existing tools. Table 3 shows how scoring levels on the DWS correspond to comparable language levels for writing by Clay (2006) and forms of writing by Sulzby et al. (1989), illustrating gaps in the other measures that the DWS could fill. Copying was the only construct on one of the other measures not represented in the DWS scale, and copying is not a construct that is appropriate when the goal is to encourage students to produce original writing samples. This table also demonstrates that the DWS, which targets qualitative growth in linguistic development, aligns more closely with Clay's (2006) language levels for writing than with the early forms of writing described by Sulzby et al., which emphasized spelling development.

Table 4 shows how the results would differ if the three different systems were applied to the same samples. It illustrates how

Table 2. Developmental writing scale for beginning writers[a]

Levels	Scoring Criteria	Description	Accommodations
1	Drawing	Lines and curves that appear to represent objects	Selection of a picture by a child who cannot hold a traditional pencil or marker
2	Scribbling	Continuous vertical, circular, or wavy lines arranged linearly across the page, which may include letter-like forms, but with the majority of shapes not recognizable as letters.	If a child uses a keyboard, this level would not be used
3	Letter strings (no groups)	Handwritten or typed strings of letters but not grouped into words. Examples: *ttttt* *kshpppns*	Alphabet display (e.g., paper copy) and standard or electronic keyboard access (e.g., on screen keyboard or AAC system)
4	Letter strings grouped in words	Strings of letters grouped into "words" (i.e., with spaces between at least two groups of letters) but with no intelligible words. Example: *iLCR6a iLKVKC CPRSB WRKe BRKe*	Alphabet display (e.g., paper copy) and standard or electronic keyboard access (e.g., on screen keyboard or AAC system)
5	One intelligible word	Strings of letters grouped into "words," with only one possible real word (i.e., two or more letters in length) set apart, written repeatedly (e.g., dog, dog, dog), or embedded in a string of letters. Example: *IMPlCOTheC (I am playing outside on the swing.)*	Word bank or word prediction software
6	Two to three intelligible words	Two or three different intelligible words embedded in strings, separated by spaces, or in a list format. Single letter words such as "I" and "a" must be separated by spaces to count as an intelligible word. Example: *IYTKTOSMNTHETR (I like to swim under the water.)*	Word bank or word prediction software
7	Three or more different intelligible words in a list	Three or more related words. Example: *Lions Detroit football*	Word bank or word prediction software

(*continues*)

Table 2. Developmental writing scale for beginning writers[a] (*Continued*)

Levels	Scoring Criteria	Description	Accommodations
8	Partial sentence of more than three words	More than three different intelligible words, with at least two of them in a partially formed sentence (i.e., grammatically related parts of a phrase, clause, or sentence). Example: *MYDADDYWASLIEGAGARILA (My daddy was like a gorilla.)*	Word bank or word prediction software
9	One to two complete sentences	Sentences have a subject phrase and a verb phrase. End punctuation is not necessary. Example: *I am hpe Easter is here. I cw the Easter bnny. (I am happy Easter is here. I saw the Easter bunny.)*	Word bank or word prediction software
10	Three or more unrelated sentences (neither coherent nor cohesive)	Sentences have no coherent topic (i.e., sentences are not related) *I play a game. I went to my fnid house. I went to get a egg to eat. I went to chansh on Sun day. I kiss my momer sun day. I can walk my dog. I sat in my house. I went to the saing in ring.*	Word bank or word prediction software
11	Three or more related sentences (coherent but limited cohesive)	Organized writing with three or more sentences on a coherent topic but with limited cohesion between sentences (i.e., sentences can be reordered without changing meaning). Example: *Frogs are eggs. Frog are cool. I no how a frog grows egg then grow mory. Frog eat lot of things that we don't eat like bugs. I want a frog to play with. I thak frogs are mumloss because thae swim.*	Word bank or word prediction software
12	Three or more related sentences that cannot be reordered (coherent and cohesive)	Organized writing with a coherent topic (i.e., on a consistent theme) and use of cohesive devices (e.g., pronoun or synonym replacement, logical connectors, subordinating conjunctions, conclusions that refer to prior content) across three or more sentences so that sentences cannot be reordered without changing meaning (see Supplemental Digital Content [available at http://links.lww.com/TLD/A10] Appendix A [available at http://links.lww.com/TLD/A10] for examples)	Word bank or word prediction software
13	Two coherent paragraphs of at least three cohesive sentences each	Organized writing with a coherent main topic and two cohesive subsections (subtopics or story parts), with at least two sentences elaborating the meaning of each (see Supplemental Digital Content Appendix A [available at http://links.lww.com/TLD/A10] for examples)	Word bank or word prediction software
14	Three or more coherent paragraphs of at least three cohesive sentences each	Organized writing with a coherent main topic and at least three cohesive subsections (subtopics or story parts), with at least two sentences elaborating the meaning of each (see Supplemental Digital Content Appendix A [available at http://links.lww.com/TLD/A10] for examples)	Word bank or word prediction software

