Differentiated Writing Interventions for High-Achieving Urban African American Elementary Students

Jennifer L. Geisler
Walnut Creek Elementary School

Terri Hessler
The Ohio State University, Newark

Ralph Gardner, III
Temple S. Lovelace
The Ohio State University

Introduction

A free and effective education that prepares children to become productive members of society is considered essential to a democracy (Kauffman, Conroy, Gardner, & Oswald, 2008). To accomplish this goal, schools must meet the challenge of effectively educating all children. An irrefutable fact is that African American children, particularly those in urban schools, are not faring well. African American children are at higher risk than other children for receiving a special education disability label, and they are less likely to be placed in gifted education (Artiles, Trent, & Palmer, 2004; Hosp & Reschly, 2004; Losen & Orfield, 2002).

African American students are overrepresented in special education and underrepresented in gifted education. This is in large part due to students' poor performance in core academic areas such as reading, math, and writing. Differentiating instruction in early grades could assist in closing the writing performance gap between African American and majority students, with the intended outcome of increasing the likelihood of students achieving their academic potential. Research-based strategies that improve fluency and vocabulary—potentially improving writing quality—include students counting the total number of words they have written and self-monitoring their progress. We examined the effects of self-counting (which included viewing graphs of the word counts students had written) and student use of synonym lists (which provided alternative words for students to use in their writing) on the length and quality of writing of 5 high-achieving urban African American first graders whose instruction took place in a first/second-grade split classroom. All 5 students demonstrated improved writing outcomes. The results of this study support the use of differentiated interventions for high-achieving students in order to better increase the likelihood that they will achieve in written expression at a level commensurate with their abilities.

summary

Regardless of whether a student has a disability, is typically developing, or is high achieving, one goal of education is to assist students, including those from culturally diverse families, to perform at their highest potential. According to the U.S. Department of Education (USDE) in its report *National Excellence: A Case for Developing America's Talent* (1993), students with disabilities or who are academically at risk are provided a multitude of interventions and resources to help them progress, while high-achieving students are often neglected because of their early mastery of skills. Despite the recognition more than 2 decades ago that far too often high-achieving students fail to achieve their full potential (National Commission on Excellence in Education, 1983), underachievement is still a common occurrence among our most promising students (Delisle & Galbraith, 2002).

The gap between ability and achievement is especially apparent in African American populations in urban school districts (Ford & Harris, 1993). *National Excellence* (USDE, 1993) exposed the fact that economically disadvantaged African Americans are offered the fewest opportunities for academic advancement and are the most at risk for underachievement. African American children from impoverished families often begin school already at an academic disadvantage, due to a limited vocabulary (Hart & Risley, 1995). A primary reason children fall further and further behind in school is poor literacy skills (Stanovich & Siegel, 1994).

There is a strong relationship between vocabulary and a person's ability to read and write. In fact, the National Reading Panel (National Institute of Child Health and Human Development [NICHD], 2000) acknowledged the importance of vocabulary development as one of the five essential skills in teaching children to read. Blachowicz, Fisher, and Watts-Taffe (2005) also made evident the connection between vocabulary and school performance. Poverty, sadly a case for many minority students, is a clear detriment to vocabulary development and can negatively affect the performance of poor African Americans (Blachowicz et al., 2005). Fortunately, disadvantaged students can make gains in reading comprehension and written expression with vocabulary instruction that is appropriate to their age and ability

(Blachowicz et al., 2005). Castellano, Faivus, and White (2003) provided evidence that early, rather than late, identification and intervention is more effective for the success of high-achieving students from culturally diverse groups, including African American children.

Differentiation

The success of high-achieving children, like their typical peers, lies in the hands of the general education classroom teacher, despite the fact that many teachers receive little to no training regarding the needs of high-achieving students (Ohio Gifted Task Force [OGTF], 2002). Differentiated instruction, a recent trend in education (Tomlinson, 1999), is helpful—perhaps critical—for educators who teach students with a range of abilities and needs. Differentiated instruction involves the adjustment of content, instruction, and assessment to meet the needs of each individual in a heterogeneous classroom (George, 2005). Teachers must plan and implement instruction; they also must assess instructional effectiveness. The assessment of instructional effectiveness is essential in maximizing student achievement. Assessment provides the opportunity for the teacher to identify and change ineffective instruction, as well as to verify effective instruction. For teachers to assess on a regular basis, assessment tools should be user-friendly and provide useful information.

Curriculum-Based Measurement

Curriculum-based measurement (CBM) is commonly used for classroom assessments because it can be used frequently, and it provides educators with a reliable, valid, and efficient assessment of student competence in curriculum-related skills (e.g., reading, solving math problems, writing; Fuchs, Fuchs, & Compton, 2004). Examples of CBM include the number of words read, number of math problems solved correctly, and number of words written. CBM is intended to assist teachers in improving student academic performance by facilitating a database of repeated

measures across a period of time. The teacher uses the data to evaluate the extent of the student's learning and the effectiveness of the instruction by looking at the rate of change displayed in the repeated measurements (Deno, 1992). There are three benefits of using CBM: (a) it helps teachers and students focus on the concepts of basic skills, (b) it operates as an indicator of growth in basic skills, and (c) it describes incremental student growth that infrequent achievement tests fail to show (Deno, 1992). The latter benefit is particularly useful in the classroom. Teachers cannot afford to wait until an end-of-the-year test to know which students are not progressing. In fact, unless the teacher has easy access to specific results from the previous year's test, these standardized tests give a limited amount of information (e.g., Is the student performing at a level comparable to peers nationwide?); they do not provide the kind of useful, easy-to-implement progress monitoring of student performance that the frequent, short cycle curriculum-based measures do. Infrequent progress monitoring in the form of standardized tests may be contributing to the gap between culturally diverse students and those of the dominant culture. Frequent progress monitoring with CBM is a tool teachers can use to maximize instructional time. If a teacher monitors a child's reading progress, for example, and notices that the child is still reading 23 words per minute with eight errors after 4 weeks of instruction, the teacher can presume that the current instruction is not effective in improving the child's reading performance. The teacher who is armed with that knowledge early in the school year can make instructional changes to positively impact a child's reading performance instead of having to wait until an end-of-the year assessment. In other words, for an assessment tool to be useful, the teacher must correctly interpret the data and use the data to inform future academic instruction.

Written Expression

Written expression is becoming recognized as the most neglected of the basic academic skills (National Committee on Writing, 2003), and there is no consensus on the best method

for either teaching composition or evaluating it. Many educators prefer an informal assessment method. Minner, Prater, Sullavan, and Gwaltney (1989) suggested analyzing a writing sample for fluency and vocabulary, along with sentence types and ideation. Fluency refers to the ability of a student to write quickly and correctly; vocabulary involves the use of different but appropriate words. Improvement of these skills is believed to indicate advancement by the writer, as a fluent writer with good vocabulary can quickly produce relatively error-free, interesting text.

Vocabulary. Word knowledge is fundamental for writing. Blachowicz et al. (2005) provided a framework for vocabulary instruction. They emphasized the teacher's responsibility to use explicit instruction and use of new words and then differentiate the instruction based on the needs of the learners. One example of specific vocabulary instruction to expand writing is teaching synonym awareness and practice, such as categorizing words that go together, constructing personal webs for writing, and using a thesaurus.

Length of Writing. Some research indicates that the number of words written is a quality indicator of written expression. Bui, Schumaker, and Deshler (2006) reported that non-LD students in an experimental group had a 10% increase in length of writing on the posttest in addition to all quality measures improving as well; conversely, the control group wrote less during posttest while scoring lower on some of the quality measures.

De La Paz (1999) reported that middle school students across all ranges of ability (i.e., low achieving, average achieving, high achieving, and those with learning disabilities) doubled the amount of text written after being taught self-regulated strategy instruction. These students wrote more text, but importantly, "they wrote less nonfunctional text" (p. 103), and their overall quality of writing improved as well.

Tindal and Parker (1991) conducted a study of third-, fourth-, and fifth-grade students of varied abilities to assess writing methods. They reported their quantitative measures (number of total words written, correctly spelled words, and correct

word sequences) highly correlated with their qualitative measures (rubric measures of story-idea and organization-cohesion), while mechanics-conventions showed a lower correlation.

In a study of fourth-grade students, Kasper-Ferguson and Moxley (2002) examined self-counting and graphing of total words written from timed freewriting samples over the course of a school year. Improvement in the quality of writing was associated with increases in number of words written.

Holistic Scoring. Writing is a combination of many skills, few of which can be measured objectively. In essence, a writing piece is more than just a sum of its parts. An essay may be riddled with capitalization, spelling, and punctuation errors but overall still communicate a coherent, interesting message. That is why it is often recommended that written expression be evaluated on its impression as a whole (White, 1984). Classroom teachers and researchers alike recognize that rubric scores are useful. Graham, Harris, and Mason (2005); Gansle, Noell, VanDerHeyden, Naquin, and Slider (2002); Tindal and Hasbrouck (1991); and Tindal and Parker (1991) used rubric scores as evaluative measures in their studies.

Self-Monitoring

Self-monitoring can occur across various skills depending on the needs of the students. In fact, one of the goals of effective writing instruction is for students to self-monitor their ability to communicate effectively in standard English. Providing high-achieving students who self-monitor with progress feedback is an effective learning component to be modeled for their future generalization of concepts (Rogers, 1986). Schunk and Swartz (1993) taught fourth-grade gifted students a strategy goal and gave them feedback on their writing. The results demonstrated that students receiving a strategy goal plus progress feedback outperformed the other groups in all writing posttests. The researchers concluded that gifted students in particular are strategic by nature and more likely to understand and generalize a strategy for future use, and they suggested that an interven-

tion that includes feedback should be incorporated into writing instruction to produce more capable writers.

Another type of self-monitoring used to assist writers is self-counting, an assessment/instructional strategy intended to increase word expression. Moxley and Lutz (1995) used self-recorded word counts of freewriting as one strategy to provide more writing experiences with less frustration. In their study of first graders, the researchers found that the students improved from a class total of 140 words in the first session to 451 words in the last session. The median increase for number of words written was between 5 and 18 words.

Purpose

The purpose of this study was to determine the effects of self-counting and a synonym list on the number of total words written and the number of different words written by high-achieving urban African American first graders, as well as on the quality of their written expression. Quality was determined by scores assigned to writing samples based on a district rubric organized into three areas, including content, organization, and conventions, and in each area the scores can range from 1 (*minimal*) to 4 (*effective*). Higher rubric scores represent better quality writing.

Method

Participants and Setting

The participants were 5 African American first graders (4 females and 1 male). They were selected for this study because they were the only students in their class who scored at the second-grade level at the beginning of the school year as measured by school district's first-grade benchmark assessment. In this school district, students who perform one full level or more above grade level in reading and comprehension on the district

Table 1

Student Demographic Information at Entry Into First Grade

Student	Gender	Age	Benchmark Grade Level[a]	Reading?	Strong Writer?[b]
1	F	7	2.3	yes	no
2	M	7	3.2	yes	yes
3	F	6	2.3	yes	yes
4	F	7	2.4	yes	yes
5	F	6	2.3	yes	no

[a]Students were given the benchmark assessment upon entry to first grade so an at-grade-level score would be 1.1.

[b]Based on teacher assessment of writing ability.

benchmark are designated as high-achieving and receive their language arts instruction in a classroom appropriate to their benchmark grade level (e.g., first graders who score at a second-grade level upon entry into first grade receive language arts in a second-grade classroom). Unfortunately, a scheduling conflict prevented these 5 students from attending the second-grade language arts class, so the principal directed the teacher-researcher to provide 20 minutes of daily language arts enrichment. See Table 1 for each student's demographic information.

Student 1. The teacher-researcher described Student 1 as overly concerned about spelling words correctly, causing her to write slowly; she often was seen carefully erasing and rewriting words. She was not a strong writer (classroom observations, September 2005–May 2006).

Student 2. The teacher-researcher described Student 2, the only male in the study, as an advanced speller who did not particularly like to write, although he was more skilled than other students in

class and in the study (classroom observations, September 2005–May 2006).