Note. AAC = augmentative and alternative communication.

[a]Definitions of key terms are provided in Table 1. If debating between two levels, assign the lower level.

From "Outcome Measures for Beginning Writers With Disabilities," by J. M. Sturm, N. W. Nelson, M. Staskowski, and K. Cali, 2010, November, Philadelphia, PA: Miniseminar presented at the American Speech-Language-Hearing Convention; Revisions copyright 2012 by J. Sturm, K. Cali, N. Nelson, & M. Staskowski used with permission of the authors.

Table 3. Scoring correspondence for the developmental writing scale, with language levels for writing by Clay (2006), and forms of writing by Sulzby et al. (1989)

DWS Level	Clay's Language Level	Sulzby's Forms of Writing
1 Drawing	–	1 Drawing
2 Scribbling	–	2 Scribble—wavy 3 Scribble—letter-like
3 Letter strings (no groups)	1 Alphabetical (letters only)	4 Letter-like Units 5 Letters—random
4 Letters grouped in "words" with spaces	–	6 Letters—patterns 7 Letters—name-like elements
–	–	8 Copying
5 One intelligible word	2 Word (any recognizable word)	9 Invented spelling—syllabic
6 Two to three intelligible words	–	10 Invented spelling—intermediate
7 More than three intelligible words in a list	–	11 Invented spelling—full 12[a] Conventional spelling
8 Partial sentence	3 Word group (any two-word phrase)	–
9 One to two sentences	4 Sentence (any simple sentence)	–
10 Three or more sentences (not coherent)	5 Punctuated story (of two or more sentences)	–
11 Three or more sentences (coherent but limited cohesive)	–	–
12 Three or more sentences in one paragraph (coherent + cohesive)	–	–
13 Two paragraphs (coherent + cohesive)	6 Paragraphed story (two themes)	–
14 Three paragraphs (coherent + cohesive)	–	–

Note. DWS = Developmental Writing Scale.
[a]Sulzby et al. (1989) also included a Level 13, described as "other," which is not represented in this table.

Table 4. Samples produced by beginning writers, coded for writing development by four methods

Writing Sample[a]	End-of-Grade Test Score	Language Levels (Clay, 2006)	Forms of Writing (Sulzby et al., 1989)	DWS
AT CAT DOG	–	Level 2 (recognizable word)	Level 12 (conventional spelling)	Level 6 (two to three intelligible words)
FoRstIll TaLe You weN ItAPPAND She Was ALIITL bAbY So ShE ASCTHAR MothR BUtHeRMOtHR Sad She was To LIITL NOW Ill TeL You wot MADE HOR want to DO It Beecose she HAD A BIG BROtHR AND He CODRIDE HIS AND that's ALL	–	Level 5 (two or more sentence story)	Level 11 (invented spelling)	Level 12 (more than three sentences in a paragraph; coherent + cohesive)

(*continues*)

Table 4. Samples produced by beginning writers, coded for writing development by four methods (*Continued*)

Writing Sample[a]	End-of-Grade Test Score	Language Levels (Clay, 2006)	Forms of Writing (Sulzby et al., 1989)	DWS
SROD UF DEF POT 1. Hir I whs sentn din rind. SROD UF DEF POT 1.	Score 1	Level 1 (letters only)	Level 5 (random letters)	Level 4 (letters grouped in words)
My role model is my dad because he waled on tools, moshens, and matel baerols. I think it's verey intuorasting to do because I like to fix thing's, and it's fun to do waleding.	Score 1	Level 5 (two or more sentence story)	Level 11 or 12 (invented or conventional spelling)	Level 11 (3+ sentences—coherent but limited cohesion)
Pow Pow Pow, I think it would be a good idea for teachers to have a gun permit. It would reduce violence in schools and outside of schools. It will also protect themselves as well as the students. I think it would be a great idea for it.	Score 1	Level 5 (two or more sentence story)	Level 12 (conventional spelling)	Level 12 (more than three sentences in a paragraph; coherent + cohesive)

DWS = Developmental Writing Scale.

[a]The first two samples in this table were used by Sulzby et al. (1989); the final three are from the Tennessee Department of Education (2010) and are used, with permission, to introduce this article.

the linguistic and communicative constructs that underpin the DWS categories differ from the spelling accuracy of the forms of writing described by Sulzby et al. (1989) as moving from drawing and scribbling to conventional spelling. The DWS, in contrast, is meaning-based, focusing on linguistic development at the letter, word, sentence, and paragraph level. This difference in focus is revealed in scoring of the two writing samples drawn from Sulzby et al. that are shown in this table.

Table 5 shows how the results would differ for three of these samples using the analytical scoring method of TWW compared with using the 14 levels of the DWS. Using total words, Sample A, which has the least number of words (24), would appear to be the least sophisticated of the three writing samples. However, using the DWS, Sample A is assessed as the most sophisticated because it demonstrates organized writing on a topic that is both coherent and cohesive. In contrast, Sample B, with 34 words, is evaluated as the least sophisticated sample within this set because the sentences are not related, giving it a rating of Level 10. Sample C, on the other hand, is more cohesive than Sample B because each sentence is about the same topic, earning a Level 11 score. Sample A is rated as being more coherent than Sample C because it has a clear beginning, middle, and end, and its sentences cannot be rearranged without changing the meaning of the passage; therefore, it is rated as a Level 12.

The comparisons in Table 5 also show that DWS ratings could provide specific information on a student's current developmental level and direction for what instruction should target to help the student move to the next level. This supports the validity of the tool for meeting Purposes 2, which is to provide a *formative assessment measure that is instructionally relevant*, and 3, which is to *provide a means of measuring either periodic probes of student's independent writing abilities or naturally occurring writing artifacts that students compose within their classroom writing activities.* Related to Purpose 2, for example, the next goal for student A would be to target writing an organized, coherent topic containing two cohesive subsections. Work with student B could target writing three or more topically related sentences that are organized and coherent. Student C would need a goal aimed at producing organized writing on a topic that is both coherent and cohesive.

Scoring reliability and specialized scoring rules

The results of the substudy of inter-rater reliability based on independent scoring by two trained graduate assistants showed a percentage agreement of 91%. The correlation

Table 5. Writing samples comparing total words to the DWS level

Writing Sample	Text	Total Words Written	DWS Level
A	On monday my frid came over my house. We played and we had fun. She lath. She what houm I clin up my mast.	24	12
B	Happy Birthday Matthew. I like chocolate please Mom. I have a new school. Am 14. A new pet is a puppies and a dog and a cat and a shirt and a new baby.	34	10
C	I love to watch the garbageman to pick up our trash can to. I don't watch the garman out to my window to. I love to watch the recycling person to get my recycling from my house to	38	11

Note. DWS = Developmental Writing Scale.

between scores using Cohen's κ (Cohen, 1960; Hayes & Hatch, 1999) also was strong, with $\kappa = .898$ ($p < .001$).