Student 3. The teacher-researcher described Student 3 as a child who was an advanced speller, as well as thoughtful about writing topics. She could write complete sentences before the start of the study (classroom observations, September 2005–May 2006).

Student 4. The teacher-researcher described Student 4 as meticulous, writing slowly to avoid making mistakes, although she was able to write complete sentences at the start of the study. She was also easily distracted by other students and often complained of illness (classroom observations, September 2005–May 2006).

Student 5. The teacher-researcher described Student 5 as a slow writer concerned about neatness who frequently asked for help with spelling. This student was not able to write in complete sentences at the start of the study (classroom observations, September 2005–May 2006).

The study was conducted in a first/second-grade split classroom composed of 21 African American students in an urban elementary school in a Midwestern metropolitan school district in a K–5 building with an approximate enrollment of 350 students. Of the other 15 students in the class, none met the district's criteria as high achievers at the beginning of the year. All students in the class completed the writing activities, but only the participants were introduced to the interventions of self-counting and use of a synonym list (as this constituted their enrichment).

Data Collectors

The primary data collector was the teacher-researcher. A doctoral student in special education was the secondary data collector, trained by the teacher-researcher to collect word counts for the number of total words and the number of different words.

The doctoral student met training completion criterion after two training sessions, scoring 100% agreement with the teacher-researcher in each session.

For rubric scoring of the generalization probe essays, two evaluators were selected by the researchers for their familiarity with the writing of urban elementary students. Evaluator 1 was a second-grade teacher with 8 years of experience. Evaluator 2 was a reading facilitator in the building who also prepared students for standardized writing tests and provided professional development on writing standards to the staff. Evaluator 2 also had 8 years of experience. To prevent influences in scoring, they were not informed of the study's purpose until after its completion.

Independent Variables

Self-Counting. Self-counting was the first intervention and involved two elements: (a) counting the words and (b) viewing graphs of the number of words. Students met individually with the teacher-researcher after each writing session to count the words in their writing sample. Students and data collectors only counted words written during the first 3 minutes of writing, a common practice in writing research at the elementary level (e.g., Gansle et al., 2002; Jewell & Malecki, 2005; Tindal & Parker, 1991). They recorded their counts for the number of total words and the number of different words at the bottom of the writing samples.

The number of different words was graphed after each writing session by the teacher-researcher and shown privately to each student before the next writing session. The graphs were only shown to students during self-counting and self-counting plus synonym list conditions.

Self-Counting Plus Synonym List. The second intervention involved a synonym list developed by the teacher-researcher (see Figure 1). The list was designed to provide students with alternate word choices to assist them in writing a greater number of different words. Synonyms for the list were chosen based on

Synonym List

big	going/went (to somewhere)	going (to do something)
large	visited	ready
huge	took a trip	will
enormous	drove	
giant	walked	
gigantic		
good/great	happy	knew/know
wonderful	excited	realize(d)
terrific	joyous	aware of
like	sad	said/say
enjoy	disappointed	shouted
love	unhappy	whispered
favorite	upset	asked
	miserable	cried
small	very	then
little	extremely	next,
tiny	really	after that,
miniature		eventually,

Watch out for these most frequently repeated words
and
went
would
then
so
very

Figure 1. The synonym list, created by the teacher-researcher, used in the second intervention phase.

the most frequently repeated words from their baseline and self-counting writing samples. In addition to using the synonym list, students still continued to self-count and view their graphs of the number of different words. The synonym list was given to students before the writing session began and collected afterwards.

Dependent Variables

Number of Different Words. The number of different words is a measure of vocabulary use. Both students and data collectors counted the words written during the first 3 minutes of each writing sample. Students counted directly after the writing period as part of the intervention. Students were instructed to place a checkmark above each word. They also were directed to circle the second occurrence (or third or fourth, etc.) of a word and then skip over circled words (i.e., not place a checkmark above circled—or repeated—words). Students then counted the number of checkmarks, which was the number of different words. In short, if a word was repeated, it was counted only the first time it appeared in that session's writing sample.

Number of Total Words. Students and data collectors counted the number of total words in each writing sample. The same guidelines for counting the number of different words were used for counting the number of total words. For the total count of words, though, every word was counted, whether it was repeated or not. A word was defined as a letter or series of letters separated by a space next to another letter or series of letters, including misspelled words.

Rubric Scores on Generalization Probes. Because the researchers wanted a more traditional and comprehensive measure of quality (as opposed to one related only to vocabulary, i.e., number of different words) and because it is important that any positive outcomes from the intervention manifest in (i.e., generalize to) more authentic writing situations, generalization probes were implemented. Students were provided writing prompts

three times during the course of the study. The conditions under which they wrote in response to these probes were different than their weekly writing sessions in that (a) writing was untimed, although limited by the length of the class period, (b) students went through the entire writing process, (c) the self-counting and synonym list materials were not provided, although they were neither encouraged to nor discouraged from self-counting, and (d) students were given the following topics and could choose which to write about:

- a dream they had recently,
- a trip they took or would like to take,
- a super power they wished they had,
- a city they had visited,
- a memorable spring break or summer vacation,
- a pet they had or would like to have, or
- a concert they had attended or would like to attend.

Evaluators used the district K–2 writing rubric to score the generalization probe essays that had been stripped of identifying information. The district rubric is a matrix that helps scorers rate the quality of writing on a scale of 1–4 (1 = *Minimal*; 2 = *Partial*; 3 = *Adequate*; 4 = *Effective*) in the areas of content, organization, and conventions. The highest rubric score possible was 12 and the minimum was 3. Example criteria from the content area of the rubric that would be rated *effective* (and therefore, 4) include developing and supporting a main idea and writing in complete sentences. Example criteria from the organization area of the rubric that would be rated *effective* include developing a beginning, middle, and end and using a variety of vocabulary. Example criteria from the conventions area of the rubric that would be rated *effective* include legible handwriting and correct punctuation, capitalization, and spelling.

Multiple-Baseline Experimental Design

The researchers chose a single-case research design to examine these interventions. There are two major characteristics of

single-case research: (a) there is continuous assessment of the target behavior over time, and (b) student outcomes are examined in different conditions, and the conditions are then replicated across subjects (Kazdin, 1982). Multiple-baseline designs are one approach for single-case research.

In a multiple baseline across subjects design, the researcher introduces the intervention to different persons at different times. The significance of this is that if a behavior changes only after the intervention is presented, and this behavior change is seen successively in each subject's data, the effects can more likely be credited to the intervention itself as opposed to other variables. Multiple-baseline designs do not require the intervention to be withdrawn. Instead, each subject's own data are compared between intervention and nonintervention behaviors, resulting in each subject acting as his or her own control (Kazdin, 1982). An added benefit of this design, and all single-case designs, is the immediacy of the data. Instead of waiting until postintervention to take measures on the behavior, single-case research prescribes continuous data collection and visual monitoring of that data displayed graphically, allowing for immediate instructional decision-making. Students, therefore, do not linger in an intervention that is not working for them, making the graphic display of single-case research combined with differentiated instruction responsive to the needs of students. This is especially important for high-achieving students who often languish too long in situations in which their abilities are not challenged.

Procedure

Mini-Lesson

Before the onset of baseline, the teacher-researcher provided the entire class with a mini-lesson to introduce the concept of using different words to enhance the quality of its writing. The teacher-researcher told a simple story repeating the word "sad" many times. The students agreed that this was not a very excit-

ing story. The teacher-researcher asked students to suggest other words that could be used instead of "sad." The students provided a short list of synonyms such as "unhappy" and "miserable." The teacher-researcher retold the same simple story using some of the alternative word choices for "sad." The students agreed that this story sounded better. The teacher-researcher repeated this process with another word. The students were then given examples of simple sentences (i.e., "It was fun.") and asked to change words in the sentence to make it more interesting. The teacher-researcher told students that using different words in their writing would help make it more interesting and exciting to read. The first generalization probe was written after this mini-lesson.

Baseline

The teacher-researcher used a script for all conditions of the study. In baseline, all of the students in the class were told to write as many words and as many different words as possible and were given a writing prompt or story starter (e.g., "My lottery ticket had the winning number!"). Students were given a 1-minute think period for brainstorming, followed by the signal to begin writing. After 3 minutes, they were instructed to mark their papers with a line after the last word they wrote and continue writing for a total of 8 minutes. The teacher periodically provided general praise and/or prompts, such as, "Keep writing," "You can go to the next page if you need to," and "Great job! You are writing a lot!" After 8 minutes, students were told to stop writing.

Intervention Phases

Self-Counting. The self-counting strategy was introduced to the participants as a mini-lesson before the initial intervention session. During these procedures the other student participants were engaged in literacy activities at various stations throughout the room. The teacher-researcher explained to each subject the process of self-counting and reiterated the goal to write more

total words and a greater number of different words. Students then followed the same procedures for writing as in baseline.

After the first writing session, students were trained to self-count and record the results. They were given blank transparency paper to attach with paper clips over their writing journals and a black overhead projector marker to mark the words as they counted. These tools allowed them to self-count immediately after the writing session while maintaining a "clean" copy of the writing sample for the data collectors. The number of total words written and the number of different words written from the beginning to the 3-minute mark were counted and noted at the bottom/end of the paper.

The teacher-researcher created a graph for each student's total words written and number of different words written. Before the following writing sessions, participants were shown their own graphs for each writing sample produced in the last session.

Self-Counting Plus Synonym List. The use of self-counting plus synonym list was implemented when self-counting data stabilized (see Figure 2). The line between baseline and the first intervention indicates when data stabilized for each student. The teacher-researcher provided each student with the synonym list to use while writing and told him or her, "You may use this word list to help you write different words you want to use." This intervention phase used the same writing procedure as the self-counting phase. Students continued to self-count and view the graphs of the number of total words and number of different words that they wrote.

Generalization Probes. The students wrote essays in response to three generalization probes throughout the study: the first after the introductory mini-lesson but before baseline sessions began, the second after finishing the self-counting intervention but before beginning the self-counting plus synonym list intervention, and the third after the final writing session. Students did not self-count or use a synonym list during the writing of the generalization probe essays. Before baseline began, students were asked to produce a narrative essay on a topic of their choice and
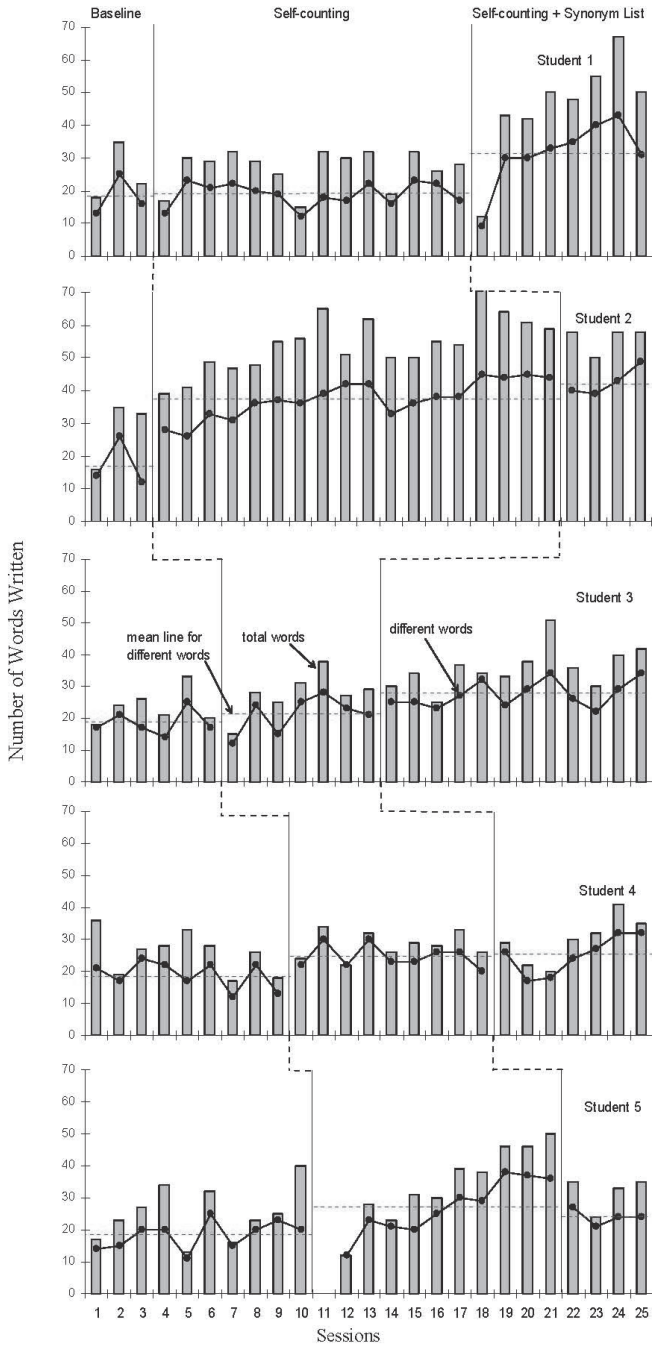
Figure 2. Graphs showing number of total words written (bars) and number of different words written (lines) for Students 1–5 in 3-minute writing samples.