The procedures for developing the DWS also revealed some areas in which challenging situations called for special scoring rules. In general, the score assigned to the sample is the one that best fits the description at a particular level. We found that scoring agreement could be improved if a scorer debating between two levels assigned the lower level being considered. Another rule established through this process was that the scorer should focus on the nature of the student's writing (or prewriting) and not the spatial placement of text on a page (e.g., paragraph spacing, indentation, or margins). In addition, we found it helpful to remind ourselves that the concepts of word, sentence, and paragraph represented in this scale are meant to be primarily linguistic in nature. One should look beyond technical accuracy when assigning scores. For example, if a student produces one large paragraph, examination may reveal that three cohesive and coherent subsections are present and a Level 14 is the best score. Another student might have a true word (e.g., *the*) embedded within random letters. This student would be assigned a Level 5. If the same word is repeated in a list format (e.g., *dog, dog, dog*) the student also would be assigned a Level 5. Student names at the top of the page (denoting who wrote it) are not counted; however, student names in the body of the text are scored on the scale.

Another challenging scoring element relates to judgments of word intelligibility. DWS scoring allows examiners to use graphic content, such as hand-drawn pictures or pictures selected from a picture bank, to support "reading" of the student's text. Figure 1 provides an example in which the picture in Sample 1 makes it possible to detect the words, "to" and "field trip," and the picture in Samples 2 makes it evident that the student was writing "I'm playing basketball." A caution is that examiners should use graphic content only (i.e., context embedded in the work to communicate with an absent audience) to aid in interpreting children's text. To the extent possible, they should avoid being influenced by additional context provided orally by the student from the author's chair or in face-to-face communication about the work because such context would not be available to an absent audience. The rationale is that the scoring should be based on the messages that can be gleaned by a remote audience assessing the written artifacts the student has produced only, and not oral or gestural communication.

Evidence from the pilot survey for DWS ease, efficiency, and utility

The brief pilot survey on the ease and utility of the DWS was completed by two teachers and two speech–language pathologists

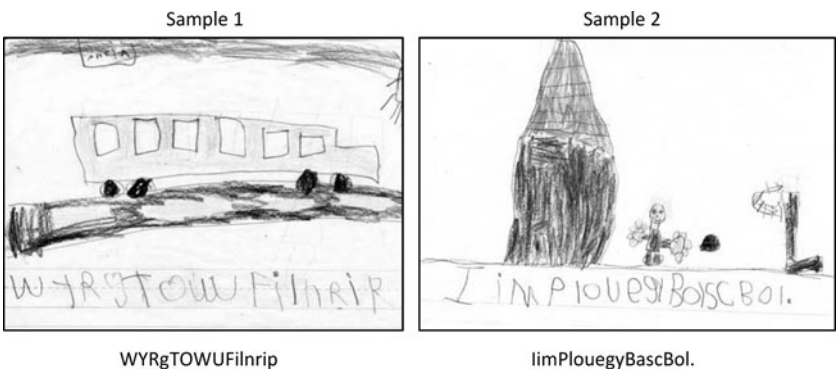

Figure 1. Examples of students' writing in which graphic illustrations contribute to the intelligibility of students' written content.

(SLP). All four participants had been using the DWS to examine writing samples produced by students with DD and CCN across a 1-year period. Results of this small pilot study are summarized in Table 6, which also provides excerpts from qualitative comments written in the open-ended sections of the survey. Overall, the teachers' and speech–language pathologists' reported perceptions of the DWS indicated that they found it easy, efficient, and useful when used to examine the writing outcomes of students with DD.

DISCUSSION

The DWS described in this article was designed as a comprehensive measure of qualitative change in beginning writers that can capture refined changes in growth over time. The construct and content validity of the tool are supported by the recursive development process with foundations in existing literature on early writing development (e.g., Sulzby et al., 1989), modified on the basis of empirical evidence drawn both from young typically developing students and students with severe and complex disabilities of a wide range of ages. These results provide preliminary evidence for the summative and formative uses of this assessment tool for quantifying developmental advances in the beginning writing of students with and without disabilities.

Educational implications for beginning writers

Currently, measures for monitoring the progress of typically developing beginning writers tend to be either literacy profiles that measure multiple constructs of both reading and writing or curriculum-based measures that focus on quantitative measures of writing fluency (e.g., TWW). One advantage of the DWS is that it provides teachers with specific information about students' conceptual understandings of written language that cannot be determined by measures of TWW or other quantitative measures of writing progress. This is consistent with Purposes 1 (to *distinguish variations in beginning writing skills*) and 5 (to *make it possible to celebrate students' positive change as writers*) summarized previously.