progress through the writing process. It was an untimed probe during which students planned, drafted, and edited their pieces. The generalization probe writing occurred over the course of 2 to 3 days. Students were neither encouraged nor discouraged from self-counting (i.e., it was not mentioned as a part of the directions for the probe) and were not provided the synonym list. At the conclusion of the self-counting phase, but before synonym lists were added, another generalization probe was implemented using the same procedure. A third generalization probe was completed at the conclusion of the last intervention using the same procedure.

The students wrote in response to each probe over the course of 3 days: They brainstormed and organized ideas for a beginning, middle, and end on Day 1; they combined those parts into an initial draft on Day 2; and they edited a final copy on Day 3. The students were allowed to use any of or the entire 20 minute enrichment time on each of the 3 days.

Results

To determine the effects of the two interventions on the written expression skills of 5 high-achieving first-grade African American students, the number of total words and different words written was examined. The results for both of these dependent variables are displayed in Figure 2. The data lines in the graphs represent the number of different words written and the bars represent the number of total words written. A horizontal dotted line also is provided to show the mean number of different words written in each condition.

Number of Different Words

The number of different words written is represented by the line graph in Figure 2. The mean line shows that the number of different words the 5 students used increased—albeit some only slightly—the mean number of different words written during the first intervention phase compared to baseline.

Student 1. Student 1 wrote a slightly higher number of different words during self-counting compared to baseline but wrote noticeably more when using the synonym list.

Student 2. Student 2 had the greatest increase in the number of different words written during the self-counting condition compared to baseline and a slight increase continued with the synonym list.

Student 3. Student 3 had the greatest increase in number of different words written when using the synonym list, although a slight increase did occur from baseline to the self-counting condition.

Student 4. Student 4 wrote a greater number of different words in self-counting compared to baseline and then had a less noticeable increase when using a synonym list.

Student 5. Student 5 wrote a greater number of different words when self-counting but then wrote slightly fewer after using a synonym list. She still wrote a greater number of different words using the synonym list than in baseline.

Number of Total Words

The number of total words written is represented by the bar graph in Figure 2. All 5 students increased the number of words they wrote in the first intervention phase, self-counting, compared to the baseline condition. Two students either decreased or increased only slightly when given a synonym list to use in the second intervention phase.

Generalization Probes

Quantitative Data. Table 2 shows the number of different words written and number of total words written by each student on each of the three generalization probes. All of the students wrote

Table 2

Word Counts on the Three Generalization Probes

Student	1st Generalization Probe		2nd Generalization Probe		3rd Generalization Probe	
	Different Words	Total Words	Different Words	Total Words	Different Words	Total Words
1	24	35	33	60	54	95
2	52	76	72	112	81	158
3	26	35	53	84	97	155
4	40	64	48	67	98	213
5	69	105	72	126	112	189

a greater number of different words and more total words in each successive generalization probe: from baseline to self-counting and then again from self-counting to self-counting plus synonym list.

Qualitative Data. Table 3 shows the rubric scores given by each evaluator for each of the three generalization probes. Evaluators could give a low score of 3, a high score of 12, or anything within that range depending on their professional opinion of the student's performance in three areas: content, conventions, and organization. Evaluator 1 scored three of the participants' last essay higher than the first and the other two students' last essays lower than the first. Evaluator 2 scored only one student's last essay higher than the first. She scored three of the students' last essays lower than the first and scored one student's first and last essay identically. Except for the identical score, three of the remaining four sets of essays' scores were within one point and the last was within two points.

Scores were very close across the three probe essays, so much so that evaluators were also asked to rank each of the three essays (by student) on a scale of 1 to 3, 1 being the best. Table 4 shows each evaluator's ranking of each student's three essays. Evaluator 1 ranked 4 out of the 5 students' first essays as being the best and the fifth student's last essay as being the best. Evaluator 2's rankings were more varied: she ranked 2 students' first essays as their

Table 3

Rubric Scores by Both Evaluators
for Generalization Probes Essays

Student	1st Generalization Probe		2nd Generalization Probe		3rd Generalization Probe	
	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2	Evaluator 1	Evaluator 2
1	8	6	9	7	9	8
2	10	9	12	8	11	7
3	12	10	11	8	11	9
4	12	10	9	9	11	10
5	9	8	11	7	10	7

Note. Evaluators scored the essays from 1 to 4 points across three areas: content, conventions, and organization. This table reflects the composite score of all areas, so the minimum score can be no less than 3 with a maximum score of 12.

Table 4

Evaluators' Rankings of Each Student's
Three Generalization Probe Essays

Student	Evaluator 1 Rankings			Evaluator 2 Rankings		
	1st Probe	2nd Probe	3rd Probe	1st Probe	2nd Probe	3rd Probe
1	3	2	1	3	2	1
2	1	2	3	3	1	2
3	1	3	2	1	2	3
4	1	3	2	1	3	2
5	1	2	3	3	1	2

Note. Evaluators assigned a rank to each essay on a scale of 1 to 3, 1 being the best.

best, 2 students' second essays as their best, and 1 student's last essay as being her best.

Interobserver Agreement (IOA)

The teacher-researcher served as primary data collector and completed data collection for each student after each writ-

ing session. She removed names from the writing samples and replaced them with initials. The secondary data collector randomly chose 10 out of the 25 writing samples each week to collect for interobserver agreement (IOA) data.

The primary and secondary data collectors obtained IOA data on 40% of the writing samples for each quantitative dependent variable. The mean percentage of IOA for the number of different words written was 98.8% with a range from 93.3% to 100%. IOA was 100% for 32 of the 50 observed writing samples and greater than 95% on all but 1. The mean percentage of IOA for the number of total words written was 99.7% with a range from 95.4% to 100%. IOA was 100% for 45 of the 50 observed writing samples and greater than 95% on the remaining 5.

Procedural Integrity

A procedural checklist was followed by the teacher-researcher during each session to maintain a consistent delivery of instruction. The secondary data collector observed 9 out of the 25 sessions (36%) and completed the same procedural checklist that the teacher researcher used for delivery of instruction, determining that the teacher-researcher implemented instruction with 100% accuracy.

Social Validity

In accordance with the university's internal review board's guidelines for human subjects' protection, the secondary data collector (without the teacher-researcher present) individually asked each student after completion of the last writing session which strategy he or she preferred and why (Internal Review Board Policy Committee, 2006). In addition, she asked questions to identify if the student recognized any change in his or her writing. All answers were recorded in writing by the data collector. Table 5 shows the questions and each student's answers. When asked which intervention was liked better, 3 chose self-counting and 2 chose the synonym list. All 5 students believed their writing had improved since they started the study, that see-

Table 5

Students' Answers to the Social Validity Survey Questions

Question	Student 1	Student 2	Student 3	Student 4	Student 5
What did you like better? Why?	SC You got to write check marks and write on a plastic sheet.	SC You get to count your own writing.	SL It helped me get more different words.	SL It helps you get other words that mean the same.	SC You got to be different from everyone.
If you had to write again, which strategy would you use?	SL	SC	SC	SL	SC
Do you think your writing has improved since you began this project? Why?	Yes. I wrote a lot more words than before.	Yes. The SL helped me get better and better.	Yes. I got more words to write in other stories.	Yes. Because of the synonym list.	Yes. You don't have to look at the list and get time taken away.
Did seeing your graph make a difference when you wrote? How did it affect you?	Yes. My graph almost went off the paper.	Yes . . . see if I went higher on my graph.	Yes. It made feel proud of myself.	Yes. I wrote more different words.	Yes. At first I didn't like [it] but then I got used to it.
Do you use better words now when you write?	Yes	Yes	Yes	Yes	Yes

Note. SC stands for self-counting; SL stands for synonym list plus self-counting.

ing their graphs made a difference, and that they now use better words when writing.

Discussion

This study examined the effects of self-counting and synonym lists on the number of different words written, the number of total words written, and the quality of written expression of 5 high-achieving urban African American first graders.

Number of Total Words Written

All 5 students increased the amount of writing they produced in the intervention phases compared to baseline results. All students averaged a greater number of total words written when self-counting compared to baseline, and all but one increased the average number of total words written then again in the self-counting plus synonym condition. This outcome is important for two reasons. First, decades of research indicate longer writing pieces are strongly associated with higher quality writing (e.g., Hallenbeck, 1996; Hillocks, 1986; Reid & Lienemann, 2006). Second, because a teacher should provide instruction on what students need to know, a student must produce enough writing to allow the teacher to see which skills to target and then in turn produce enough writing to generate opportunities to practice newly learned skills. It should be emphasized that having students write more just for volume's sake is not the goal. A student can just as easily repeat errors while writing a large quantity of text, so frequent teacher monitoring and feedback is essential in the writing process.

Number of Different Words Written

All students averaged a greater number of different words in self-counting than in baseline. All but one student displayed increased averages in the number of different words during self-

counting plus synonym list condition than both the baseline and self-counting conditions. This outcome is important because the use of different words in writing is one indicator of a child's expanding vocabulary. Increasing vocabulary is one of the five essential elements for effective reading instruction as determined by the National Reading Panel (NICHD, 2000). Given the complementary nature of reading and writing (Graham & Perin, 2007), using varied vocabulary in writing can be considered desirable as well. As the teacher-researcher demonstrated in her vocabulary mini-lesson, even children know that better word choices make better writing.

Generalization

The increases in number of different words written and total words written generalized to the probe essays the students wrote periodically throughout the study. All 5 students increased the number of different words written and total words written across the three generalization probes. This is especially exciting because it suggests that the students were able to transfer their skills to untrained experiences. That is, the teacher-researcher observed that students were more attentive in their writing. They made an effort to increase the number of different words written (which makes their writing more interesting) and increased the length of their writing (perhaps by adding more details).

Rankings and Rubric Scores and of Generalization Probe Essays. The rubric scores and rankings of the generalization probe essays underscore the difficulty of evaluating a subjective skill such as written expression. The two evaluators had very different results in their analysis of the students' writings. The first evaluator, the school's reading specialist, tended to score the students' *earlier* writing more favorably. The second evaluator, a second-grade teacher, tended to score the students' *later* writing more favorably (except in one case, Student 1), which was more in line with what the researchers informally predicted; students were expected to write better later in the study after having weeks of exposure to

self-monitoring and the use of synonym lists. Evaluator 1's rankings surprised the researchers but were explained shortly after the study ended when the Evaluator 1 expressed to the teacher-researcher that she disliked "wordy papers" and so scored those lower (personal communication, May 26, 2006). Because the goal of the study was for students to produce more writing, this comment indicates that the later writing of the students was at a disadvantage. Her perspective is also counter to the research across recent decades that indicate longer writing pieces are strongly associated with higher quality writing (e.g., Hallenbeck, 1996; Hillocks, 1986; Reid & Lienemann, 2006).