In summary, as the previously mentioned examples illustrate, the DWS can be used to distinguish variations in the writing quality of beginning writers that are instructionally relevant. Results of pilot research on the ease, efficiency, and utility of the DWS, based on the small survey, are also promising, suggesting that teachers may find it relatively easy to use the DWS for periodic probes of student progress. Displaying DWS outcomes in tables for students, educational teams, and families provides a way to celebrate student writers. An anecdotal example of this naturally occurred when a student excitedly shared his outcome book with his parents as part of a "Meet the Author" event in an Enriched Writers' Workshop taught by the first author (see Sturm, 2012). The student's outcome book contained tables of his writing progress and his writing for the year. The student looked at each table and showed them to his parents. During this sharing moment clinicians were able to use the DWS to talk easily with parents about the changes their child had made and to show how the DWS levels were reflected in the writing samples.

Limitations and future directions

A limitation of the research reported in this article was sample size, particularly for the initial survey of only four participants. Preliminary evidence regarding teacher and SLP perceptions about ease, efficiency, and utility of the DWS was reported for only four participants. These participants work closely with the first author; therefore, an additional limitation is response bias based on social desirability. Because the participant responses could not be kept completely anonymous, their responses may have been influenced by what they thought the researchers wanted to hear. Future research could examine perceptions of the DWS with a larger group of general education teachers, special education teachers, SLPs, and university students.

Table 6. Results of pilot survey on the ease, efficiency, and utility of the DWS

Survey Item	Participant Responses (N = 4)	Participant Comments
1. The DWS is easy to use	Very much agree = 3 Agree = 1	Clear instructions and very detailed, helpful examples were provided. The description of each benchmark (level) is clear. The DWS is easy and straightforward. Best of all, it allows me to plot my students no matter what tools and supports my students use to write.
2. The DWS is fast.	Very much agree = 4	Rating the samples goes fast. So far my samples are Levels 1–5. Most writing samples take less than 2 min to assign a writing level.
3. The DWS is useful.	Very much agree = 4	The DWS is useful in many ways—it helps to see, "at a glance," the skills of a student writer and to consider the next level that the student is working toward. The DWS is very helpful when working with writing across the curriculum—all instruction is centered on developing writing with the scale benchmarks (levels) as goals. The DWS is really useful as a point of reference when team teaching or when working with educators teaching other subjects. It is extremely useful to have a scale that reflects writing development across learners.
4. The DWS is easy to learn.	Very much agree = 4	The examples are so helpful when learning the tool! Learning the DWS required a bit of training. The examples are really helpful. Very easy and straightforward.
5. The DWS will help me with my writing instruction.	Very much agree = 4	I've written IEP goals with the DWS as the measurement tool (progress indicator). It has helped center instruction for our educational team. Very clear for all members of the team and I am hoping it will be easily transferred to new teams as students transition to new classrooms. It helps me to know where my students are in their writing abilities and where I should help them go.
6. The DWS will make my job easier.	Very much agree = 3 Agree = 1	The DWS is helpful in explaining writing development to parents and IEP teams It is very useful in addressing student writing goals.
7. I will use the DWS again.	Very much agree = 4	I will continue to use the DWS to inform student written language assessment.
8. How often would you use the DWS to measure your students' writing?	Quarterly = 1 Monthly = 1 Biweekly and monthly = 1 Monthly and quarterly = 1	I use it quarterly at trimester, progress reporting times. I could increase the frequency. I'd use it at least monthly, perhaps also during team meetings, or parent conferences during the year. Biweekly and monthly—it depends, based on the level of writer. I use it more often as a student's writing progresses. I work with preschool-age children, so, depending on the child, I may use it monthly. I would also use it quarterly, so I can look at several samples at one time.

Note. DWS = Developmental Writing Scale.