Further underscoring the difficulty in assessing written expression is the outcome of very little variability in the rubric scores of the generalization probe essays; all but three of the essay scores were within one point of the next chronological essay (the exceptions being two 2-point and one 3-point difference), giving no clear indication that later essays were of a higher quality than earlier ones. This could be an indication that student writing did not improve in such a way as to be noticed when using a rubric. In other words, perhaps rubric scores are not a sensitive enough measure of student writing growth, especially over the course of only weeks. Yet changes did, indeed, occur as demonstrated by the graphs in Figure 2. Another possibility is that even with rubric scoring, written expression assessment is highly subjective in nature. Another indication of the subjective nature of assessing written expression is the comparison of the ranks to the scores: The evaluators matched on only 3 out of 5 students, meaning they gave the highest scores to the essays they identified as the best for only 3 of the students (Students 1, 3, and 4). Conversely, Evaluator 1 gave the lowest scores to the essays she identified as the best for the other 2 students (Student 2 and Student 5).

One way to minimize the subjectivity would have been to use an anchoring procedure to increase the likelihood that the evaluators approached their analysis of the students' writing using the same frame of reference. Anchoring is a training procedure in which, prior to scoring actual student writing samples, evaluators collaboratively rank sample papers in order to more closely

align their reasons for scoring and/or ranking papers. See White (1984) for more details about anchoring.

Because the rubric scoring and ranking outcomes were likely due to both the insensitive nature of rubrics *and* the subjective nature of assessing written expression, a sensitive assessment tool that reduces subjectivity could be useful to teachers, especially if the students themselves could be taught to employ it.

Curriculum-Based Measurement

As demonstrated in this study, CBM is sensitive enough to "pick up" incremental improvements (e.g., the change in the number of different words from baseline to the first intervention and then again to the next intervention) that otherwise might not be noticed. Just as important, CBM can be used to inform instruction. In the study, the researchers waited for one student's data to stabilize in the self-counting condition before providing a synonym list for the next condition and repeated that process across students to demonstrate experimental control. The use of this multiple-baseline design to evaluate the two writing strategies contributes to the growing body of literature that includes single-subject experimental design. However, in the classroom, a teacher does not need to adhere to experimental procedures. Thus, he or she can immediately change instructional strategies as the data dictate. Such data-based instructional decision-making is encouraged by the federal legislation (i.e., No Child Left Behind, 2001 and IDEA, 2004).

Implications for Practice

The results of the study demonstrate that students respond differently to specific instructional strategies, requiring the need for differentiation of the curriculum by the teacher. High-achieving students especially need a challenging curriculum that extends beyond what is appropriate for their typically achieving peers. Challenging students to write a greater number of differ-

ent words—and having them self-monitor by self-counting—
increased the number of different and total words written in the
first 3 minutes of their writing, indicating that these could be
important tools for other young writers. The addition of a syn-
onym list to the self-counting strategy produced increased num-
bers for some participants, indicating another potentially useful
tool for emerging writers. This study also demonstrated that
high-achieving students can generalize the concept of increasing
different and total words to other written expression opportuni-
ties (i.e., the generalization probe essays).

Further, this study demonstrated that students can more
actively participate in their own education by self-monitoring
their writing using curriculum-based measures. This is an easily
implemented strategy and provides the teacher with data with-
out requiring the teacher to do additional work. This availabil-
ity of data is important because teachers can make instructional
decisions that positively impact instruction when data are avail-
able to support or refute classroom practices. For example, after a
certain number of writing sessions in which students self-count,
teachers can determine from the graphs which students are not
succeeding with the self-counting strategy (e.g., those whose
number of different words is stable) and which ones should
continue just self-counting (e.g., those whose number of differ-
ent words is steadily, even if only incrementally, increasing). For
those not succeeding, teachers can relatively quickly try another
strategy (e.g., a synonym list) in response to data indicating a
need for a change. This is an important way in which CBM is
more responsive to teachers' (and therefore, students') needs
than standardized and/or infrequent evaluations. However, an
increase of different and total words may not necessarily produce
better holistic writing. Therefore, teachers should allow students
opportunities to progress through all of the stages of the writing
process and provide meaningful feedback so that students can
implement that feedback to improve their overall writing skills.

Additionally, because this intervention occurred at the
beginning of the students' formal training in written expression,
there is the potential for cumulative benefits as they matriculate

through their schooling. Students who write lengthier stories provide teachers with more opportunities to provide constructive feedback that can be used to inform future written expression opportunities. Therefore, the more mature the students' writing is at an early age, the greater the opportunities for skill development. The students in this study demonstrated the ability to use self-observation and feedback to improve their writing skills.

Limitations and Future Research

An important potential limitation of this study is that the teacher is one of the researchers; however, care was taken to eliminate any bias of the teacher-researcher. IOA data certainly suggest there was no bias in the scoring of weekly writing samples, but there is always the possibility that bias may interfere with the conduct of a study when the teacher is the researcher.

Care should be taken not to generalize the results of the relatively small number of participants in this study to the larger population, especially since they were not randomly selected. A related suggestion for future research is replication with a greater number of participants. Similarly, future research could implement a group design with one control group and two experimental groups with one experimental group self-counting and the other using a synonym list. This would enable the researcher to compare the two interventions by eliminating the possibility of practice effects from self-counting to self-counting plus a synonym list. Similarly, the self-counting condition was really a package that included viewing a graph, and both components may have played key roles in the increase in the number of words written; self-counting may have made students more aware of repetitive words and encouraged ownership of their writing, while the graphing may have provided a visual motivation for them. In future research, one might want to compare the effects of self-counting and graphing independently of each other. A variation would be to teach students to create their own graphs.

Several procedural issues arose during the study that could have affected the results; some changes, therefore, might be desirable. First, students should be directed to cross out mistakes and continue. The lack of this direction resulted in students taking time to erase words. Second, teachers should emphasize that students do not need to have correct spelling in free-writing samples. However, a researcher needs to consider whether misspelled words should be scored as repeated words, especially when they are used in the same context but spelled differently (e.g., "If I could" versus "If I coud").

Summary

Although further research on several aspects of this study is needed, the results show that self-counting and the use of synonym lists are effective forms of differentiated instruction for high-achieving students to increase the length of written expression, as well as to improve the variety of their vocabulary.