Although evidence of reliability using Cohen's κ was reported for analysis of 285 samples, future research should examine scoring reliability using a greater number of writing samples from both developmentally disabled and typically developing students. It should also examine scoring reliability for teachers who have received minimal training. Such research should also examine the length of training time and type of training needed to use the DWS with greater reliability. In addition, reliability studies should address reliability issues related to scoring words with questionable intelligibility. For example, the sample "IKTO the BC" has been scored as a 6 and a 3 by two different individuals (the authors would score it as a 6).

Currently, research is being conducted to validate the DWS with typically developing kindergarten and first-grade writers (Cali, manuscript in preparation). Future research should be conducted to validate the levels of the DWS for beginning writers with disabilities as they change longitudinally. Naturalistic samples, from both groups of students, could be examined to further validate each level of the scale and to understand the range of writing levels of students in beginning general education classrooms. To further validate the DWS, teachers and SLP could be asked to rank order sets of samples to confirm the overall developmental sequence of the tool.

CONCLUSIONS

This article began with three writing samples that all scored a "1" on a state assessment, highlighting the multiple limitations of the scoring system. In contrast, the DWS, as a formative and summative assessment, provides refined developmental levels that allow teachers to measure small amounts of progress across grade levels with disabled and nondisabled writers. This was illustrated in Table 4, in which the same samples were assigned DWS levels of 4, 11, and 12. As a formative assessment, the DWS provides teachers with information about student's writing ability within a natural context. Teachers can use this information to plan instruction that supports the student in moving to the next level. The DWS is grounded in research on the development of beginning writers and is focused on growth in linguistic quality. Therefore, the face validity for this measure is enhanced by making it intuitive and easy for educators to use and implement. As the Common Core State Standards (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010) are implemented, there is an expectation that writing instruction will be differentiated for learners while also expecting high levels of teacher accountability for student progress. As a summative measure, the sensitivity of the DWS measure will support educators in achieving that goal. The DWS is not only for the educator; it also allows students to see positive growth in their own writing and can be used to celebrate their writing gains and foster intrinsic motivation to write. The preliminary data presented here suggest that the DWS may be a powerful comprehensive measure of qualitative change in beginning writers and will be useful to enhance both assessment and instruction.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA/APA/NCME). (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bedrosian, J., Lasker, J., Speidel, K., & Politsch, A. (2003). Enhancing literacy in individuals with autism and severe communication impairments. *Topics in Language Disorders*, *23*(4), 305–324.

Bereiter, C. (1980). Development in writing. InW. Gregg, & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 73–93). Hillsdale, NJ: Lawrence Erlbaum Associates.

Berninger, V., & Gans, B. (1986). Language profiles in nonspeaking individuals of normal intelligence with

severe cerebral palsy. *Augmentative and Alternative Communication*, *2*, 45–50.

Broun, L. (2009). Taking the pencil out of the process. *Teaching Exceptional Children*, *42*, 14–21.

Cali, K. S. (manuscript in preparation). *The validity and reliability of the Developmental Writing Scale for beginning writers*.

Clay, M. (1975). *What did I write?* Exeter, NH: Heinemann.

Clay, M. (2006). *An observation survey of early literacy achievement: Revised second edition*. Portsmouth, NH: Heinemann.

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.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Dyson, A. H. (1985). Individual differences in emerging writing. InM. Farr (Ed.), *Advances in writing research: Children's early writing development* (pp. 59–126). Norwood, NJ: Ablex.

Dyson, A. H. (1986). Transitions and tensions: Interrelationships between the drawing, talking, and dictating of young children. *Research in the Teaching of English*, *20*, 379–409.

Ehren, B. J. (1994). New directions for meeting the needs of adolescents with language learning disabilities. InG. P. Wallach, & K. G. Butler (Eds.), *Language learning disabilities in school age children and adolescents* (pp. 393–417). New York: Merrill.

Espin, C. A., Weissenburger, J. W., & Benson, B. J. (2004). Assessing the writing performance of students in special education. *Exceptionality*, *12*(1), 55–67.

Ferreiro, E., & Teberosky, A. (1982). *Literacy before schooling*. Portsmouth, NH: Heinemann.

Fitzgerald, J., & Spiegel, D. L. (1986). Textual cohesion and coherence in children's writing. *Research in the Teaching of English*, *20*, 263–280.

Fitzgerald, J., & Spiegel, D. L. (1990). Textual cohesion and coherence in children's writing. Revisited. *Research in the Teaching of English*, *24*, 48–66.

Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Naquin, G. A., & Slider, N. J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternative measures for curriculum-based measurement in writing. *School Psychology Review*, *31*(4), 477–498.

Graham, S., & Harris, K. (2005). *Writing better: Effective strategies for teaching students with learning difficulties*. Baltimore, MD: Brookes Publishing Co., Inc.

Graham, S., MacArthur, C., Schwartz, S., & Page-Voth, V. (1992). Improving the compositions of students with learning disabilities using a strategy involving product and process goal setting. *Exceptional Children*, *58*, 322–334.

Halliday, M., & Hasan, R. (1976). *Cohesion in English*. London: Longman.

Harris, D. (1982). Communicative interaction processes involving nonvocal physically handicapped children. *Topics in Language Disorders*, *22*, 21–37.

Hayes, J. R., & Hatch, J. (1999). Issues in measuring reliability: Correlation versus percentage of agreement. *Written Communication*, *16*, 354–367.

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Washington, DC: Chief Council of State School Officers.

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*(1), 27–44.

Joseph, L. M., & Konrad, M. (2009). Teaching students with intellectual or developmental disabilities to write: a review of the literature. *Research in Developmental Disabilities*, *30*(1), 1–19.

Kaderavek, J., & Rabidoux, P. (2004). Interactive to independent literacy: A model for designing literacy goals for children with atypical communication. *Reading and Writing Quarterly*, *20*(3), 237–260.

Koke, S., & Neilson, J. (1987). *The effect of auditory feedback on the spelling of nonspeaking physically disabled individuals*. *Unpublished master's thesis*, University of Toronto, Toronto, Ontario, Canada.

Koppenhaver, D. A., & Erickson, K. A. (2003). Natural emergent literacy supports for preschoolers with autism and severe communication impairments. *Topics in Language Disorders*, *23*(4), 283–292.

Koppenhaver, D. A., & Yoder, D. E. (1993). Classroom literacy instruction for children with severe speech and physical impairments (SSPI): What is and what might be. *Topics in Language Disorders*, *13*(2), 1–15.

Kress, G. (1982/1994). *Learning to write*. London: Routledge.

Langer, J. A. (1986). *Children reading and writing: Structures and strategies*. Norwood, NJ: Ablex.

MacArthur, C., Schwartz, S. S., & Graham, S. (1991). A model for writing instruction into a process approach to writing. *Learning Disabilities Practice*, *6*, 230–236.

Massachusetts Department of Education (2012). *Massachusetts comprehensive assessment system: Scoring guides for MCAS English Language Arts Composition*. Retrieved August 28, 2012, from http://www.doe.mass.edu/mcas/student/elacomp_scoreguide.html

McCabe, A., & Bliss, L. S. (2003). *Patterns of narrative discourse: A multicultural, life span approach*. Boston, MA: Allyn & Bacon.

McMaster, K., Du, X., Yeo, S., Deno, S.L., Parker, D., & Ellis, T. (2011). Curriculum-based measures of beginning writing: Technical features of the slope. *Exceptional Children*, 77(2), 185–206.

Michigan Department of Education. (2000). *Michigan literacy progress profile*. Retrieved June 2, 2012, from http://www.misd.net/MLPP/assessments/writing/Writing-A.pdf

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: Author. Retrieved September 1, 2012, from, http://www.corestandards.org/the-standards.

Nelson, N.W. (1992). Performance is the prize: Language competence and performance among AAC users. *Augmentative and Alternative Communication*, *8*, 3–18.

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*, 578–593.

Newkirk, T. (1987). The non-narrative writing of young children. *Research in the Teaching of English*, *21*(2), 121–144.

Newkirk, T. (1989). *More than stories: The range of children's writing*. Portsmouth, NH: Heinemann.

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labeling. *Journal of Child Language*, *5*(1), 1–15.