References

Artiles, A. J., Trent, S. C., & Palmer, J. (2004). Culturally diverse students in special education: Legacies and prospects. In J. A. Banks & C. M. Banks (Eds.), *Handbook of research on multicultural education* (2nd ed., pp. 716–735). San Francisco: Jossey-Bass.

Blachowicz, C. L. Z., Fisher, P. J., & Watts-Taffe, S. (2005). *Integrated vocabulary instruction: Meeting the needs of diverse learners in grades K–5*. Naperville, IL: North Central Regional Educational Lab. (ERIC Document Reproduction Service No. ED489512)

Bui, Y. N., Schumaker, J. B., & Deshler, D. D. (2006). The effects of a strategic writing program for students with and without learning disabilities in inclusive fifth-grade classes. *Learning Disabilities Research & Practice, 21,* 244–260.

Castellano, J. A., Faivus, A., & White, W. (2003). Servicing the economically disadvantaged in gifted education. In J. A. Castellano

(Ed.), *Special populations in gifted education: Working with diverse gifted learners* (pp. 135–149). Boston: Pearson Education.

De La Paz, S. (1999). Self-regulated strategy instruction in regular education settings: Improving outcomes for students with and without learning disabilities. *Learning Disabilities Research & Practice, 14,* 92–106.

Delisle, J. R., & Galbraith, J. (2002). *When gifted kids don't have all the answers.* Minneapolis, MN: Free Spirit.

Deno, S. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36,* 5–10.

Ford, D. Y., & Harris, J. J. (1993). Educational reform and the focus on gifted African-American students. *Roeper Review, 15,* 200–204.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Monitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71,* 7–21.

Gansle, K. A., Noell, G. H., VanDerHeyden, A. M., Naquin, G. M., & Slider, N. J. (2002). Moving beyond total words written: The reliability, criterion validity, and time cost of alternate measures in curriculum-based measurements in writing. *School Psychology Review, 31,* 477–497.

George, P. S. (2005). A rationale for differentiating instruction in the regular classroom. *Theory Into Practice, 44,* 185–193.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30,* 207–241.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools—A report to Carnegie Corporation of New York.* Washington, DC: Alliance for Excellent Education.

Hallenbeck, M. J. (1996). The cognitive strategy in writing: Welcome relief for adolescents with learning disabilities. *Learning Disabilities Research and Practice, 11,* 107–119.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Brookes.

Hillocks, G. (1986). *Research on written composition: New directions for teaching.* Urbana, IL: National Council of Teachers of English.

Hosp, J. L., & Reschly, D. J. (2004). Disproportionate representation of minority students in special education: Academic, demographic, and economic predictors. *Exceptional Children, 70,* 185–199.

Individuals with Disabilities Education Improvement Act of 2004 (IDEA), Public L. No. 108-446, § 118 Stat. 2647

Internal Review Board Policy Committee. (2006). *Human research protection program policies and procedures: Exempt research.* Retrieved from http://orrp.osu.edu/irb/forms/ExemptResearchSOP.pdf

Jewell, J., & Malecki, C. K. (2005). The utility of CBM written language in dices: An investigation of production-dependent, production-independent, and accurate-production scores. *School Psychology Review, 34,* 27–44.

Kasper-Ferguson, S., & Moxley, R. A. (2002). Developing a writing package with student graphing of fluency. *Education and Treatment of Children, 25,* 249–267.

Kauffman, J. M., Conroy, M., Gardner, R., & Oswald, D. (2008). Cultural sensitivity in the application of behavior principles to education. *Education and Treatment of Children, 31,* 239–262.

Kazdin, A. E. (1982). *Single-case research designs: Methods for clinical and applied settings.* New York: Oxford University Press.

Losen, D. J., & Orfield, G. (2002). *Racial inequality in special education.* Cambridge, MA: Harvard Education Press.

Minner, S., Prater, G., Sullavan, C., & Gwaltney, W. (1989). Informal assessment of written expression. *Teaching Exceptional Children, 21,* 76–79.

Moxley, R. A., & Lutz, P. A. (1995). Self-recorded word counts of freewriting in grades 1–4. *Education & Treatment of Children, 18,* 138–157.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: Government Printing Office.

National Committee on Writing in America's Schools and Colleges (2003). *The neglected "R": The need for a writing revolution.* Retrieved from http://www.writingcommission.org/prod_downloads/writingcom/neglectedr.pdf

National Institute of Child Health and Human Development. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

No Child Left Behind Act, 20 U.S.C. §6301 (2001).

Ohio Gifted Task Force. (2002). *Gifted in the 21st century: A report of findings and recommendations.* Retrieved from http://www.ode.state.oh.us/exceptional_children/gifted_children/final11802.asp

Reid, R., & Lienemann, T. O. (2006). Self-regulated strategy develop-ment for written expression with students with attention deficit/hyperactivity disorder. *Exceptional Children, 73,* 53–68.

Rogers, K. B. (1986). Do the gifted think and learn differently? A review of recent research and its implications for instruction. *Journal for the Education of the Gifted, 10,* 17–39.

Schunk, D. H., & Swartz, C. W. (1993). Writing strategy instruction with gifted students: Effects of goals and feedback on self-efficacy. *Roeper Review, 15,* 225–230.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance pro-file of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86,* 24–53.

Tindal, G., & Hasbrouck, J. (1991). Analyzing student writing to develop instructional strategies. *Learning Disabilities Research & Practice, 6,* 237–245.

Tindal, G., & Parker, R. (1991). Identifying measures for evaluating written expression. *Learning Disabilities Research & Practice, 6,* 211–218.

Tomlinson, C. A. (1999). *The differentiated classroom: Responding to the needs of all learners.* Alexandria, VA: Association for Supervision and Curriculum Development.

U.S. Department of Education, Office of Educational Research and Improvement. (1993). *National excellence: A case for developing America's talent.* Washington, DC: U.S. Government Printing Office.

White, E. M. (1984). Holisticism. *College Composition and Communica-tion, 35,* 400–409.