North Carolina Department of Public Instruction. (2009). *North Carolina K-2 literacy assessment. English language arts resources*. Retrieved June 2, 2012, from http://www.ncpublicschools.org/docs/curriculum/languagearts/elementary/k2literacy/2009k2-literacy.pdf

Northwest Regional Educational Laboratory. (2010). *Beginning Writer's Continuum. 6+1 Trait® Rubrics (aka Scoring Guides)*. Retrieved June 2, 2012, from http://educationnorthwest.org/webfm_send/772

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics*. New York: Plenum.

Prelock, P. A. (2006). *Autism spectrum disorders: Issues in assessment and intervention*. Austin, TX: ProEd.

Salahu-Din, D., Persky, H., & Miller, J. (2008). *The Nation's Report Card: Writing 2007 (NCES 2008–468)*. Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved June 2, 2012, from http://nces.ed.gov/nationsreportcard/pdf/main2007/2008468.pdf

Scott, C. M. (1989). Problem writers: Nature, assessment, and intervention. InA. G. Kamhi, & H. W. Catts (Eds.), *Reading disabilities: A developmental language perspective* (pp. 303–344). Boston, MA: Allyn & Bacon.

Scott, C. M. (2012). Learning to write. InA. G. Kamhi, & H. W. Catts (Eds.), *Language and reading disabilities* (3rd ed., pp. 248–268). Boston, MA: Pearson.

Smith, A., Thurston, S., Light, J., Parnes, P., & O'Keefe, B. (1989). The form and use of written communication produced by physically disabled individuals using microcomputers. *Augmentative and Alternative Communication*, *5*, 115–124.

Sturm, J. M. (2012). An enriched writers' workshop for beginning writers with developmental disabilities. *Topics in Language Disorders*, *32*(4), 1–35.

Sturm, J. M., & Clendon, S. A. (2004). AAC, language, and literacy: Fostering the relationship. *Topics in Language Disorders*, *24*(1), 76–91.

Sturm, J. M., Erickson, K. A., & Yoder, D. E. (2003). State of the science: Enhancing literacy participation through AAC technologies. *Journal of Assistive Technology*, *14*, 45–54.

Sturm, J. M., Knack, L., & Hall, J. (2011, November). *Writing instruction in primary classrooms: Implications for students with disabilities*. Poster session presented at the American Speech-Language-Hearing Convention, San Diego, CA.

Sturm, J. M., Nelson, N. W., Staskowski, M., & Cali, K. (2010, November). *Outcome measures for beginning writers with disabilities*. Miniseminar presented at the American Speech-Language-Hearing Convention, Philadelphia, PA.

Sulzby, E., Barnhart, J., & Hieshima, J. (1989). *Forms of writing and re-rereading from writing: A preliminary report (Technical Report No. 20)*. Berkeley, CA: National Center for the Study of Writing and Literacy. Retrieved November 15, 2010, from http://www.nwp.org/cs/public/print/resource/606

Sulzby, E., & Teale, W. (1991). Emergent literacy. InR. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. II, pp. 727–757). New York: Longman.

Tennessee Department of Education. (2010). *Tennessee comprehensive assessment program writing assessments*. Retrieved January 2, 2011, from http://www.tn.gov/education/assessment/writing.shtml

Tennessee Department of Education. (2012). *Writing assessment scoring information*. Retrieved August 28, 2012, from http://www.tn.gov/education/assessment/writing_scoring.shtml

Tolchinsky, L. (2006). The emergence of writing. InS. Graham, C. MacArthur, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 83–95). New York: Guilford Press.

Udwin, O., & Yule, W. (1990). Augmentative communication systems taught to cerebral-palsied children—A longitudinal study. 1. The acquisition of signs and symbols, and syntactic aspects of their use over time. *British Journal of Disorders of Communication*, *25*, 295–309.

van Balkom, H., & Welle Donker-Grimbrère, M. (1996). A psycholinguistic approach to graphic language use. InS. von Tetzchner (Ed.), *Augmentative and Alternative Communication: European Perspectives* (pp. 153–170). London: Whurr.

Vandervelden, M., & Siegel, L. (1999). Phonological processing and literacy in AAC users and children with motor speech impairments. *Augmentative and Alternative Communication*, *15*, 191–211.